



(10) **DE 20 2013 012 241 U1** 2016.05.19

(12) **Gebrauchsmusterschrift**

(21) Aktenzeichen: **20 2013 012 241.2**
(22) Anmeldetag: **15.03.2013**
(67) aus Patentanmeldung: **EP 13 79 3997.1**
(47) Eintragungstag: **18.01.2016**
(45) Bekanntmachungstag im Patentblatt: **19.05.2016**

(51) Int Cl.: **C12N 15/11 (2006.01)**

(30) Unionspriorität:

61/652,086	25.05.2012	US
61/716,256	19.10.2012	US
61/757,640	28.01.2013	US
61/765,576	15.02.2013	US

(73) Name und Wohnsitz des Inhabers:

**Charpentier, Emmanuelle, 38124 Braunschweig,
DE; The Regents of The University of California,
Oakland, Calif., US; Universität Wien, Wien, AT**

(74) Name und Wohnsitz des Vertreters:

Uexküll & Stolberg, 22607 Hamburg, DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Modell vorhanden

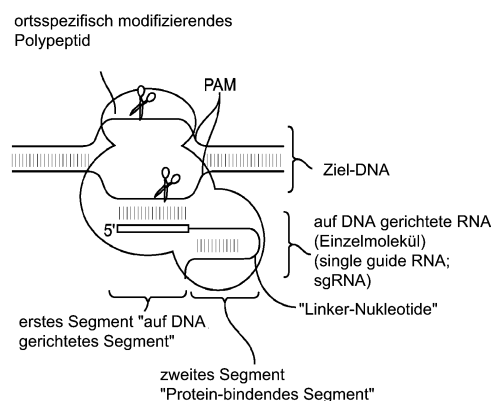
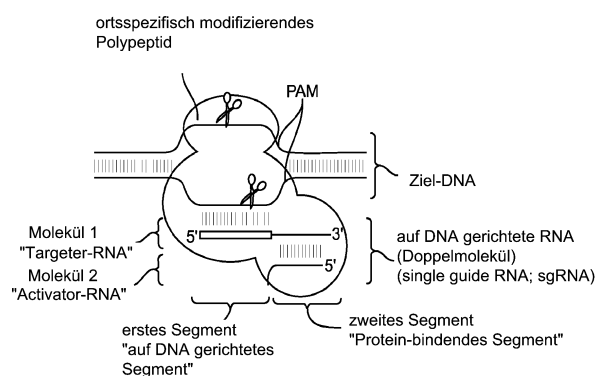
(54) Bezeichnung: **Zusammensetzungen für die durch RNA gesteuerte Modifikation einer Ziel-DNA und für die durch RNA gesteuerte Modulation der Transkription**

(57) Hauptanspruch: Zusammensetzung, die Folgendes umfasst:

(i) ein Cas9-Polypeptid; und
(ii) ein auf DNA-gerichtetes RNA-Einzelmolekül, das Folgendes umfasst:

(a) ein auf DNA-gerichtetes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist, und

(b) ein Protein-bindendes Segment, das mit einem Cas9-Polypeptid interagiert, wobei das Protein-bindende Segment zwei komplementäre Abschnitte von Nukleotiden umfasst, die so hybridisieren, dass ein doppelsträngiger RNA(dsRNA)-Duplex gebildet wird, wobei der dsRNA-Duplex komplementäre Nukleotide einer tracrRNA und einer CRISPR-RNA (crRNA) umfasst, und wobei die zwei komplementären Abschnitte von Nukleotiden kovalent durch zwischenliegende Nukleotide miteinander verbunden sind.



Beschreibung

Querverweise

[0001] Diese Anmeldung beruft sich auf US Provisional Patentanmeldung 61/652,086, die am 25. Mai 2012 eingereicht wurde, 61/716,256, die am 19. Oktober 2012 eingereicht wurde, 61/757,640, die am 28. Januar 2013 eingereicht wurde und 61/765,576, die am 15. Februar 2013 eingereicht wurde, wobei jede dieser Anmeldungen vorliegend vollumfänglich durch Querverweis eingeschlossen ist.

Hinweis zu staatlich unterstützter Forschung

[0002] Diese Erfindung wurde mit Unterstützung der Regierung unter Grant Nr. GM081879, der durch das National Institute of Health vergeben wurde, getätigt. Der Regierung stehen bestimmte Rechte an dieser Erfindung zu.

Einschluss des als Text-Datei bereitgestellten Sequenzprotokolls durch Bezugnahme

[0003] Ein Sequenzprotokoll, das am 13. März 2013 erstellt wurde und eine Größe von 7645 KB aufweist, wird vorliegend als Text-Datei "BERK-187W0-SeqList_ST25.txt" bereitgestellt. Der Inhalt dieser Text-Datei ist vorliegend in seiner Gesamtheit durch Bezugnahme eingeschlossen.

Hintergrund

[0004] Über 60% der Bakterien und 90% der Archaeen besitzen CRISPR(clustered regularly interspaced short palindromic repeats)/CRISPR-assoziierte (Cas) Systeme, um eine Resistenz gegen fremde DNA-Elemente zu etablieren. Das CRISPR-System vom Typ II aus *Streptococcus pyogenes* umfasst lediglich ein einzelnes Gen, das für das Cas9-Protein kodiert, sowie für zwei RNAs – eine reife CRISPR-RNA (crRNA) und eine partiell komplementäre trans-wirkende RNA (tracrRNA) – die für das RNA-gesteuerte Silencing von fremden DNAs erforderlich und ausreichend sind.

[0005] In den vergangenen Jahren haben künstlich erstellte Nuklease-Enzyme, die so erzeugt wurden, dass sie auf spezifische DNA-Sequenzen gerichtet sind, große Beachtung als wirksame Werkzeuge für die genetische Manipulation von Zellen und ganzen Organismen erfahren, wobei diese die gezielte Deletion, das Ersetzen und die Reparatur von Genen sowie auch die Insertion von exogenen Sequenzen (Transgenen) in das Genom ermöglichen. Es haben sich zwei wichtige Technologien für die künstliche Erstellung von ortsspezifischen DNA-Nukleasen herausgebildet, wobei beide auf der Konstruktion von chimären Endonuklease-Enzymen basieren, in denen eine nicht sequenzspezifische DNA-Endonuklease-Domäne an eine künstlich erstellte DNA-bindende Domäne fusioniert ist. Jedoch erfordert die Ausrichtung auf jeden neuen genomischen Locus die Erzeugung eines neuen Nuklease-Enzyms, was dieses Vorgehen sowohl zeit- als auch kostenintensiv gestaltet. Darüber hinaus leiden beide Technologien unter einer bedingten Genauigkeit, was zu unberechenbaren, unspezifischen Wirkungen führen kann.

[0006] Das systematische Abfragen von Genomen und die genetische Reprogrammierung von Zellen umfasst das Targeting von Gruppen von Genen für die Expression oder Repression. Die gegenwärtig am häufigsten verwendete Vorgehensweise für das Targeting von beliebigen Genen für die Regulation ist die Verwendung von RNA-Interferenz (RNAi). Diese Methode hat jedoch Einschränkungen. Beispielsweise kann RNAi signifikante unspezifische Wirkungen und Toxizität aufweisen.

[0007] Es gibt in diesem Bereich der Technik ein Bedürfnis nach einer Technologie, die ein genaues Targeting einer Nuklease-Aktivität (oder anderer Protein-Aktivitäten) an unterschiedlichen Stellen innerhalb einer Ziel-DNA in der Weise ermöglicht, dass die Erzeugung eines neuen Proteins für jede neue Ziel-Sequenz nicht erforderlich ist. Darüber hinaus gibt es ein Bedürfnis im Stand der Technik nach Verfahren zur Kontrolle der Gen-Expression mit minimalen unspezifischen Wirkungen.

ZUSAMMENFASSUNG

[0008] Die vorliegende Offenbarung stellt eine auf DNA gerichtete RNA bereit, die eine Targeting-Sequenz umfasst und gemeinsam mit einem modifizierenden Polypeptid für eine ortsspezifische Modifikation einer Ziel-DNA und/oder eines Polypeptids, das mit der Ziel-DNA assoziiert ist, sorgt. Die vorliegende Offenbarung stellt ferner ortsspezifisch modifizierende Polypeptide bereit. Die vorliegende Offenbarung stellt ferner Verfahren für

die ortsspezifische Modifikation einer Ziel-DNA und/oder eines Polypeptids, das mit der Ziel-DNA assoziiert ist, bereit. Die vorliegende Offenbarung stellt Verfahren zur Modulierung der Transkription einer Ziel-Nukleinsäure in einer Ziel-Zelle bereit, was im Allgemeinen das Inkontaktbringen der Ziel-Nukleinsäure mit einem enzymatisch inaktiven Cas9-Polypeptid und einer auf DNA gerichteten RNA umfasst. Kits und Zusammensetzungen für die Durchführung der Verfahren werden ebenfalls bereitgestellt. Die vorliegende Offenbarung stellt genetisch modifizierte Zellen bereit, die Cas9 erzeugen; und Cas9-transgene, nicht-humane multizelluläre Organismen.

MERKMALE

[0009] Merkmale der vorliegenden Offenbarung umfassen eine auf DNA gerichtete RNA, die Folgendes umfasst: (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert. In einigen Fällen umfasst das erste Segment 8 Nukleotide, die 100% Komplementarität zu einer Sequenz in der Ziel-DNA aufweisen. In einigen Fällen umfasst das zweite Segment eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–682 (z. B. 431–562). In einigen Fällen umfasst das zweite Segment eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 563–682. In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Abschnitten in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346 aufweist.

[0010] Merkmale der vorliegenden Offenbarung umfassen ein DNA-Polynukleotid, das eine Nukleotidsequenz aufweist, die für die auf DNA gerichtete RNA kodiert. In einigen Fällen umfasst ein rekombinanter Expressionsvektor das DNA-Polynukleotid. In einigen Fällen ist die Nukleotidsequenz, welche für die auf DNA gerichtete RNA kodiert, in funktionsfähiger Weise mit einem Promotor verknüpft. In einigen Fällen ist der Promotor ein induzierbarer Promotor. In einigen Fällen umfasst die Nukleotidsequenz, welche für die auf DNA gerichtete RNA kodiert, ferner eine Multiple Cloning Site. Merkmale der vorliegenden Offenbarung umfassen eine in vitro genetisch modifizierte Wirtszelle, die das DNA-Polynukleotid umfasst.

[0011] Merkmale der vorliegenden Erfindung umfassen einen rekombinanten Expressionsvektor, der Folgendes umfasst: (i) eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) eine Nukleotidsequenz, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, die Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf die DNA gerichtete RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität zeigt, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0012] Merkmale der vorliegenden Offenbarung umfassen einen rekombinanten Expressionsvektor, der Folgendes umfasst (i) eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) eine Nukleotidsequenz, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0013] Merkmale der vorliegenden Offenbarung umfassen eine Variante des ortsspezifisch modulierenden Polypeptids, die Folgendes umfasst: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) einen Aktivitätsteil, der eine verringerte ortsspezifische enzymatische Aktivität zeigt, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen umfasst die Variante des ortsspezifisch modifizierenden Polypeptids eine H840A-Mutation der *S. pyogenes*-Sequenz von SEQ ID NO: 8 oder die entsprechende Mutation in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen umfasst die Variante des ortsspezifisch modifizierenden Polypeptids eine D10A-Mutation der *S. pyogenes*-Sequenz von SEQ ID NO: 8 oder die entsprechende

Mutation in einer beliebigen der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen umfasst die Variante des ortsspezifisch modifizierenden Polypeptids sowohl (i) eine D10A-Mutation der *S. pyogenes* Sequenz von SEQ ID NO: 8 oder die entsprechende Mutation in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346; und (ii) eine H840A-Mutation der *S. pyogenes*-Sequenz von SEQ ID NO: 8 oder die entsprechende Mutation in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346.

[0014] Merkmale der vorliegenden Offenbarung umfassen ein chimäres ortsspezifisch modifizierendes Polypeptid, das Folgendes umfasst: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität zeigt, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen umfasst das chimäre ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 aufweist oder zu den entsprechenden Teilen in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen umfasst die auf DNA gerichtete RNA ferner eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–682 (z. B. SEQ ID NOs: 563–682). In einigen Fällen umfasst die auf DNA gerichtete RNA ferner eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–562. In einigen Fällen modifiziert die enzymatische Aktivität des chimären ortsspezifisch modifizierenden Polypeptids die Ziel-DNA. In einigen Fällen ist die enzymatische Aktivität des chimären ortsspezifisch modifizierenden Polypeptids eine Nuklease-Aktivität, eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine DNA-Reparatur-Aktivität, eine DNA-Schädigungs-Aktivität, eine Deaminierungs-Aktivität, eine Dismutase-Aktivität, eine Alkylierungs-Aktivität, eine Depurinierungs-Aktivität, eine Oxidations-Aktivität, eine Pyrimidin-Dimer-bildende Aktivität, eine Integrase-Aktivität, eine Transposase-Aktivität, eine Rekombinase-Aktivität, eine Polymerase-Aktivität, eine Ligase-Aktivität, eine Helicase-Aktivität, eine Photolyase-Aktivität oder eine Glykosylase-Aktivität. In einigen Fällen ist die enzymatische Aktivität des chimären ortsspezifisch modifizierenden Polypeptids eine Nuklease-Aktivität. In einigen Fällen führt die Nuklease-Aktivität einen Doppelstrang-Bruch in die Ziel-DNA ein. In einigen Fällen modifiziert die enzymatische Aktivität des chimären ortsspezifisch modifizierenden Polypeptids ein Ziel-Polypeptid, das mit einer Ziel-DNA assoziiert ist. In einigen Fällen ist die enzymatische Aktivität des chimären ortsspezifisch modifizierenden Polypeptids eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität.

[0015] Merkmale der vorliegenden Offenbarung umfassen ein Polynukleotid, das eine Nukleotidsequenz umfasst, welche für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert. In einigen Fällen ist das Polynukleotid ein RNA-Polynukleotid. In einigen Fällen ist das Polynukleotid ein DNA-Polynukleotid. Merkmale der vorliegenden Offenbarung umfassen einen rekombinanten Expressionsvektor, der das Polynukleotid umfasst. In einigen Fällen ist das Polynukleotid in funktionsfähiger Weise mit einem Promotor verknüpft. In einigen Fällen ist der Promotor ein induzierbarer Promotor. Merkmale der vorliegenden Offenbarung umfassen eine in vitro genetisch modifizierte Wirtszelle, die das Polynukleotid umfasst.

[0016] Merkmale der vorliegenden Offenbarung umfassen ein chimäres ortsspezifisch modifizierendes Polynukleotid, das Folgende umfasst: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist, und (ii) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen erhöht der Aktivitätsteil die Transkription innerhalb der Ziel-DNA. In einigen Fällen verringert der Aktivitätsteil die Transkription innerhalb der Ziel-DNA.

[0017] Merkmale der vorliegenden Offenbarung umfassen eine genetisch modifizierte Zelle, die ein rekombinantes ortsspezifisch modifizierendes Polypeptid umfasst, welches einen RNA-bindenden Teil umfasst, der mit einer auf DNA gerichteten RNA interagiert; und einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität zeigt, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten

Aminosäuresequenz von Cas9/Csn1 aufweist oder zu den entsprechenden Teilen in den Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen ist die Zelle ausgewählt aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer bakteriellen Zelle, einer eukaryontischen Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzenzelle, einer Algenzelle, einer Tierzelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fischzelle, Froschzelle, einer Vogelzelle, einer Säugetierzelle, einer Schweinezelle, einer Kuhzelle, einer Ziegenzelle, einer Schafzelle, einer Nagetierzelle, einer Rattenzelle, einer Mauszelle, einer Zelle aus einem nicht-menschlichen Primaten und einer menschlichen Zelle.

[0018] Merkmale der vorliegenden Offenbarung umfassen einen transgenen nicht-menschlichen Organismus, dessen Genom ein Transgen umfasst, welches eine Nukleotidsequenz umfasst, die für ein rekombinantes ortsspezifisch modifizierendes Polypeptid kodiert, welches Folgendes umfasst: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert; und (ii) einem Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität zeigt, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz mit mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in den Aminosäuresequenzen von SEQ ID NOs: 1–246 und 795–1346. In einigen Fällen ist der Organismus ausgewählt aus der Gruppe bestehend aus einem Archaeon, einem Bakterium, einem eukaryontischen einzelligen Organismus, einer Alge, einer Pflanze, einem Tier, einem Invertebraten, einer Fliege, einem Wurm, einem Nesseltier, einem Vertebraten, einem Fisch, einem Frosch, einem Vogel, einem Säugetier, einem Huftier, einem Nagetier, einer Ratte, einer Maus und einem nicht-menschlichen Primaten.

[0019] Merkmale der vorliegenden Offenbarung umfassen eine Zusammensetzung, die Folgendes umfasst: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbiges kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) ein RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen umfasst das erste Segment der auf DNA gerichteten RNA 8 Nukleotide, die mindestens 100% Komplementarität zu einer Sequenz in der Ziel-DNA aufweisen. In einigen Fällen umfasst das zweite Segment der auf DNA gerichteten RNA eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–682 (z. B. SEQ ID NOs: 563–682). In einigen Fällen umfasst das zweite Segment der auf DNA gerichteten RNA eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–562. In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz mit mindestens etwa 75% Sequenzidentität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen modifiziert die enzymatische Aktivität die Ziel-DNA. In einigen Fällen ist die enzymatische Aktivität eine Nuklease-Aktivität, eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine DNA-Reparatur-Aktivität, eine DNA-Schädigungs-Aktivität, eine Deaminierungs-Aktivität, eine Dismutase-Aktivität, eine Alkylierungs-Aktivität, eine Depurinierungs-Aktivität, eine Oxidations-Aktivität, eine Pyrimidin-Dimer-bildende Aktivität, eine Integrase-Aktivität, eine Transposase-Aktivität, eine Rekombinase-Aktivität, eine Polymerase-Aktivität, eine Ligase-Aktivität, eine Helicase-Aktivität, eine Photolyase-Aktivität oder eine Glykosylase-Aktivität. In einigen Fällen ist die enzymatische Aktivität eine Nuklease-Aktivität. In einigen Fällen führt die Nuklease-Aktivität einen Doppelstrang-Bruch in die Ziel-DNA ein. In einigen Fällen modifiziert die enzymatische Aktivität ein Ziel-Polypeptid, das mit einer Ziel-DNA assoziiert ist. In einigen Fällen ist die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität. In einigen Fällen ist das Ziel-Polypeptid ein Histon und die enzymatische Aktivität ist eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, oder eine deubiquitinierende Aktivität. In einigen Fällen ist die auf DNA gerichtete RNA ein auf DNA gerichtetes RNA-Doppelmolekül, und die Zusammensetzung umfasst sowohl eine Targeter-RNA und eine Aktivator-RNA, wobei die Duplex-bildenden Segmente derselben komplementär

zueinander sind und hybridisieren, wodurch das zweite Segment der auf die DNA gerichteten RNA gebildet wird. In einigen Fällen umfasst das Duplex-bildenden Segment der Aktivator-RNA eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–682.

[0020] Merkmale der vorliegenden Offenbarung umfassen eine Zusammensetzung die Folgendes umfasst: (i) eine auf DNA gerichtete RNA der vorliegenden Offenbarung oder ein DNA-Polynukleotid, welches für selbiges kodiert; und (ii) einen Puffer zur Stabilisierung von Nukleinsäuren. Merkmale der vorliegenden Offenbarung umfassen eine Zusammensetzung, die Folgendes umfasst: (i) ein ortsspezifisch modifizierendes Polypeptid der vorliegenden Offenbarung oder ein Polynukleotid, welches für selbiges kodiert; und (ii) einen Puffer zur Stabilisierung von Nukleinsäuren und/oder Proteinen. Merkmale der vorliegenden Offenbarung umfassen eine Zusammensetzung, die Folgendes umfasst: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, welches für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid, oder ein Polynukleotid, welches für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen erhöht der Aktivitätsteil die Transkription innerhalb der Ziel-DNA. In einigen Fällen verringert der Aktivitätsteil die Transkription innerhalb der Ziel-DNA. Merkmale der vorliegenden Offenbarung umfassen eine Zusammensetzung, die Folgendes umfasst: (i) ein ortsspezifisch modifizierendes Polypeptid oder ein Polynukleotid, welches für selbiges kodiert; und (ii) einen Puffer zur Stabilisierung von Nukleinsäuren und/oder Proteinen.

[0021] Merkmale der vorliegenden Offenbarung umfassen ein Verfahren zur ortsspezifischen Modifikation einer Ziel-DNA, wobei das Verfahren Folgendes umfasst: Inkontaktbringen der Ziel-DNA mit: (i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, welches für selbiges kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, welches für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität zeigt. In einigen Fällen liegt die Ziel-DNA extrachromosomal vor. In einigen Fällen umfasst die Ziel-DNA eine PAM-Sequenz des komplementären Stranges, die wie folgt 5'-CCY-3' lautet, wobei Y ein beliebiges DNA-Nukleotid ist, und Y unmittelbar 5' von der Ziel-Sequenz des komplementären Stranges der Ziel-DNA liegt. In einigen Fällen ist die Ziel-DNA Teil eines in vitro vorliegenden Chromosoms. In einigen Fällen ist die Ziel-DNA Teil eines in vivo vorliegenden Chromosoms. In einigen Fällen ist die Ziel-DNA Teil eines Chromosoms in einer Zelle. In einigen Fällen ist die Zelle ausgewählt aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer bakteriellen Zelle, einer eukaryontischen Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzenzelle, einer Algenzelle, einer Tierzelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fischzelle, Froschzelle, einer Vogelzelle, einer Säugetierzelle, einer Schweinezelle, einer Kuhzelle, einer Ziegenzelle, einer Schafzelle, einer Nagetierzelle, einer Rattenzelle, einer Mauszelle, einer Zelle aus einem nicht-menschlichen Primaten und einer menschlichen Zelle. In einigen Fällen umfasst die auf DNA gerichtete RNA eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Aminosäuresequenzen von SEQ ID NOs: 431–682 (z. B. SEQ ID NOs: 563–682). In einigen Fällen umfasst die auf DNA gerichtete RNA eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Aminosäuresequenzen von SEQ ID NOs: 431–562. In einigen Fällen umfasst das DNA-modifizierende Polypeptid eine Aminosäuresequenz mit mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen modifiziert die enzymatische Aktivität die Ziel-DNA. In einigen Fällen ist die enzymatische Aktivität eine Nuklease-Aktivität, eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine DNA-Reparatur-Aktivität, eine DNA-Schädigungs-Aktivität, eine Deaminierungs-Aktivität, eine Dismutase-Aktivität, eine Alkylierungs-Aktivität, eine Depurinierungs-Aktivität, eine Oxidations-Aktivität, eine Pyrimidin-Dimer-bildende Aktivität, eine Integrase-Aktivität, eine Transposase-Aktivität, eine Rekombinase-Aktivität, eine Polymerase-Aktivität, eine Ligase-Aktivität, eine Helicase-Aktivität, eine Photolyase-Aktivität oder eine Glykosylase-Aktivität. In einigen Fällen ist die DNA-modifizierende enzymatische Aktivität eine Nuklease-Aktivität. In einigen Fällen führt die Nuklease-Aktivität einen Doppelstrang-Bruch in die Ziel-

DNA ein. In einigen Fällen erfolgt das Inkontaktbringen unter Bedingungen, die ein nicht-homologes End-Joining oder eine Homologie-gesteuerte Reparatur ermöglichen. In einigen Fällen umfasst das Verfahren das Inkontaktbringen der Ziel-DNA mit einem Donor-Polynukleotid, wobei das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA integriert. In einigen Fällen umfasst das Verfahren nicht das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid, wobei die Ziel-DNA so modifiziert wird, dass Nukleotide innerhalb der Ziel-DNA deletiert werden. In einigen Fällen modifiziert die enzymatische Aktivität ein Ziel-Polypeptid, das mit einer Ziel-DNA assoziiert ist. In einigen Fällen ist die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität. In einigen Fällen ist das Ziel-Polypeptid ein Histon und die enzymatische Aktivität ist eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, oder eine deubiquitinierende Aktivität. In einigen Fällen umfasst der Komplex ferner eine Aktivator-RNA. In einigen Fällen umfasst die Aktivator-RNA eine Nukleotidsequenz mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der Nukleotidsequenzen von SEQ ID NOs: 431–682.

[0022] Merkmale der vorliegenden Erfindung umfassen ein Verfahren zum Modulieren einer ortsspezifischen Transkription innerhalb einer Ziel-DNA, wobei das Verfahren das Inkontaktbringen der Ziel-DNA mit Folgendem umfasst: (i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) ein Aktivitätsteil, der die Transkription moduliert, wobei das Inkontaktbringen zu einer modulierten Transkription innerhalb der Ziel-DNA führt. In einigen Fällen wird die Transkription innerhalb der Ziel-DNA erhöht. In einigen Fällen wird die Transkription innerhalb der Ziel-DNA verringert.

[0023] Merkmale der vorliegenden Erfindung umfassen ein Verfahren zur ortsspezifischen Modifikation einer Ziel-DNA, wobei das Verfahren das Inkontaktbringen der Ziel-DNA mit Folgendem umfasst: (i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, welches für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) ein Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert. In einigen Fällen erhöht das ortsspezifisch modifizierende Polypeptid die Transkription innerhalb der Ziel-DNA. In einigen Fällen verringert das ortsspezifisch modifizierende Polypeptid die Transkription innerhalb der Ziel-DNA.

[0024] Merkmale der vorliegenden Offenbarung umfassen ein Verfahren zur Förderung der ortsspezifischen Spaltung und Modifikation einer Ziel-DNA in einer Zelle, wobei das Verfahren Schritte umfasst, bei dem man Folgendes in die Zelle einbringt: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) ein ortsspezifisch modifizierendes Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine Nuklease-Aktivität zeigt, die einen Doppelstrang-Bruch in der Ziel-DNA erzeugt; wobei die Stelle des Doppelstrang-Bruchs durch die auf DNA gerichtete RNA bestimmt wird, das Inkontaktbringen unter Bedingungen erfolgt, die ein nicht-homologes End Joining oder eine Homologie-gesteuerte Reparatur ermöglichen und die Ziel-DNA gespalten und wieder zusammengefügt wird, sodass eine modifizierte DNA-Sequenz erzeugt wird. In einigen Fällen umfasst das Verfahren ferner das Inkontaktbringen der Ziel-DNA mit einem Donor-Polynukleotid, wobei das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA integriert. In einigen Fällen umfasst das Verfahren nicht das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid, wobei die Ziel-DNA so modifiziert wird, dass Nukleotide innerhalb der Ziel-DNA deletiert werden. In einigen Fällen ist die

Zelle ausgewählt aus der Gruppe bestehend aus einer Archaeen-Zelle, einer bakteriellen Zelle, einer eukaryontischen Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzenzelle, einer Algenzelle, einer Tierzelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fischzelle, Froschzelle, einer Vogelzelle, einer Säugetierzelle, einer Schweinezelle, einer Kuhzelle, einer Ziegenzelle, einer Schafzelle, einer Nagetierzelle, einer Rattenzelle, einer Mauszelle, einer Zelle aus einem nicht-menschlichen Primaten und einer menschlichen Zelle. In einigen Fällen liegt die Zelle in vitro vor, in einigen Fällen liegt die Zelle in vivo vor.

[0025] Merkmale der vorliegenden Offenbarung umfassen ein Verfahren zur Herstellung einer genetisch modifizierten Zelle in einem Subjekt, bei dem man: (I) Folgendes in eine Zelle einbringt: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) ein ortsspezifisch modifizierendes Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine Nuklease-Aktivität zeigt, welche einen Doppelstrang-Bruch in der Ziel-DNA erzeugt; wobei die Stelle des Doppelstrang-Bruchs durch die auf DNA gerichtete RNA bestimmt wird, das Inkontaktbringen unter Bedingungen erfolgt, die ein nicht-homologes End Joining oder eine Homologie-gesteuerte Reparatur ermöglichen und die Ziel-DNA gespalten und wieder zusammengefügt wird, sodass eine modifizierte DNA-Sequenz erzeugt wird; wodurch die genetisch modifizierte Zelle erzeugt wird; und (II) die genetisch modifizierte Zelle in das Subjekt transplantiert. In einigen Fällen umfasst das Verfahren ferner das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid, wobei das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA integriert. In einigen Fällen umfasst das Verfahren nicht das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid, wobei die Ziel-DNA so modifiziert wird, dass Nukleotide innerhalb der Ziel-DNA deletiert werden. In einigen Fällen ist die Zelle ausgewählt aus der Gruppe bestehend aus einer Archaeen-Zelle, einer bakteriellen Zelle, einer eukaryontischen Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzenzelle, einer Algenzelle, einer Tierzelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fischzelle, einer Amphibien-Zelle, einer Vogelzelle, einer Säugetierzelle, einer Huftier-Zelle, einer Nagetierzelle, einer Zelle aus einem nicht-menschlichen Primaten und einer menschlichen Zelle.

[0026] Merkmale der vorliegenden Erfindung umfassen ein Verfahren zur Modifizierung einer Ziel-DNA in einer genetisch modifizierten Zelle, die eine Nukleotidsequenz umfasst, welche für ein exogenes ortsspezifisch modifizierendes Polypeptid kodiert, wobei das Verfahren Schritte umfasst, bei denen man eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, in eine genetisch modifizierte Zelle einbringt, wobei: (i) die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine Nuklease-Aktivität aufweist. In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Aminosäuresequenz von Cas9/Csn1 aufweist oder zu den entsprechenden Teilen in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346. In einigen Fällen ist die Zelle ausgewählt aus der Gruppe bestehend aus einer Archaeen-Zelle, einer bakteriellen Zelle, einer eukaryontischen Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzenzelle, einer Algenzelle, einer Tierzelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fischzelle, einer Amphibien-Zelle, einer Vogelzelle, einer Säugetierzelle, einer Huftier-Zelle, einer Nagetierzelle, einer Zelle aus einem nicht-menschlichen Primaten und einer menschlichen Zelle. In einigen Fällen liegt die Zelle in vivo vor. In einigen Fällen liegt die Zelle in vitro vor. In einigen Fällen steht die Expression des ortsspezifisch modifizierenden Polypeptids unter der Kontrolle eines induzierbaren Promotors. In einigen Fällen steht die Expression des ortsspezifisch modifizierenden Polypeptids unter der Kontrolle eines Zelltyp-spezifischen Promotors.

[0027] Merkmale der vorliegenden Erfindung umfassen ein Kit, das Folgendes umfasst: eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, und ein Reagenz für die Rekonstituierung und/oder Verdünnung. In einigen Fällen umfasst das Kit ferner ein Reagenz, das ausgewählt ist aus der Gruppe bestehend aus: einem Puffer für das Einbringen der auf DNA gerichteten RNA in Zellen, einem Waschpuffer, einem Kontroll-Reagenz, einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid, einem Reagenz für die Transkription der auf DNA gerichteten RNA ausgehend von DNA und Kombinationen derselben.

[0028] Merkmale der vorliegenden Erfindung umfassen ein Kit, das Folgendes umfasst: ein ortsspezifisch modifizierendes Polypeptid der vorliegenden Erfindung oder ein Polynukleotid, das für selbiges kodiert; und ein Reagenz für die Rekonstituierung und/oder Verdünnung. In einigen Fällen umfasst das Kit ferner ein Reagenz, das ausgewählt ist aus der Gruppe bestehend aus: einem Puffer zur Einbringung des ortsspezifisch modifizierenden Polypeptids in Zellen, einem Waschpuffer, einem Kontroll-Reagenz, einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid, einem Reagenz für die in vitro-Herstellung des ortsspezifisch modifizierenden Polypeptids ausgehend von DNA und Kombinationen davon.

[0029] Merkmale der vorliegenden Erfindung umfassen ein Kit, das Folgendes umfasst: ein ortsspezifisch modifizierendes Polypeptid der vorliegenden Offenbarung oder ein Polynukleotid, das für selbiges kodiert; und ein Reagenz für die Rekonstituierung und/oder Verdünnung. Merkmale der vorliegenden Offenbarung umfassen ein Kit, das Folgendes umfasst: eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0030] Merkmale der vorliegenden Erfindung umfassen ein Kit, das Folgendes umfasst: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, umfassend: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0031] Merkmale der vorliegenden Offenbarung umfassen ein Kit, das Folgendes umfasst: (i) einen der oben beschriebenen rekombinanten Expressionsvektoren; und (ii) ein Reagenz für die Rekonstituierung und/oder Verdünnung. Merkmale der vorliegenden Offenbarung umfassen ein Kit, das Folgendes umfasst: (i) einen der oben beschriebenen rekombinanten Expressionsvektoren; und (ii) einen rekombinanten Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität zeigt, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. Merkmale der vorliegenden Offenbarung umfassen ein Kit, das Folgendes umfasst: (i) einen der oben beschriebenen rekombinanten Expressionsvektoren; und (ii) einen rekombinanten Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0032] Merkmale der vorliegenden Offenbarung umfassen ein Kit für das Targeting einer Ziel-DNA, das Folgendes umfasst: zwei oder mehr auf DNA gerichtete RNAs oder DNA-Polynukleotide, die für selbige kodieren, wobei das erste Segment von mindestens einer der zwei oder mehreren auf DNA gerichteten RNAs sich durch mindestens ein Nukleotid von dem ersten Segment von mindestens einer anderen der zwei oder mehreren auf DNA gerichteten RNAs unterscheidet.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0033] Fig. 1A–B zeigen eine schematische Darstellung von zwei beispielhaften auf DNA gerichteten RNAs, wobei jede mit einem ortsspezifisch modifizierenden Polypeptid und mit einer Ziel-DNA assoziiert ist.

[0034] Fig. 2 zeigt die Veränderung der Ziel-DNA durch Brüche in der doppelsträngigen DNA, die unter Verwendung eines ortsspezifisch modifizierenden Cas9/Csn1-Polypeptids und einer auf DNA gerichteten RNA eingebracht wurden.

[0035] Fig. 3A–B zeigen die Aminosäuresequenz eines Cas9/Csn1-Proteins aus *Streptococcus pyogenes* (SEQ ID NO: 8). Cas9 weist Domänen auf, die sowohl zur Endonuklease HNH als auch zur Endonuklease RuvC homolog sind. (A) Die Motive 1–4 sind mit Oberstrichen dargestellt. (B) Die Domänen 1 und 2 sind mit Oberstrichen dargestellt.

[0036] Fig. 4A–B zeigen die prozentuale Identität zwischen den Cas9/Csn1-Proteinen aus verschiedenen Spezies. (A) Sequenzidentität relativ zu *Streptococcus pyogenes*. Beispielsweise entspricht Domäne 1 den Aminosäuren 7–166 und Domäne 2 den Aminosäuren 731–1003 von Cas9/Csn1 aus *Streptococcus pyogenes*, wie es in **Fig. 3B** dargestellt ist. (B) Sequenzidentität relativ zu *Neisseria meningitidis*. Beispielsweise entspricht Domäne 1 den Aminosäuren 13–139 und Domäne 2 den Aminosäuren 475–750 von Cas9/Csn1 aus *Neisseria meningitidis* (SEQ ID NO: 79).

[0037] Fig. 5 zeigt ein multiples Sequenz-Alignment der Motive 1–4 aus Cas9/Csn1-Proteinen von unterschiedlichen Spezies, die aus der phylogenetischen Tabelle von **Fig. 32** ausgewählt sind (siehe **Fig. 32**, **Fig. 3A** und Tabelle 1) (*Streptococcus pyogenes* (SEQ ID NO: 8) *Legionella pneumophila* (SEQ ID NO: 17), *Gamma proteobacterium* (SEQ ID NO: 107), *Listeria innocua* (SEQ ID NO: 3), *Lactobacillus gasseri* (SEQ ID NO: 152), *Eubacterium rectale* (SEQ ID NO: 99), *Staphylococcus lugdunensis* (SEQ ID NO: 185), *Mycoplasma synoviae* (SEQ ID NO: 22), *Mycoplasma mobile* (SEQ ID NO: 16), *Wolinella succinogenes* (SEQ ID NO: 10), *Flavobacterium columnare* (SEQ ID NO: 235), *Fibrobacter succinogenes* (SEQ ID NO: 121), *Bacteroides fragilis* (SEQ ID NO: 21), *Acidothermus cellulolyticus* (SEQ ID NO: 42), und *Bifidobacterium dentium* (SEQ ID NO: 131).

[0038] Fig. 6A–B stellen Alignments der natürlich vorkommenden tracrRNA("Aktivator-RNA")-Sequenzen aus verschiedenen Spezies bereit (*L. innocua* (SEQ ID NO: 268); *S. pyogenes* (SEQ ID NO: 267); *S. mutans* (SEQ ID NO: 269); *S. thermophilus1* (SEQ ID NO: 270); *M. mobile* (SEQ ID NO: 274); *N. meningitidis* (SEQ ID NO: 272); *P. multocida* (SEQ ID NO: 273); *S. thermophilus2* (SEQ ID NO: 271); und *S. pyogenes* (SEQ ID NO: 267). (A) Multiples Sequenz-Alignment von ausgewählten tracrRNA-Orthologen (AlignX, VektorNTI-Paket, Invitrogen), die mit CRISPR/Cas-Loci ähnlichen Aufbaus und hochgradig ähnlichen Cas9/Csn1-Sequenzen assoziiert sind. Die schwarzen Kästchen repräsentieren gemeinsame Nukleotide (B) Multiples Sequenz-Alignment von ausgewählten tracrRNA-Orthologen (AlignX, VektorNTI-Paket, Invitrogen), die mit CRISPR/Cas-Loci unterschiedlichen Aufbaus und nicht eng verwandten Cas9/Csn1-Sequenzen assoziiert sind. Zu beachten ist die Sequenzähnlichkeit der tracrRNA-Orthologe von *N. meningitidis* und *P. multocida*. Die schwarzen Kästchen repräsentieren gemeinsame Nukleotide. Für weitere exemplarische Sequenzen von Aktivator-RNAs siehe SEQ ID NOs: 431–562.

[0039] Fig. 7A–B stellen Alignments von natürlich vorkommenden, Duplex-bildenden Segmenten von crRNA("Targeter-RNA")-Sequenzen aus unterschiedlichen Spezies bereit (*L. innocua* (SEQ ID NO: //); *S. pyogenes* (SEQ ID NO: //); *S. mutans* (SEQ ID NO: //); *S. thermophilus1* (SEQ ID NO: //); *C. jejuni* (SEQ ID NO: //); *S. pyogenes* (SEQ ID NO: //); *F. novicida* (SEQ ID NO: //); *M. mobile* (SEQ ID NO: //); *N. meningitidis* (SEQ ID NO: //); *P. multocida* (SEQ ID NO: //); und *S. thermophilus2* (SEQ ID NO: //). (A) Multiple Sequenz-Alignments von beispielhaften Duplex-bildenden Segmenten von Targeter-RNA-Sequenzen (AlignX, VektorNTI-Paket, Invitrogen), die mit Loci ähnlichem Aufbau und hochgradig ähnlichen Cas9/Csn1-Sequenzen assoziiert sind. (B) Multiple Sequenz-Alignments von beispielhaften, Duplex-bildenden Segmenten von Targeter-RNA-Sequenzen (AlignX, VektorNTI-Paket, Invitrogen), die mit Loci unterschiedlicher Aufbaus und abweichenden Cas9/Csn1-Sequenzen assoziiert sind. Die schwarzen Boxen repräsentieren gemeinsame Nukleotide. Für weitere beispielhafte Duplex-bildende Segmente von Targeter-RNA-Sequenzen siehe SEQ ID NOs: 563–679.

[0040] Fig. 8 zeigt schematisch die Hybridisierung von natürlich vorkommenden Duplex-bildenden Segmenten der crRNA("Targeter-RNA") mit den Duplex-bildenden Segmenten des entsprechenden tracrRNA-Orthologe("Aktivator-RNA"). Obere Sequenz, Targeter-RNA; untere Sequenz, Duplex-bildendes Segment der entsprechenden Aktivator-RNA. Die CRISPR-Loci gehören zum Typ II (Nmeni/CASS4) CRISPR/Cas-System. Die Nomenklatur folgt der CRISPR-Datenbank (CRISPR DB). *S. pyogenes* (SEQ ID NO: // und //); *S. mutans* (SEQ ID NO: // und //); *S. thermophilus1* (SEQ ID NO: // und //); *S. thermophilus2* (SEQ ID NO: // und //); *L. innocua* (SEQ ID NO: // und //); *T. denticola* (SEQ ID NO: // und //); *N. meningitidis* (SEQ ID NO: // und //); *S. gordonii* (SEQ ID NO: // und //); *B. bifidum* (SEQ ID NO: // und //); *L. salivarius* (SEQ ID NO: // und //); *F. tularensis* (SEQ ID NO: // und //); und *L. pneumophila* (SEQ ID NO: // und //). Es ist zu beachten, dass einige Spezies jeweils zwei Typ II-CRISPR-Loci umfassen. Für weitere beispielhafte Aktivator-RNA-Sequenzen siehe SEQ ID NOs: 431–562. Für weitere beispielhafte Duplex-bildende Segmente von Targeter-RNA-Sequenzen, siehe SEQ ID NOs: 563–679.

[0041] Fig. 9 zeigt beispielhafte tracrRNA(Aktivator-RNA)- und crRNA(Targeter RNA)-Sequenzen aus zwei Spezies. Es besteht ein gewisser Grad an Austauschbarkeit; beispielsweise ist das Cas9/Csn1-Protein von *S. pyogenes* mit der tracrRNA und der crRNA von *L. innocua* funktionsfähig. (|) weist auf eine kanonische Watson-Crick-Basenpaarung hin, während (•) auf ein G-U-Wobble-Basenpaar verweist. "Variable 20nt" oder "20 nt" stellt das auf die DNA gerichtete Segment dar, welches komplementär zu einer Ziel-DNA ist (diese Region kann bis zu etwa 100nt lang sein). Ebenfalls gezeigt ist der Aufbau eines auf DNA gerichteten RNA-Einzelmoleküls, das Merkmale der Targeter-RNA und der Aktivator-RNA einschließt. (Cas9/Csn1-Protein-Sequenzen von einer Vielzahl von verschiedenen Spezies sind in Fig. 3 dargestellt und werden als SEQ ID NOs: 1–256 und 795–1346 bereitgestellt). *Streptococcus pyogenes*: von oben nach unten: (SEQ ID NO: //, //, //); *Listeria innocua*: von oben nach unten: (SEQ ID NO: //, //, //). Die bereitgestellten Sequenzen sind nicht einschränkende Beispiele und sollen lediglich veranschaulichen, wie auf DNA gerichtete RNA-Einzelmoleküle und auf DNA gerichtete RNA-Doppelmoleküle auf Basis der natürlich vorkommenden Sequenzen von einer großen Vielzahl von Spezies erstellt werden können. Verschiedene Beispiele für geeignete Sequenzen aus einer Vielzahl von Spezies werden wie folgt bereitgestellt: (Cas9-Protein: SEQ ID NOs: 1–259; tracrRNAs: SEQ ID NOs: 431–562 oder die komplementäre Sequenzen davon; crRNAs: SEQ ID NOs: 563–679 oder die komplementären Sequenzen davon; und beispielhafte auf DNA gerichtete RNA-Einzelmoleküle: SEQ ID NOs: 680–682).

[0042] Fig. 10A–Fig. 10E zeigen, dass Cas9 eine DNA-Endonuklease ist, die von zwei RNA-Molekülen geleitet wird. Figur E (von oben nach unten, SEQ ID NOs: 278–280, und //).

[0043] Fig. 11A–B zeigen, dass Cas9 zwei Nuklease-Domänen verwendet, um die beiden Stränge in der Ziel-DNA zu spalten.

[0044] Fig. 12A–E veranschaulichen, dass die Cas9-katalysierte Spaltung der Ziel-DNA eine aktivierende Domäne in der tracrRNA erfordert und von einer Seed-Sequenz in der crRNA gesteuert wird. Fig. 12C (von oben nach unten: SEQ ID NOs: 278–280, und //); Fig. 12D (von oben nach unten: SEQ ID NOs: 281–290) und Fig. 12E (von oben nach unten: SEQ ID NOs: 291–292, 283, 293–298).

[0045] Fig. 13A–C zeigen, dass ein PAM benötigt wird, um die Spaltung der Ziel-DNA durch den Cas9-tracrRNA:crRNA-Komplex zuzulassen.

[0046] Fig. 14A–C veranschaulichen, dass das Cas9 unter Verwendung eines einzelnen, künstlich erstellten RNA-Moleküls, das Eigenschaften von tracrRNA und crRNA miteinander kombiniert, programmiert werden kann. Chimäre A (SEQ ID NO: 299); Chimäre B (SEQ ID NO: 300).

[0047] Fig. 15 zeigt den RNA-vermittelten CRISPR/Cas-Immunweg vom Typ II.

[0048] Fig. 16A–B zeigen die Reinigung von Cas9-Nukleasen.

[0049] Fig. 17A–C zeigen, dass ein Cas9, die durch eine Dual-tracrRNA:crRNA gesteuert wird, ein Protospacer-Plasmid und Oligonukleotid-DNA spaltet. Fig. 17B (von oben nach unten, SEQ ID NOs: 301–303, und //); und Fig. 17C (von oben nach unten, SEQ ID NOs: 304–306, und //).

[0050] Fig. 18A–B zeigen, dass Cas9 eine von Mg²⁺ abhängige Endonuklease mit 3'–5'-Exonukleaseaktivität ist.

[0051] Fig. 19A–C veranschaulichen, dass die Dual-tracrRNA:crRNA-gesteuerte Cas9-Spaltung von Ziel-DNA ortsspezifisch ist. Fig. 19C (von oben nach unten SEQ ID NOs: 307–309, //, 337–339, und //).

[0052] Fig. 20A–B zeigen das die Dual-tracrRNA:crRNA-gesteuerte Cas9-Spaltung von Ziel-DNA schnell und effizient ist.

[0053] Fig. 21A–B zeigen, dass die HNH- und RuvC-ähnlichen Domänen von Cas9 die Spaltung des komplementären bzw. des nicht-komplementären DNA-Strangs steuern.

[0054] Fig. 22 zeigt, dass tracrRNA für die Erkennung der Ziel-DNA erforderlich ist.

[0055] Fig. 23A–B zeigen, dass eine minimale Region der tracrRNA in der Lage ist, die von der Dual-tracrRNA:crRNA vermittelte Spaltung der Ziel-DNA zu steuern.

- [0056]** Fig. 24A–D zeigen, dass die Dual-tracrRNA:crRNA-gesteuerte Spaltung der Ziel-DNA durch Cas9 spezifisch für eine Spezies sein kann.
- [0057]** Fig. 25A–C zeigen, dass eine Seed-Sequenz in der crRNA die von der Dual-tracrRNA:crRNA-gesteuerte Spaltung der Ziel-DNA durch Cas9 bestimmt. **Fig. 25A:** Ziel-DNA-Sonde 1 (SEQ ID NO: 310; Spacer 4 crRNA (1–42) (SEQ ID NO: 311); tracrRNA (15–89) (SEQ ID NO: //). **Fig. 25B** linkes Schaubild (SEQ ID NO: 310).
- [0058]** Fig. 26A–C zeigen, dass die PAM-Sequenz für die Spaltung einer Protospacer-Plasmid-DNA durch Cas9-tracrRNA:crRNA und für die Cas9-vermittelte Plasmid-DNA-Interferenz in bakteriellen Zellen essentiell ist. **Fig. 26B** (von oben nach unten: SEQ ID NOs: 312–314); und **Fig. 26C** (von oben nach unten, SEQ ID NOs: 315–320).
- [0059]** Fig. 27A–C zeigen, dass Cas9, die von einer einzelnen chimären RNA, welche die duale tracrRNA:crRNA nachahmt, geleitet wird, Protospacer-DNA spaltet. **Fig. 27C** (von oben nach unten: SEQ ID NOs: 121–324).
- [0060]** Fig. 28A–D zeigen die de novo-Erstellung von chimären RNAs, die auf die Gensequenz des grün fluoreszierenden Proteins (GFP) gerichtet sind. **Fig. 28B** (von oben nach unten: SEQ ID NOs: 325–326). **Fig. 28C:** GFP1 Ziel-Sequenz (SEQ ID NO: 327); GFP2 Ziel-Sequenz (SEQ ID NO: 328); GFP3 Ziel-Sequenz (SEQ ID NO: 329); GFP4 Ziel-Sequenz (SEQ ID NO: 330); GFP5 Ziel-Sequenz (SEQ ID NO: 331); GFP1 chimäre RNA (SEQ ID NO: 332); GFP2 chimäre RNA (SEQ ID NO: 333); GFP3 chimäre RNA (SEQ ID NO: 334); GFP4 chimäre RNA (SEQ ID NO: 335); GFP5 chimäre RNA (SEQ ID NO: 336).
- [0061]** Fig. 29A–E zeigen, dass die Co-Expression von Cas9 und Guide-RNA in humanen Zellen Brüche in doppelsträngiger DNA am Ziel-Locus erzeugt. **Fig. 29C** (von oben nach unten: SEQ ID NOs: 425–428).
- [0062]** Fig. 30A–B zeigen, dass Zelllysate aktive Cas9:sgRNA enthalten und die ortsspezifische Spaltung der DNA unterstützen.
- [0063]** Fig. 31A–B zeigen, dass die 3'-Extension von sgRNA-Konstrukten die ortsspezifische von NHEJ vermittelte Mutagenese verstärkt. **Fig. 31A** (von oben nach unten: SEQ ID NOs: 428–430).
- [0064]** Fig. 32A–B zeigen einen phylogenetischen Baum repräsentativer Cas9-Sequenzen von verschiedenen Organismen (A) sowie den Aufbau des Cas9-Locus für die Hauptgruppen des Baums (B).
- [0065]** Fig. 33A–E zeigen den Aufbau von Typ II-CRISPR-Cas von ausgewählten bakteriellen Spezies.
- [0066]** Fig. 34A–B zeigen die Co-Prozessierung von tracrRNA und prä-crRNA in ausgewählten Typ II-CRISPR-Cas-Systemen. **Fig. 34A** (von oben nach unten, SEQ ID NO: //, //, //, //, //, //, //, //); **Fig. 34B** (von oben nach unten, SEQ ID NO: //, //, //, //).
- [0067]** Fig. 35 zeigt ein Sequenz-Alignment von tracrRNA-Orthologen, das die Diversität der tracrRNA-Sequenzen aufzeigt.
- [0068]** Fig. 36A–F zeigen die Expression von bakteriellen tracrRNA-Orthologen und crRNAs, wie sie durch Deep-RNA-Sequenzierung erhalten wurden.
- [0069]** Fig. 37A–O führen alle tracrRNA-Orthologe und reife crRNAs auf, die durch Sequenzierung der untersuchten bakteriellen Spezies erhalten wurden, einschließlich Koordinaten (Bereich von Interesse) und entsprechenden cDNA-Sequenzen (von 5' nach 3').
- [0070]** Fig. 38A–B zeigen eine Tabelle von bakteriellen Spezies, die Typ II-CRISPR-Cas-Loci enthalten, welche durch das Vorhandensein des charakteristischen cas9-Gens charakterisiert sind. Diese Sequenzen wurden für die phylogenetischen Analysen verwendet.
- [0071]** Fig. 39A–B zeigen den Aufbau des CRISPR-Interferenz(CRISPRi)-Systems.
- [0072]** Fig. 40A–E zeigen, dass CRISPRi die Elongation und Initiation der Transkription wirksam unterdrückt.

- [0073]** Fig. 41A–B zeigen, dass CRISPRi durch Blockierung der Elongation der Transkription wirkt.
- [0074]** Fig. 42A–C zeigen die Targeting-Spezifität des CRISPRi-Systems.
- [0075]** Fig. 43A–F zeigen die Charakterisierung von Faktoren, welche die Wirksamkeit des Silencing beeinflussen.
- [0076]** Fig. 44A–C zeigen ein funktionales Profiling eines komplexen regulatorischen Netzwerks unter Verwendung von CRISPRi-Gen-Knockdown.
- [0077]** Fig. 45A–B zeigen das Gen-Silencing unter Verwendung von CRISPRi in Säugetierzellen.
- [0078]** Fig. 46 zeigt den Mechanismus des Typ II-CRISPRi-Systems von *S. pyogenes*.
- [0079]** Fig. 47A–B zeigen die Wachstumskurven von *E. coli*-Zellkulturen, die mit dCas9 und sgRNA co-transformiert wurden.
- [0080]** Fig. 48 zeigt, dass CRISPRi die Expression eines Reporter-Gens auf einem Multi-Kopien-Plasmid unterdrücken konnte.
- [0081]** Fig. 49A–C zeigen die RNA-seq-Daten von Zellen mit sgRNAs, die auf verschiedene Gene gerichtet sind.
- [0082]** Fig. 50A–E zeigen die Silencing-Wirkungen von sgRNAs mit benachbarten Doppel-Fehlpaarungen.
- [0083]** Fig. 51A–C zeigen die kombinatorischen Silencing-Wirkungen bei Verwendung von zwei sgRNAs zur Regulierung eines einzelnen Gens.
- [0084]** Fig. 52 zeigt, dass die sgRNA-Repression abhängig von den Ziel-Loci sowie von der relativen Entfernung vom Transkriptions-Startpunkt ist.
- [0085]** Fig. 53A–C zeigen experimentelle Ergebnisse, die belegen, dass eine Variante des ortsspezifischen Cas9-Polypeptids (dCas9) für die vorliegenden Verfahren funktionsfähig ist, wenn dCas9 eine verringerte Aktivität lediglich in der RuvC1-Domäne (z. B. D10A), lediglich in der HNH-Domäne (z. B. H840A) oder in beiden Domänen (z. B. D10A und H840A) aufweist.
- [0086]** Fig. 54A–C führen Beispiele für geeignete Fusionspartner (oder Fragmente davon) für eine bestimmte Variante des ortsspezifischen Cas9-Polypeptids auf. Beispiele umfassen die dort aufgeführten, sind jedoch nicht auf diese beschränkt.
- [0087]** Fig. 55A–D zeigen, dass ein chimäres ortsspezifisches Polypeptid verwendet werden kann, um die Transkription in humanen Zellen zu aktivieren (zu erhöhen).
- [0088]** Fig. 56 zeigt, dass ein chimäres ortsspezifisches Polypeptid verwendet werden kann, um die Transkription in humanen Zellen zu reprimieren (zu verringern).
- [0089]** Fig. 57A–B zeigen, dass künstliche Sequenzen, die etwa 50% Identität mit natürlich vorkommenden tracrRNAs und crRNAs teilen, gemeinsam mit Cas9 wirken können, um Ziel-DNA zu spalten, solange die Struktur der Protein-bindenden Domäne der auf DNA gerichteten RNA konserviert ist.

DEFINITIONEN – TEIL I

[0090] Die Begriffe "Polynukleotid" und "Nukleinsäure" werden vorliegend austauschbar verwendet und beziehen sich auf eine polymere Form von Nukleotiden beliebiger Länge, entweder Ribonukleotide oder Desoxyribonukleotide. Somit umfassen diese Begriffe (sind jedoch nicht beschränkt auf) einzelsträngige, doppelsträngige oder mehrsträngige DNA oder RNA, genomische DNA, cDNA, DNA-RNA-Hybride oder ein Polymer, das Purin- und Pyrimidin-Basen oder andere natürlich, chemische oder biochemisch modifizierte, nicht-natürlich oder derivatisierte Nukleotid-Basen umfasst. "Oligonukleotide" bezeichnet im Allgemeinen Polynukleotide zwischen etwa 5 und etwa 100 Nukleotiden von einzelsträngiger oder doppelsträngiger DNA. Für die Zwecke der vorliegenden Offenbarung jedoch gibt es keine obere Grenze bezüglich der Länge eines Oligonukleotids.

Oligonukleotide sind ferner als "Oligomere" oder "Oligos" bekannt und können von Genen isoliert werden oder chemisch mit im Stand der Technik bekannten Verfahren synthetisiert werden. Die Begriffe "Polynukleotid" und "Nukleinsäure" sollten dahingehend verstanden werden, dass sie, soweit es für die beschriebenen Ausführungsformen zutreffend ist, einzelsträngige (wie z. B. Sense oder Anti-Sense) und doppelsträngige Polynukleotide umfassen.

[0091] Eine "Haarnadelstruktur" (stem loop structure) bezeichnet eine Nukleinsäure mit einer Sekundärstruktur, die eine Region von Nukleotiden umfasst, von denen bekannt ist, oder bei denen vorhergesagt wird, dass sie einen Doppelstrang bilden (Stamm-Teil), der an einer Seite mit einer Region von vorwiegend einzelsträngigen Nukleotiden verknüpft ist (Schleifen-Teil). Die Begriffe "Haarnadel"-Struktur und "Fold-Back"-Struktur werden ebenfalls vorliegend verwendet, um auf die Haarnadelstrukturen hinzuweisen. Solche Strukturen sind im Stand der Technik bekannt, und die Begriffe werden im Einklang mit ihren üblichen Bedeutungen verwendet. Wie es im Stand der Technik bekannt ist, benötigt eine Haarnadelstruktur keine exakte Basenpaarung. Der Stamm kann daher eine oder mehrere Basen-Fehlpaarungen aufweisen. Alternativ dazu kann die Basenpaarung genau sein, d. h. keine Fehlpaarungen enthalten.

[0092] Mit "hybridisierbar" oder "komplementär" oder "im Wesentlichen komplementär" ist gemeint, dass eine Nukleinsäure (z. B. RNA) eine Sequenz von Nukleotiden umfasst, die es ihr ermöglichen, nicht-kovalent, d. h. sich durch Bildung von Watson-Crick-Basenpaarungen und/oder durch G/U-Basenpaarungen, an eine andere Nukleinsäure "anzulagern" oder zu "hybridisieren", wobei dies in Sequenz-spezifischer, antiparalleler Weise (d. h. eine Nukleinsäure bindet spezifisch an eine komplementäre Nukleinsäure) unter geeigneten in vitro- und/oder in vivo-Bedingungen von Temperatur und Ionenstärke der Lösung erfolgt. Wie im Stand der Technik bekannt ist, umfassen Standard-Watson-Crick-Basenpaarungen: Adenin (A) das sich mit Thymin (T) paart, Adenin (A), das sich mit Uracil (U) paart, und Guanin (G), das sich mit Cytosin (C) paart. [DNA, RNA]. Darüber hinaus ist es ferner im Stand der Technik bekannt, dass bei der Hybridisierung zwischen zwei RNA-Molekülen (z. B. dsRNA) sich die Base Guanin (G) mit Uracil (U) paart. Beispielsweise ist die G/U-Basenpaarung teilweise für die Degeneriertheit (d. h. die Redundanz) des genetischen Codes im Zusammenhang mit den tRNA-Anti-Codon-Basenpaarungen mit Codons in der mRNA verantwortlich. Im Zusammenhang mit der vorliegenden Offenbarung wird ein Guanin (G) eines Protein-bindenden Segments (dsRNA-Duplex) eines vorliegenden auf DNA gerichteten RNA-Moleküls als komplementär zu einem Uracil (U) angesehen und umgekehrt. Wenn daher ein G/U-Basenpaar an einer bestimmten Nukleotidposition eines Protein-bindenden Segments (dsRNA-Duplex) eines vorliegenden auf DNA gerichteten RNA-Moleküls gebildet werden kann, wird diese Position nicht als nicht-komplementär angesehen, sondern wird vielmehr als komplementär angesehen.

[0093] Bedingungen für die Hybridisierung und das Waschen sind hinreichend bekannt und werden in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), insbesondere Kapitel 11 und Tabelle 11.1 darin; und Sambrook, J. and Russell, W., *Molecular Cloning: A Laboratory Manual*, 3. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (2001), beispielhaft aufgeführt. Die Bedingungen von Temperatur und Ionenstärke bestimmen die "Stringenz" der Hybridisierung.

[0094] Die Hybridisierung erfordert, dass die beiden Nukleinsäuren komplementäre Sequenzen umfassen, obwohl Fehlpaarungen zwischen Basen möglich sind. Geeignete Bedingungen für die Hybridisierung zwischen zwei Nukleinsäuren sind abhängig von der Länge der Nukleinsäuren und vom Grad der Komplementarität, wobei diese Variablen im Stand der Technik gut bekannt sind. Je größer der Grad an Komplementarität zwischen zwei Nukleotidsequenzen ist, desto größer ist der Wert für die Schmelztemperatur (T_m) von Hybriden aus Nukleinsäuren, die diese Sequenzen aufweisen. Bei Hybridisierungen zwischen Nukleinsäuren mit kurzen komplementäre Abschnitten (z. B. Komplementarität über 35 oder weniger, 30 oder weniger, 25 oder weniger, 22 oder weniger, 20 oder weniger, oder 18 oder weniger Nukleotide) wird die Position von Fehlpaarungen wichtig (siehe Sambrook et al., siehe oben, 11.7–11.8). Üblicherweise beträgt die Länge für eine hybridisierbare Nukleinsäure mindestens etwa 10 Nukleotide. Beispielhafte minimale Länge für eine hybridisierbare Nukleinsäure sind: mindestens etwa 15 Nukleotide, mindestens etwa 20 Nukleotide, mindestens etwa 22 Nukleotide, mindestens etwa 25 Nukleotide und mindestens etwa 30 Nukleotide. Darüber hinaus wird der Fachmann erkennen, dass die Temperatur und die Salzkonzentration der Wasch-Lösung bei Bedarf im Hinblick auf Faktoren wie die Länge der komplementären Regionen und der Grad der Komplementarität angepasst werden können.

[0095] Es wird im Stand der Technik verstanden, dass die Sequenz eines Polynukleotids nicht 100% komplementär zu der Sequenz seiner Ziel-Nukleinsäure sein muss, um spezifisch hybridisierbar oder hybridisierbar zu sein. Darüber hinaus kann ein Polynukleotid über ein oder mehrere Segmente so hybridisieren, dass zwischenliegende oder benachbarte Segmente nicht an dem Hybridisierungsergebnis beteiligt sind (z. B. ei-

ne Schleifenstruktur oder eine Haarnadelstruktur). Ein Polynukleotid kann mindestens 70%, mindestens 80%, mindestens 90%, mindestens 95%, mindestens 99% oder 100% Sequenzkomplementarität zu einer Zielregion innerhalb der Ziel-Nukleinsäuresequenz, auf die sie gerichtet sind, aufweisen. Beispielsweise würde eine Antisense-Nukleinsäure, in der 18 von 20 Nukleotiden der Antisense-Verbindung komplementär zu einem Zielbereich sind und daher spezifisch hybridisieren würden, 90% Komplementarität aufweisen. In diesem Beispiel können die verbleibenden nicht-komplementären Nukleotide im Cluster vorliegen oder zwischen den komplementären Nukleotiden verteilt sein, und sie brauchen nicht miteinander oder mit komplementären Nukleotiden zusammenhängen. Die prozentuale Komplementarität zwischen bestimmten Abschnitten von Nukleinsäuresequenzen innerhalb von Nukleinsäuren kann routinemäßig unter Verwendung von im Stand der Technik bekannten BLAST-Programmen (basic local alignment search tools) und PowerBLAST-Programmen (Altschul et al., J. Mol. Biol., 1990, 215, 403–410; Zhang und Madden, Genome Res., 1997, 7, 649–656) bestimmt werden, oder durch Verwendung des Gap-Programms (Wisconsin Sequence Analysis Package, Version 8 für Unix, Genetics Computer Group, University Research Park, Madison Wis.) unter Verwendung der vorgegebenen Einstellungen, die die Algorithmen von Smith & Waterman verwenden (Adv. Appl. Math., 1981, 2, 482–489).

[0096] Die Begriffe "Peptide", "Polypeptide" und "Protein" werden vorliegend austauschbar verwendet und beziehen sich auf eine polymere Form aus Aminosäuren von beliebiger Länge, die kodierte und nicht-kodierte Aminosäuren, chemisch oder biochemisch modifizierte oder derivatisierte Aminosäuren und Polypeptide mit modifiziertem Peptid-Rückgrat umfassen können.

[0097] "Bindend" bezeichnet wie vorliegend verwendet (z. B. in Bezug auf eine RNA-bindende Domäne eines Polypeptids) eine nicht-kovalente Interaktion zwischen Makromolekülen (z. B. zwischen einem Protein und einer Nukleinsäure). Es wird gesagt, dass die Makromoleküle im Zustand einer nicht-kovalenten Interaktion "assoziiert" sind oder "interagieren" oder "binden" (wenn z. B. gesagt wird, dass ein Molekül X mit einem Molekül Y interagiert, so ist gemeint, dass das Molekül X an das Molekül Y in nicht-kovalenter Weise bindet). Nicht alle Bestandteile einer Bindungsinteraktion müssen Sequenz-spezifisch sein (z. B. Kontakte mit Phosphatresten in einem DNA-Rückgrat), wobei einige Teile einer bindenden Interaktion Sequenz-spezifisch sein können. Bindende Interaktionen sind im Allgemeinen durch eine Dissoziationskonstante (K_d) von weniger als 10^{-6} M, weniger als 10^{-7} M, weniger als 10^{-8} M, weniger als 10^{-9} M, weniger als 10^{-10} M, weniger als 10^{-11} M, weniger als 10^{-12} M, weniger als 10^{-13} M, weniger als 10^{-14} M, oder weniger als 10^{-15} M charakterisiert. "Affinität" bezieht sich auf die Stärke der Bindung, wobei eine erhöhte Bindungsaffinität mit einem geringeren K_d -Wert korreliert ist.

[0098] Mit "bindende Domäne" ist eine Proteindomäne gemeint, die in der Lage ist, nicht-kovalent an ein anderes Molekül zu binden. Eine bindende Domäne kann beispielsweise an ein DNA-Molekül binden (ein DNA-bindendes Protein), an ein RNA-Molekül (ein RNA-bindendes Protein) und/oder an ein Protein-Molekül (ein Protein-bindendes Protein). Im Falle eines Proteins, das an eine Proteindomäne bindet, kann dieses Protein an sich selbst binden (wobei Homodimere, Homotrimere, usw. gebildet werden) und/oder es kann an ein oder mehrere Moleküle eines anderen Proteins oder anderer Proteine binden.

[0099] Der Begriff "konservative Aminosäure-Substitution" bezieht sich auf die Austauschbarkeit von Aminosäure-Resten mit ähnlicher Seitenkette in Proteinen. Beispielsweise besteht eine Gruppe von Aminosäuren, die über aliphatische Seitenketten verfügt, aus Glycin, Alanin, Valin, Leucin und Isoleucin; eine Gruppe von Aminosäuren, die über aliphatische Hydroxyl-Seitenketten verfügt, besteht aus Serin und Threonin; eine Gruppe von Aminosäuren, die über Amid-haltige Seitenketten verfügt, besteht aus Asparagin und Glutamin; eine Gruppe von Aminosäuren, die über aromatische Seitenketten verfügt, besteht aus Phenylalanin, Tyrosin und Tryptophan; eine Gruppe von Aminosäuren, die über basische Seitenketten verfügt, besteht aus Lysin, Arginin und Histidin; eine Gruppe von Aminosäuren, die über saure Seitenketten verfügt, besteht aus Glutamat und Aspartat; und eine Gruppe von Aminosäuren, die über schwefelhaltige Seitenketten verfügt besteht aus Cystein und Methionin. Beispielhafte konservative Aminosäure-Substitutionsgruppen sind Valin-Leucin-Isoleucin, Phenylalanin-Tyrosin, Lysin-Arginin, Alanin-Valin und Asparagin-Glutamin.

[0100] Ein Polynukleotid oder Polypeptid hat eine bestimmte prozentuale "Sequenzidentität" zu einem anderen Polynukleotid oder einem anderen Polypeptid, was bedeutet, dass der prozentuale Anteil von Basen oder Aminosäuren in einem Alignment gleich ist und diese in der gleichen relativen Position vorliegen, wenn die beiden Sequenzen verglichen werden. Die Sequenzidentität kann auf verschiedene Weise bestimmt werden. Um die Sequenzidentität zu bestimmen, können Sequenzen durch Verwendung verschiedener Verfahren und Computerprogramme (z. B. BLAST, T-COFFEE, MUSCLE, MAFFT usw.), die über das World Wide Web auf Websites wie ncbi.nlm.nih.gov/BLAST, ebi.ac.uk/Tools/msa/tcoffee/, ebi.ac.uk/Tools/msa/muscle/, mafft.cbrc.jp/alignment/software/ verfügbar sind, miteinander verglichen werden. Siehe z. B. Altschul et al. (1990), J. Mol. Biol. 215: 403–10.

[0101] Eine DNA-Sequenz, die für eine bestimmte RNA "kodiert" ist eine DNA-Nukleinsäuresequenz, die in RNA transkribiert wird. Ein DNA-Polynukleotid kann für eine RNA (mRNA) kodieren, die in Protein translatiert wird, oder ein DNA-Polynukleotid kann für eine RNA kodieren, die nicht in Protein translatiert wird (z. B. tRNA, rRNA oder eine auf DNA gerichtete RNA; auch als "nicht-kodierende" RNA oder "ncRNA" bezeichnet).

[0102] Eine "Protein-kodierende Sequenz" oder eine Sequenz, die für ein bestimmtes Protein oder Polypeptid kodiert, ist eine Nukleinsäuresequenz, die in mRNA transkribiert wird (im Falle von DNA) und in vitro oder in vivo in ein Polypeptid translatiert wird (in Falle von mRNA), wenn sie unter Kontrolle von geeigneten regulatorischen Sequenzen gestellt wird. Die Grenzen der kodierenden Sequenz werden durch ein Start-Codon am 5'-Terminus (N-Terminus) und ein Translations-Stopp-Nonsense-Codon am 3'-Terminus (C-Terminus) bestimmt. Eine kodierende Sequenz kann cDNA von prokaryontischer oder eukaryontischer mRNA, genomische DNA-Sequenzen von prokaryontischer oder eukaryontischer DNA und synthetische Nukleinsäuren umfassen, ist jedoch nicht auf diese beschränkt. Eine Transkriptions-Terminations-Sequenz wird üblicherweise 3' von der kodierenden Sequenz lokalisiert sein.

[0103] Wie vorliegend verwendet ist eine "Promotor-Sequenz" ein DNA-regulatorischer Bereich, der in der Lage ist, RNA-Polymerase zu binden und eine Transkription einer stromabwärts (in 3'-Richtung) gelegenen, kodierenden oder nicht-kodierenden Sequenz auszulösen. Zum Zwecke der Definition der vorliegenden Erfindung wird die Promotor-Region an ihrem 3'-Terminus durch eine Transkriptions-Initiationsstelle begrenzt und erstreckt sich stromaufwärts (in 5'-Richtung), so dass sie die minimale Anzahl von Basen oder Elementen einschließt, die für die Initiation der Transkription in einem Ausmaß, das den Hintergrund überschreitet, erforderlich sind. Innerhalb der Promotor-Sequenz wird eine Transkriptions-Initiationsstelle gefunden werden, sowie auch Protein-bindende Domänen, die für die Bindung einer RNA-Polymerase verantwortlich sind. Eukaryontische Promotoren werden oftmals (jedoch nicht immer) "TATA"-Boxen und "CAT"-Boxen umfassen. Verschiedene Promotoren, einschließlich induzierbarer Promotoren, können verwendet werden, um die verschiedenen Vektoren der vorliegenden Erfindung zu steuern.

[0104] Ein Promotor kann ein konstitutiv aktiver Promotor sein (d. h. ein Promotor, der sich konstitutiv in einem aktiven Zustand/"AN" – Zustand befindet), es kann ein induzierbarer Promotor sein (d. h. ein Promotor, dessen Zustand aktiv/"AN" oder inaktiv/"AUS" durch einen externen Stimulus kontrolliert wird, z. B. durch das Vorhandensein einer bestimmten Temperatur, einer Verbindung oder eines Proteins), es kann ein räumlich limitierter Promotor sein (d. h. transkriptionale Kontroll-Elemente, Enhancer, usw.) (z. B. ein Gewebe-spezifischer Promotor, ein Zelltyp-spezifischer Promotor, usw.), und es kann ein zeitlich limitierter Promotor sein (d. h. der Promotor ist während bestimmter Phasen der embryonischen Entwicklung oder während bestimmter Phasen eines biologischen Prozesses, z. B. Haarfollikelzyklus in Mäusen, im Zustand "AN" oder "AUS").

[0105] Geeignete Promotoren können von Viren abgeleitet sein und können daher als virale Promotoren bezeichnet werden, oder sie können von beliebigen Organismen abgeleitet werden, einschließlich prokaryontischer und eukaryontischer Organismen. Geeignete Promotoren können verwendet werden, um die Expression durch eine beliebige RNA-Polymerase zu steuern (z. B. pol I, pol II, pol III). Beispielhafte Promotoren umfassen (sind jedoch nicht beschränkt auf) den SV40-Early-Promotor, den Maus-Brusttumorstromavirus-Long-Terminal-Repeat(LTR)-Promotor, den Adenovirus-Major-Late-Promotor (Ad MLP), einen Herpes-Simplex-Virus(HSV)-Promotor, einen Cytomegalovirus(CMV)-Promotor, wie z. B. die CMV-Immediate-Early-Promotorregion (CMVIE), einen Rous-Sarkoma-Virus(RSV)-Promotor, einen humanen U6-Small-Nuclear-Promotor (U6) (Miyagishi et al., Nature Biotechnology 20, 497–500 (2002)), einen verstärkten U6-Promotor (z. B. Xia et al., Nucleic Acids Res. 2003 Sep 1; 31(17)), einen humanen H1-Promotor (H1), und ähnliche.

[0106] Beispiele für induzierbare Promotoren umfassen (sind jedoch nicht beschränkt auf) den T7-RNA-Polymerase-Promotor, den T3-RNA-Polymerase-Promotor, den Isopropyl-Beta-D-Thiogalactopyranosid(IPTG)-regulierten Promotor, den Laktose-induzierten Promotor, den Hitzeschock-Promotor, den Tetracyclin-regulierten Promotor, den Steroid-regulierten Promotor, den Metall-regulierten Promotor, den Östrogenrezeptor-regulierten Promotor usw. Induzierbare Promotoren können daher durch Moleküle reguliert werden, die Doxycyclin; RNA-Polymerase, z. B. T7-RNA-Polymerase; einen Östrogenrezeptor; eine Östrogenrezeptor-Fusion; usw. umfassen (jedoch nicht auf diese beschränkt sind).

[0107] In einigen Ausführungsformen ist der Promotor ein räumlich limitierter Promotor (d. h. ein Zelltyp-spezifischer Promotor, ein Gewebe-spezifischer Promotor, usw.), so dass der Promotor in einem multizellulären Organismus in einer Untergruppe von spezifischen Zellen aktiv ist (d. h. "AN"). Räumlich limitierte Promotoren können auch als Verstärker, Transkriptions-Kontrollelemente, Kontrollsequenzen, usw. bezeichnet werden. Jeder beliebige räumlich limitierte Promotor kann verwendet werden, und die Auswahl eines geeigneten Pro-

motors (z. B. eines Gehirn-spezifischen Promotor, eines Promotors, der die Expression in einer Untergruppe von Neuronen steuert, eines Promotor, der die Expression in einer Keimlinie steuert, eines Promotors, der die Expression in den Lungen steuert, eines Promotors, der die Expression in den Muskeln steuert, eines Promotors, der die Expression in Inselzellen der Bauspeicheldrüse steuert, usw.) wird vom Organismus abhängen. Beispielsweise sind verschiedene räumlich limitierte Promotoren für Pflanzen, Fliegen, Würmer, Säugetiere, Mäuse, usw. bekannt. Somit kann ein räumlich limitierter Promotor verwendet werden, um die Expression einer Nukleinsäure, die für ein vorliegendes ortsspezifisch modifizierendes Polypeptid kodiert, in einer Vielzahl von verschiedenen Geweben und Zelltypen in Abhängigkeit vom Organismus zu regulieren. Einige räumlich limitierte Promotoren sind auch in zeitlicher Hinsicht limitiert, so dass der Promotor während bestimmter Phasen der embryonalen Entwicklung oder während bestimmter Phasen eines biologischen Prozesses (z. B. Haarfollikelzyklus in Mäusen) im Zustand "AN" oder im Zustand "AUS" ist.

[0108] Beispiele für räumlich limitierte Promotoren umfassen (sind jedoch nicht beschränkt auf) Neuronen-spezifische Promotoren, Adipozyten-spezifische Promotoren, Kardiomyozytenspezifische Promotoren, für glatte Muskelzellen spezifische Promotoren, Photorezeptor-spezifische Promotoren, usw. Neuronen-spezifische, räumlich limitierte Promotoren umfassen (sind jedoch nicht beschränkt auf) einen Neuronen-spezifischen-Enolase(NSE)-Promotor (siehe z. B., EMBL HSENO2, X51956); einen Promotor der Decarboxylase für aromatische Aminosäuren (AADC); einen Neurofilament-Promotor (siehe z. B. GenBank HUMNFL, L04147); einen Synapsin-Promotor (z. B. GenBank HUMSYNIB, M55301); einen thy-1-Promotor (siehe z. B. Chef et al. (1987) Cell 51: 7–19; und Llewellyn, et al. (2010) Nat. Med. 16(10): 1161–1166); einen Serotonin-Rezeptor-Promotor (siehe z. B. GenBank S62283); einen Thyrosin-Hydroxylase-Promotor (TH) (siehe z. B. Oh et al. (2009) Gene Ther 16: 437; Sasaoka et al. (1992) Mol. Brain Res. 16: 274; Boundy et al. (1998) J. Neurosci. 18: 9989; und Kaneda et al. (1991) Neuron 6: 583–594); einen GnRH-Promotor (siehe z. B. Radovick et al. (1991) Proc. Natl. Acad. Sci. USA 88: 3402–3406); einen L7-Promotor (siehe z. B. Oberdick et al. (1990) Science 248: 223–226); einen DNMT-Promotor (siehe z. B. Bartge et al. (1988) Proc. Natl. Acad. Sci. USA 85: 3648–3652); einen Enkephalin-Promoter (siehe z. B. Comb et al. (1988) EMBO J. 17: 3793–3805); einen Myelin-basisches-Protein (MBP)-Promoter; einen Ca²⁺ Calmodulin-abhängige Proteinkinase II-alpha (CamKIIα)-Promoter (siehe z. B. Mayford et al. (1996) Proc. Natl. Acad. Sci. USA 93: 13250; und Casanova et al. (2001) Genesis 31: 37); einen CMV-Enhancer/Blutplättchen-abgeleiteter Wachstumsfaktor-β-Promoter (siehe z. B. Liu et al. (2004) Gene Therapy 11: 52–60); und ähnliche.

[0109] Adipozyten-spezifische räumlich limitierte Promotoren umfassen (sind jedoch nicht beschränkt auf) einen aP2-Gen-Promotor/Enhancer, z. B. eine Region von –5,4 kb bis +21 bp des humanen aP2-Gens (siehe z. B. Tozzo et al. (1997) Endocrinol. 138: 1604; Ross et al. (1990) Proc. Natl. Acad. Sci. USA 87: 9590; und Pavjani et al. (2005) Nat. Med. 11: 797); einen Glukose-Transporter-4(GLUT4)-Promoter (siehe z. B. Knight et al. (2003) Proc. Natl. Acad. Sci. USA 100: 14725); einen Fettsäure-Translokase(FAT/CD36)-Promoter (siehe z. B. Kuriki et al. (2002) Biol. Pharm. Bull. 25: 1476; und Sato et al. (2002) J. Biol. Chem. 277: 15703); einen Stearoyl-CoA-Desaturase-1(SCD1)-Promoter (Tabor et al. (1999) J. Biol. Chem. 274: 20603); einen Leptin-Promotor (siehe z. B. Mason et al. (1998) Endocrinol. 139: 1013; und Chen et al. (1999) Biochem. Biophys. Res. Comm. 262: 187); einen Adiponectin-Promoter (siehe z. B. Kita et al. (2005) Biochem. Biophys. Res. Comm. 331: 484; und Chakrabarti (2010) Endocrinol. 151: 2408); einen Adipsin-Promoter (siehe z. B. Platt et al. (1989) Proc. Natl. Acad. Sci. USA 86: 7490); einen Resistin-Promoter (siehe z. B. Seo et al. (2003) Molec. Endocrinol. 17: 1522); und ähnliche.

[0110] Kardiomyozyten-spezifische, räumlich limitierte Promotoren umfassen (sind jedoch nicht beschränkt auf) Kontrollsequenzen, die von den folgenden Genen abgeleitet sind: Myosinleichte Kette-2, α-Myosin-schwere Kette, AE3, kardiäres Troponin-C, kardiäres Actin und ähnliche. Franz et al. (1997) Cardiovasc. Res. 35: 560–566; Robbins et al. (1995) Ann. N. Y. Acad. Sci. 752: 492–505; Linn et al. (1995) Circ. Res. 76: 584–591; Parmacek et al. (1994) Mol. Cell. Biol. 14: 1870–1885; Hunter et al. (1993) Hypertension 22: 608–617; und Sartorelli et al. (1992) Proc. Natl. Acad. Sci. USA 89: 4047–4051.

[0111] Für glatte Muskeln spezifische, räumlich limitierte Promotoren umfassen (sind jedoch nicht beschränkt auf) einen SM22-α-Promotor (siehe z. B. Akyürek et al. (2000) Mol. Med. 6: 983; und US Patent Nr. 7,169,874); einen Smoothelin-Promoter (siehe z. B. WO 2001/018048); einen α-glatte Muskel-Actin-Promotor; und ähnliche. Beispielsweise konnte gezeigt werden, dass eine 0,4 kb Region des SM22α-Promotors, innerhalb derer zwei CArG-Elemente liegen, die Zell-spezifische Expression im vaskulären glatten Muskel vermittelt (siehe z. B. Kim, et al. (1997) Mol. Cell. Biol. 17, 2266–2278; Li, et al., (1996) J. Cell Biol. 132, 849–859; und Moessler, et al. (1996) Development 122, 2415–2425).

[0112] Photorezeptor-spezifische, räumlich limitierte Promotoren umfassen (sind jedoch nicht beschränkt auf) einen Rhodopsin-Promotor; einen Rhodopsin-Kinase-Promotor (Young et al. (2003) Ophthalmol. Vis. Sci. 44: 4076); einen Beta-Phosphodiesterase-Gen-Promotor (Nicoud et al. (2007) J. Gene Med. 9: 1015); einen Retinitis-Pigmentosa-Gen-Promoter (Nicoud et al. (2007) siehe oben); einen Interphotorezeptor-Retinoid-Bindungsprotein(IRBP)-Gen-Enhancer (Nicoud et al. (2007) siehe oben); einen IRBP-Gen-Promoter (Yokoyama et al. (1992) Exp Eye Res. 55: 225); und ähnliche.

[0113] Die Begriffe "DNA-regulatorische Sequenzen", "Kontrollelemente" und "regulatorische Elemente" werden vorliegend austauschbar verwendet und beziehen sich auf transkriptionale und translationale Kontrollsequenzen, wie beispielsweise Promotoren, Verstärker, Polyadenylierungssignale, Terminatoren, Protein-Abbau-Signale und ähnliche, die für die Transkription einer nicht-kodierenden Sequenz (z. B. auf DNA gerichtete RNA) oder einer kodierenden Sequenz (z. B. ortsspezifisch modifizierendes Polypeptid oder Cas9/Csn1-Polypeptid) sorgen und/oder diese regulieren und/oder die Translation eines kodierten Polypeptids regulieren.

[0114] Der Begriff "natürlich vorkommend" oder "unmodifiziert", wie er vorliegend für eine Nukleinsäure, ein Polypeptid, eine Zelle oder einen Organismus verwendet wird, bezieht sich auf eine Nukleinsäure, ein Polypeptid, eine Zelle oder einen Organismus, die/das/der in der Natur vorgefunden wird. Beispielsweise ist ein Polypeptid oder eine Polynukleotidsequenz, die in einem Organismus (einschließlich Viren) vorliegt und das/die von einer natürlichen Quelle isoliert werden kann und nicht bewusst von Menschen im Labor modifiziert wurde, natürlich vorkommt.

[0115] Der Begriff "chimär", wie er vorliegend für eine Nukleinsäure oder ein Polypeptid verwendet wird, bezieht sich auf zwei Bestandteile, die durch Strukturen definiert sind, welche von verschiedenen Quellen abgeleitet sind. Wenn beispielsweise "chimär" im Zusammenhang mit einem chimären Polypeptid (z. B. einem chimären Cas9/Csn1-Protein) verwendet wird, umfasst das chimäre Polypeptid Aminosäuresequenzen, die von unterschiedlichen Polypeptiden abgeleitet sind. Ein chimäres Polypeptid kann entweder modifizierte oder natürlich vorkommende Polypeptidsequenzen umfassen (z. B. eine erste Aminosäuresequenz von einem modifizierten oder unmodifizierten Cas9/Csn1-Protein; und eine zweite Aminosäuresequenz, bei der es sich nicht um das Cas9/Csn1-Protein handelt). In ähnlicher Weise umfasst "chimär" im Zusammenhang mit einem Polynukleotid, das für ein chimäres Polypeptid kodiert, Nukleotidsequenzen, die sich aus unterschiedlichen kodierenden Regionen ableiten (z. B. eine erste Nukleotidsequenz, die für ein modifiziertes oder unmodifiziertes Cas9/Csn1-Protein kodiert; und eine zweite Nukleotidsequenz, die für ein Polypeptid kodiert, bei dem es sich nicht um ein Cas9/Csn1-Protein handelt).

[0116] Der Begriff "chimäres Polypeptid" bezeichnet ein Polypeptid, welches durch die Kombination (d. h. die "Fusion") von zwei ansonsten getrennten Segmenten von Aminosäuresequenzen hergestellt wurde, üblicherweise durch menschlichen Eingriff. Ein Polypeptid, das eine chimäre Aminosäuresequenz umfasst, ist ein chimäres Polypeptid. Einige chimäre Polypeptide können als "Fusionsvarianten" bezeichnet werden.

[0117] "Heterolog" bedeutet wie vorliegend verwendet eine Nukleotid- oder Polypeptidsequenz, die nicht in der nativen Nukleinsäure bzw. in dem nativen Protein gefunden wird. Beispielsweise kann die RNA-bindende Domäne eines natürlich vorkommenden bakteriellen Cas9/Csn1-Polypeptids (oder einer Variante davon) in einem chimären Cas9/Csn1-Protein an eine heterologe Polypeptidsequenz fusioniert sein (d. h. eine Polypeptidsequenz von einem anderen Protein als Cas9/Csn1 oder eine Polypeptidsequenz von einem anderen Organismus). Die heterologe Polypeptidsequenz kann eine Aktivität (z. B. eine enzymatische Aktivität) zeigen, die auch durch das chimäre Cas9/Csn1-Protein gezeigt werden wird (z. B. Methyltransferase-Aktivität, Acetyltransferase-Aktivität, Kinase-Aktivität, ubiquitinierende Aktivität, usw. Eine heterologe Nukleinsäuresequenz kann an eine natürlich vorkommende Nukleinsäuresequenz (oder eine Variante davon) gekoppelt sein (z. B. durch genetische Modifikation), wobei eine chimäre Nukleinsäuresequenz erzeugt wird, die für ein chimäres Polypeptid kodiert. Als weiteres Beispiel kann in einer Fusion eines ortsspezifischen Varianten-Cas9-Polypeptid ein ortsspezifisches Varianten-Cas9-Polypeptid an ein heterologes Polypeptid (d. h. ein Polypeptid, bei dem es sich nicht um Cas9 handelt) fusioniert sein, das eine Aktivität aufweist, welche auch von der Fusion des ortsspezifischen Varianten-Cas9-Polypeptids gezeigt wird. Eine heterologe Nukleinsäuresequenz kann an ein ortsspezifisches Varianten-Cas9-Polypeptid gekoppelt sein (z. B. durch genetische Modifikation), wobei eine Nukleinsäuresequenz erzeugt wird, die für eine Fusion des ortsspezifischen Varianten-Cas9-Polypeptids kodiert.

[0118] "Rekombinant" bedeutet wie vorliegend verwendet, dass eine bestimmte Nukleinsäure (DNA oder RNA) das Produkt verschiedener Kombinationen aus Klonierung, Verdau, Polymerase-Kettenreaktion (PCR) und/oder Ligationsschritten ist, was zu einem Konstrukt führt, welches eine strukturelle kodierende oder nicht-

kodierende Sequenz aufweist, die von endogenen Nukleinsäuren, die in natürlichen Systemen gefunden werden, unterscheidbar ist. DNA-Sequenzen, die für Polypeptide kodieren, können aus cDNA-Fragmenten oder aus einer Gruppe von synthetischen Oligonukleotiden zusammengesetzt werden, wobei eine synthetische Nukleinsäure entsteht, die sich von einer rekombinanten Transkriptionseinheit, die in einer Zelle vorliegt, oder in einem zellfreien Transkriptions- und Translationssystem exprimieren lässt. Genomische DNA, welche die relevanten Sequenzen umfasst, kann ferner für die Bildung eines rekombinanten Gens oder einer Transkriptionseinheit verwendet werden. Sequenzen aus nicht-translatierter DNA können 5' oder 3' vom offenen Leserahmen vorhanden sein, wobei solche Sequenzen die Manipulation oder Expression der kodierenden Regionen nicht stören, und sie können tatsächlich wirksam sein, indem sie die Herstellung des gewünschten Produkts durch verschiedene Mechanismen modulieren (siehe "DNA-regulatorische Sequenzen", siehe unten). Alternativ dazu können DNA-Sequenzen, die für RNA kodieren (z. B. auf DNA gerichtete RNA), welche nicht translatiert wird, auch als rekombinant angesehen werden. Somit betrifft der Begriff "rekombinante" Nukleinsäure eine solche, die nicht natürlich vorkommt, z. B. eine solche, die durch künstliche Kombination von zwei ansonsten getrennten Sequenz-Elementen durch menschlichen Eingriff hergestellt wird. Diese künstliche Kombination wird oftmals durch chemische Synthese oder durch künstliche Manipulation von isolierten Nukleinsäure-Segmenten, z. B. durch Verfahren der genetischen Modifikation, erreicht. Dies wird üblicherweise durchgeführt, um ein Codon gegen ein Codon, das für dieselbe Aminosäure, für eine konservative Aminosäure oder für eine nicht-konservative Aminosäure kodiert, zu ersetzen. Alternativ dazu wird es durchgeführt, um Nukleinsäure-Segmente mit erwünschten Funktionen aneinander zu fügen, um eine gewünschte Kombination von Funktionen zu erzeugen. Diese künstliche Kombination wird oftmals entweder durch chemische Synthese oder durch künstliche Manipulation von isolierten Segmenten von Nukleinsäuren erreicht, z. B. durch genetische Modifikationsverfahren. Wenn ein rekombinantes Polynukleotid für ein Polypeptid kodiert, kann die Sequenz des kodierten Polypeptids natürlich vorkommend sein ("Wildtyp"), oder es kann sich um eine Variante (z. B. eine Mutante) der natürlich vorkommenden Sequenz handeln. Somit bezieht sich der Begriff "rekombinantes" Polypeptid nicht notwendigerweise auf ein Polypeptid, dessen Sequenz nicht natürlich vorkommt. Vielmehr wird ein "rekombinantes" Polypeptid von einer rekombinanten DNA-Sequenz kodiert, wobei die Sequenz des Polypeptids natürlich vorkommen kann ("Wildtyp") oder nicht natürlich vorkommen kann (z. B. eine Variante, eine Mutante, usw.). Somit ist ein "rekombinantes" Polypeptid das Ergebnis eines menschlichen Eingriffs, kann jedoch eine natürlich vorkommende Aminosäuresequenz sein.

[0119] Ein "Vektor" oder "Expressionsvektor" ist ein Replikon, wie beispielsweise ein Plasmid, ein Phage, ein Virus oder ein Cosmid, an das ein anderes DNA-Segment, d. h. ein "Insert" angeheftet werden kann, so dass es zur Replikation des angehefteten Segments in einer Zelle kommt.

[0120] Eine "Expressions-Kassette" umfasst eine kodierende DNA-Sequenz, die in funktionsfähiger Weise mit einem Promotor verknüpft ist. "In funktionsfähiger Weise verknüpft" bezeichnet eine Aneinanderreihung, bei der die so beschriebenen Bestandteile in einer Beziehung zueinander stehen, die es diesen ermöglicht, in der beabsichtigten Weise zu funktionieren. Beispielsweise ist ein Promotor in funktionsfähiger Weise mit einer kodierenden Sequenz verknüpft, wenn der Promotor die Transkription oder Expression dieser Sequenz bewirkt.

[0121] Die Begriffe "rekombinanter Expressionsvektor" oder DNA-Konstrukt" werden vorliegend austauschbar verwendet und beziehen sich auf ein DNA-Molekül, das einen Vektor und mindestens einen Insert umfasst. Rekombinante Expressionsvektoren werden üblicherweise zum Zweck der Expression und/oder der Vervielfältigung des Inserts/der Inserts erzeugt, oder für die Konstruktion weiterer rekombinanter Nukleotidsequenzen. Das Insert/die Inserts kann/können in funktionsfähiger Weise mit einer Promotor-Sequenz verknüpft sein oder nicht, und es/sie kann/können in funktionsfähiger Weise mit regulatorischen DNA-Sequenzen verknüpft sein oder nicht.

[0122] Eine Zelle wurde mit exogener DNA, z. B. einem rekombinanten Expressionsvektor, "genetisch modifiziert" oder "transformiert" oder "transfiziert", wenn solche DNA in die Zelle eingebracht worden ist. Die Gegenwart der exogenen DNA führt zu einer permanenten oder transienten genetischen Veränderung. Die transformierende DNA kann integriert (kovalent verbunden) im Genom der Zelle vorliegen oder nicht. In Prokaryonten-, Hefe- und Säugetierzellen kann die transformierende DNA beispielsweise auf einem episomalen Element, wie beispielsweise auf einem Plasmid, gehalten werden. Bei eukaryontischen Zellen ist eine stabil transformierte Zelle eine solche, bei der die transformierende DNA in ein Chromosom integriert wurde, so dass sie durch Chromosomen-Replikation von Tochterzellen vererbt wird. Diese Stabilität wird durch die Fähigkeit der eukaryontischen Zelle angezeigt, Zelllinien oder Klone zu bilden, die eine Population von Tochterzellen umfassen, welche die transformierende DNA enthält. Ein "Klon" ist eine Population von Zellen, die durch Mitose von einer

einzelnen Zelle oder von einem gemeinsamen Vorläufer abgeleitet ist. Eine "Zelllinie" ist ein Klon einer primären Zelle, die in der Lage ist, in vitro für mehrere Generationen stabil zu wachsen.

[0123] Geeignete Verfahren der genetischen Modifikation (auch als "Transformation" bezeichnet) umfassen z. B. virale Infektion oder Bakteriophagen-Infektion, Transfektion, Konjugation, Protoplastenfusion, Lipofektion, Elektroporation, Kalziumphosphat-Präzipitation, Polyethylenimin(PEI)-vermittelte Transfektion, DEAE-Dextran-vermittelte Transfektion, Liposomen-vermittelte Transfektion, "Particle Gun"-Technologie, Kalziumphosphat-Präzipitation, direkte Mikroinjektion, Nanopartikel-vermittelte Nukleinsäure-Bereitstellung (siehe z. B. Panyam et al., *Adv Drug Deliv Rev.* 2012 Sep 13. pii: 50169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023), und ähnliche.

[0124] Die Wahl des Verfahrens der genetischen Modifikation ist im Allgemeinen vom Typ der Zelle, die transformiert wird, sowie von den Umständen, unter denen die Transformation stattfindet (z. B. in vitro, ex vivo oder in vivo), abhängig. Eine allgemeine Diskussion dieser Verfahren kann in Ausubel, et al., (*Short Protocols in Molecular Biology*, 3. Aufl., Wiley & Sons, 1995) gefunden werden.

[0125] Eine "Ziel-DNA" ist wie vorliegend verwendet ein DNA-Polynukleotid, das eine "Ziel-Stelle" oder eine "Ziel-Sequenz" umfasst. Die Begriffe "Ziel-Stelle" oder "Ziel-Sequenz" oder "Ziel-Protospacer-DNA" werden vorliegend austauschbar verwendet und beziehen sich auf eine Nukleinsäuresequenz, die in einer Ziel-DNA vorhanden ist und an die ein auf DNA gerichtetes Segment einer vorliegenden auf DNA gerichteten RNA binden wird (siehe **Fig. 1** und **Fig. 39**), vorausgesetzt, dass ausreichende Bedingungen für eine Bindung bestehen. Zum Beispiel wird die Ziel-Stelle (oder die Ziel-Sequenz) 5'-GAGCATATC-3' (SEQ ID NO: //) innerhalb einer Ziel-DNA von der RNA Sequenz 5'-GAUAUGCUC-3' (SEQ ID NO: //) anvisiert (oder sie wird von dieser gebunden oder hybridisiert mit dieser oder ist komplementär dazu). Geeignete DNA/RNA-Bindungsbedingungen umfassen physiologische Bedingungen, die normalerweise in der Zelle vorliegen. Andere geeignete DNA/RNA-Bindungsbedingungen (z. B. Bedingungen in einem zellfreien System) sind im Stand der Technik bekannt; siehe z. B. Sambrook, siehe oben. Der Strang der Ziel-DNA, der zu der auf DNA gerichteten RNA komplementär ist und mit dieser hybridisiert, wird als "komplementärer Strang" bezeichnet, und der Strang der Ziel-DNA, der komplementär zum "komplementären Strang" ist (und daher nicht komplementär zu der auf DNA gerichteten RNA ist) wird als "nichtkomplementärer Strang" oder "nicht-komplementärer Strang" bezeichnet (siehe **Fig. 12**).

[0126] Mit "ortsspezifisch modifizierendem Polypeptid" oder "RNA-bindendem, ortsspezifischem Polypeptid" oder "RNA-bindendem ortsspezifisch modifizierendem Polypeptid" oder "ortsspezifischem Polypeptid" ist ein Polypeptid gemeint, das RNA bindet und auf eine spezifische DNA-Sequenz gerichtet ist. Ein wie vorliegend beschriebenes, ortsspezifisch modifizierendes Polypeptid wird durch das RNA-Molekül, an das es gebunden ist, zu einer spezifischen DNA-Sequenz geführt. Das RNA-Molekül umfasst eine Sequenz, die komplementär zu einer Ziel-Sequenz innerhalb der Ziel-DNA ist, wodurch das gebundene Polypeptid zu einer spezifischen Stelle innerhalb der Ziel-DNA geführt wird (der Ziel-Sequenz).

[0127] Mit "Spaltung" ist das Aufbrechen des kovalenten Rückgrats eines DNA-Moleküls gemeint. Die Spaltung kann durch eine Vielzahl von Verfahren initiiert werden, einschließlich (jedoch nicht beschränkt auf) enzymatische oder chemische Hydrolyse einer Phosphodiesterbindung. Sowohl einzelsträngige Spaltung als auch doppelsträngige Spaltung ist möglich, und die doppelsträngige Spaltung kann als Ergebnis von zwei unterschiedlichen einzelsträngigen Spaltungsereignissen auftreten. Die DNA-Spaltung kann entweder zur Erzeugung von stumpfen Enden oder versetzten Enden führen. In einigen Ausführungsformen wird ein Komplex, der eine auf DNA gerichtete RNA und ein ortsspezifisch modifizierendes Polypeptid umfasst, für die gezielte Spaltung doppelsträngiger DNA verwendet.

[0128] "Nuklease" und "Endonuklease" werden vorliegend austauschbar verwendet und verweisen auf ein Enzym, das katalytische Aktivität für eine DNA-Spaltung besitzt.

[0129] Mit "Spaltungs-Domäne" oder "aktiver Domäne" oder "Nuklease-Domäne" einer Nuklease ist die Polypeptidsequenz oder Polypeptid-Domäne innerhalb einer Nuklease gemeint, welche die katalytische Aktivität für die Spaltung der DNA besitzt. Eine Spaltungs-Domäne kann in einer einzelnen Polypeptid-Kette enthalten sein, oder die Spaltungs-Aktivität kann aus der Assoziation von zwei (oder mehreren) Polypeptiden resultieren. Eine einzelne Nuklease-Domäne kann aus mehr als einem isolierten Abschnitt von Aminosäuren innerhalb eines bestimmten Polypeptids bestehen.

[0130] Das RNA-Molekül, welches an das ortsspezifisch modifizierende Polypeptid bindet und das Polypeptid zu einer spezifischen Stelle innerhalb der Ziel-DNA führt, wird vorliegend als die "auf DNA gerichtete RNA" oder "auf DNA gerichtetes RNA-Polynukleotid" (vorliegend auch "guide RNA" oder "gRNA" genannt) bezeichnet. Eine vorliegende auf DNA gerichtete RNA umfasst zwei Segmente, ein "auf DNA gerichtetes Segment" und ein "Protein-bindendes Segment". Mit "Segment" ist ein Segment/ein Abschnitt/eine Region eines Moleküls gemeint, z. B. ein zusammenhängender Abschnitt von Nukleotiden in einer RNA. Ein Segment kann auch eine Region/ein Abschnitt eines Komplexes sein, so dass ein Segment Regionen aus mehr als einem Molekül umfassen kann. Beispielsweise ist in einigen Fällen das Protein-bindende Segment (nachfolgend beschrieben) einer auf DNA gerichteten RNA ein einzelnes RNA-Molekül, und das Protein-bindende Segment umfasst daher eine Region dieses RNA-Moleküls. In anderen Fällen umfasst das Protein-bindende Segment (nachfolgend beschrieben) einer auf DNA gerichteten RNA zwei unterschiedliche Moleküle, die mittels eines komplementären Abschnitts miteinander hybridisiert sind. Als veranschaulichendes, nicht-einschränkendes Beispiel kann ein Protein-bindendes Segment einer auf DNA gerichteten RNA, das zwei getrennte Moleküle umfasst, Folgendes umfassen: (i) Basenpaare 40–75 eines ersten RNA-Moleküls, das 100 Basenpaare lang ist; und (ii) Basenpaare 10–25 eines zweiten RNA-Moleküls, das 50 Basenpaare lang ist. Die Definition von "Abschnitt" ist – sofern sie nicht anderweitig spezifisch in einem bestimmten Zusammenhang definiert ist – nicht auf eine spezifische Anzahl von Gesamt-Basenpaaren beschränkt, ist nicht auf irgendeine bestimmte Anzahl von Basenpaaren von einem bestimmten RNA-Molekül beschränkt, ist nicht auf eine bestimmte Anzahl von getrennten Molekülen in einem Komplex beschränkt, und kann Abschnitte aus RNA-Molekülen umfassen, die eine beliebige Gesamtlänge aufweisen, und kann Abschnitte mit Komplementarität zu anderen Molekülen umfassen oder nicht.

[0131] Das auf DNA gerichtete Segment (oder die "auf DNA gerichtete Sequenz") umfasst eine Nukleotidsequenz, die komplementär zu einer spezifischen Sequenz innerhalb einer Ziel-DNA ist (der komplementäre Strang der Ziel-DNA). Das Protein-bindende Segment (oder die "Protein-bindende Sequenz") interagiert mit einem ortsspezifisch modifizierenden Polypeptid. Wenn das ortsspezifisch modifizierende Polypeptid ein Cas9 oder ein mit Cas9 verwandtes Polypeptid ist (im Folgenden genauer erklärt), tritt die ortsspezifische Spaltung der Ziel-DNA an Stellen auf, die sowohl durch (i) Basenpaarungs-Komplementarität zwischen der auf DNA gerichteten RNA und der Ziel-DNA; als auch durch (ii) ein kurzes Motiv (als Protospacer-benachbartes Motiv (Protospacer adjacent motive, PAM) bezeichnet) bestimmt werden.

[0132] Das Protein-bindende Segment einer vorliegenden auf DNA gerichteten RNA umfasst zwei komplementäre Abschnitte von Nukleotiden, die miteinander hybridisieren, wobei ein doppelsträngiger DNA-Duplex (dsRNA-Duplex) gebildet wird.

[0133] In einigen Ausführungsformen umfasst eine vorliegende Nukleinsäure (z. B. eine auf DNA gerichtete RNA, eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; eine Nukleinsäure, die für ein ortsspezifisches Polypeptid kodiert usw.) eine Modifikation oder eine Sequenz, die für zusätzliche gewünschte Eigenschaft sorgt (z. B. modifizierte oder regulierte Stabilität; subzelluläres Targeting; Tracking, z. B. eine fluoreszente Markierung; eine Bindungsstelle für ein Protein oder für einen Proteinkomplex). Nicht-einschränkende Beispiele umfassen: eine 5'-Kappe (z. B. eine 7-Methylguanylat-Kappe (m7G)); einen 3'-polyadenylierten Anhang (z. B. einen 3'-Poly(A)-Anhang); eine Riboswitch-Sequenz (z. B. um eine regulierte Stabilität und/oder eine regulierte Zugänglichkeit für Proteine und/oder Protein-Komplexe zu ermöglichen); eine Stabilitäts-Kontrollsequenz; eine Sequenz, die einen dsRNA-Duplex bildet (z. B. eine Haarnadelschleife); eine Modifikation oder Sequenz, welche die RNA gezielt zu einer subzellulären Stelle führt (z. B. Nukleus, Mitochondrien, Chloroplasten, und ähnliche); eine Modifikation oder Sequenz, die für ein Tracking sorgt (z. B. direkte Konjugation an ein fluoreszentes Molekül, Konjugation an einen Rest, der eine fluoreszente Detektion ermöglicht, eine Sequenz, die eine fluoreszente Detektion ermöglicht, usw.); eine Modifikation oder Sequenz, die eine Bindungsstelle für Proteine bereitstellt (z. B. Proteine, die auf DNA wirken, einschließlich Transkriptions-Aktivatoren, Transkriptions-Repressoren, DNA-Methyltransferasen, DNA-Demethylasen, Histon-Acetyltransferasen, Histon-Deacetylasen, und ähnliche) und Kombinationen derselben.

[0134] In einigen Ausführungsformen umfasst eine auf DNA gerichtete RNA entweder am 5'-Ende oder am 3'-Ende ein zusätzliches Segment, das für eines der oben beschriebenen Merkmale sorgt. Beispielsweise kann ein geeignetes drittes Segment eine 5'-Kappe (z. B. eine 7-Methylguanylat-Kappe (m7G)); einen 3'-polyadenylierten Anhang (z. B. einen 3'-Poly(A)-Anhang); eine Riboswitch-Sequenz (z. B. um eine regulierte Stabilität und/oder eine regulierte Zugänglichkeit für Proteine und/oder Protein-Komplexe zu ermöglichen); eine Stabilitäts-Kontrollsequenz; eine Sequenz, die einen dsRNA-Duplex bildet (z. B. eine Haarnadelschleife); eine Sequenz, welche die RNA gezielt zu einer subzellulären Stelle führt (z. B. Nukleus, Mitochondrien, Chloroplasten, und ähnliche); eine Modifikation oder Sequenz, die für ein Tracking sorgt (z. B. direkte Konjugation an ein

fluoreszentes Molekül, Konjugation an einen Rest, der eine fluoreszente Detektion ermöglicht, eine Sequenz, die eine fluoreszente Detektion ermöglicht, usw.); eine Modifikation oder Sequenz, die eine Bindungsstelle für Proteine bereitstellt (z. B. Proteine, die auf DNA wirken, einschließlich Transkriptions-Aktivatoren, Transkriptions-Repressoren, DNA-Methyltransferasen, DNA-Demethylasen, Histon-Acetyltransferasen, Histon-Deacetylase, und ähnliche) und Kombinationen derselben.

[0135] Eine vorliegende auf DNA gerichtete RNA und ein vorliegendes ortsspezifisch modifizierendes Polypeptid (d. h. ein ortsspezifisches Polypeptid) bilden einen Komplex (d. h. binden über nicht-kovalente Interaktionen). Die auf DNA gerichtete RNA sorgt für eine Zielspezifität des Komplexes, indem sie eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz der Ziel-DNA ist. Das ortsspezifisch modifizierende Polypeptid des Komplexes sorgt für eine ortsspezifische Aktivität. Mit anderen Worten wird das ortsspezifisch modifizierende Polypeptid durch seine Assoziation mit dem Protein-bindenden Segment der auf DNA gerichteten RNA an eine DNA-Ziel-Sequenz (z. B. eine Ziel-Sequenz in einer chromosomalen Nukleinsäure, eine Ziel-Sequenz in einer extrachromosomalen Nukleinsäure, z. B. einer episomalen Nukleinsäure, einem minizirkulären Molekül, usw.; einer Ziel-Sequenz in einer mitochondrialen Nukleinsäure; einer Ziel-Sequenz in einer Chloroplasten-Nukleinsäure, einer Ziel-Sequenz in einem Plasmid; usw.) geführt.

[0136] In einigen Ausführungsformen umfasst die auf DNA gerichtete RNA zwei unterschiedliche RNA-Moleküle (RNA-Polynukleotide: eine "Aktivator-RNA" und eine "Targeter-RNA", siehe im Folgenden) und wird vorliegend als "auf DNA gerichtetes RNA-Doppelmolekül" oder als "aus zwei Molekülen bestehende, auf DNA gerichtete RNA" bezeichnet. In anderen Ausführungsformen ist die auf DNA gerichtete RNA ein RNA-Einzelmolekül (einzelnes RNA-Polynukleotid) und wird vorliegend als "auf DNA gerichtetes RNA-Einzelmolekül" oder als "single guide RNA" oder als "sgRNA" bezeichnet. Der Begriff "auf DNA gerichtete RNA" oder "gRNA" ist einschließend gemeint und bezeichnet sowohl auf DNA gerichtete RNA-Doppelmoleküle als auch auf DNA gerichtete RNA-Einzelmoleküle (d. h. sgRNAs).

[0137] Ein beispielhaftes auf DNA gerichtetes RNA-Doppelmolekül umfasst ein crRNA-ähnliches ("CRISPR-RNA-" oder "Targeter-RNA-" oder crRNA-" oder "crRNA Repeat-")Molekül und ein entsprechendes tracrRNA-ähnliches ("trans-wirkendes CRISPR-RNA-" oder "Aktivator-RNA-" oder "tracrRNA-")Molekül. Ein crRNA-ähnliches Molekül (Targeter-RNA) umfasst sowohl das auf DNA gerichtete Segment (einzelsträngig) der auf DNA gerichteten RNA als auch einen Abschnitt ("Duplex-bildendes Segment") von Nukleotiden, der eine Hälfte des dsRNA-Duplex des Protein-bindenden Segments der auf DNA gerichteten RNA bildet. Ein entsprechendes tracrRNA-ähnliches Molekül (Aktivator-RNA) umfasst einen Abschnitt von Nukleotiden (Duplex-bildendes Segment), der die andere Hälfte des dsRNA-Duplex des Protein-bindenden Segments der auf DNA gerichteten RNA bildet. Mit anderen Worten ist ein Abschnitt von Nukleotiden des crRNA-ähnlichen Moleküls komplementär zu und hybridisiert mit einem Abschnitt von Nukleotiden eines tracrRNA-ähnlichen Moleküls, wobei der dsRNA-Duplex der Protein-bindenden Domäne der auf DNA gerichteten RNA gebildet wird. Somit kann behauptet werden, dass jedes crRNA-ähnliche Molekül ein entsprechendes tracrRNA-ähnliches Molekül hat. Das crRNA-ähnliche Molekül stellt darüber hinaus das einzelsträngige, auf DNA gerichtete Segment bereit. Somit hybridisieren ein crRNA-ähnliches und ein tracrRNA-ähnliches Molekül (als entsprechendes Paar), wodurch eine auf DNA gerichtete RNA gebildet wird. Die exakte Sequenz eines bestimmten crRNA- oder tracrRNA-Moleküls ist charakteristisch für die Spezies, in der die RNA-Moleküle gefunden werden. Zahlreiche crRNAs und tracrRNAs sind in entsprechenden komplementären Paaren in **Fig. 8** gezeigt. Ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül kann ein beliebiges Paar aus crRNA und tracrRNA umfassen. Ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül kann ein beliebiges Paar aus crRNA und tracrRNA umfassen.

[0138] Der Begriff "Aktivator-RNA" wird vorliegend verwendet, um ein tracrRNA-ähnliches Molekül eines auf DNA gerichteten RNA-Doppelmoleküls zu bezeichnen. Der Begriff "Targeter-RNA" wird vorliegend verwendet, um ein crRNA-ähnliches Molekül eines auf DNA gerichteten RNA-Doppelmoleküls zu bezeichnen. Der Begriff "Duplex-bildendes Segment" wird vorliegend verwendet, um einen Abschnitt von Nukleotiden einer Aktivator-RNA oder einer Targeter-RNA zu bezeichnen, der zur Bildung des dsRNA-Duplex durch Hybridisierung an einen Abschnitt von Nukleotiden des entsprechenden Aktivator-RNA- oder Targeter-RNA-Moleküls beiträgt. Mit anderen Worten umfasst eine Aktivator-RNA ein Duplex-bildendes Segment, das komplementär zum Duplex-bildenden Segment der entsprechenden Targeter-RNA ist. Somit umfasst eine Aktivator-RNA ein Duplex-bildendes Segment, während eine Targeter-RNA sowohl ein Duplex-bildendes Segment als auch das auf DNA gerichtete Segment der auf DNA gerichteten RNA enthält. Somit kann ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül aus einem beliebigen korrespondierenden Paar von Aktivator-RNA und Targeter-RNA bestehen.

[0139] Eine "Wirtszelle" bezeichnet wie vorliegend verwendet eine in vivo oder in vitro vorliegende eukaryontische Zelle, eine prokaryontische Zelle (z. B. eine bakterielle oder eine Archaeen-Zelle), oder eine Zelle von einem multizellulären Organismus (z. B. eine Zelllinie), die als einzellige Einheit kultiviert wurde, wobei die eukaryontischen oder prokaryontischen Zellen als Rezipienten für Nukleinsäure verwendet werden können oder für diesen Zweck verwendet wurden, und die Nachkommen der ursprünglichen Zelle umfassen, die mit der Nukleinsäure transformiert wurde. Es sollte verstanden werden, dass die Nachkommen einer einzelnen Zelle hinsichtlich der Morphologie oder des genomischen oder Gesamt-DNA-Komplements aufgrund von natürlicher, zufälliger oder gewollter Mutation nicht notwendigerweise vollständig identisch zum ursprünglichen Elter sein müssen. Eine "rekombinante Wirtszelle" (vorliegend auch als "genetisch modifizierte Wirtszelle" bezeichnet) ist eine Wirtszelle, in die eine heterologe Nukleinsäure eingebracht wurde, z. B. ein Expressionsvektor. Beispielsweise ist eine vorliegende bakterielle Wirtszelle dadurch eine genetisch modifizierte Wirtszelle, dass man eine exogene Nukleinsäure (z. B. ein Plasmid oder einen rekombinanten Expressionsvektor) in eine geeignete bakterielle Wirtszelle eingebracht hat, und eine vorliegende eukaryontische Wirtszelle ist dadurch eine genetisch modifizierte eukaryontische Wirtszelle (z. B. eine Säugetier-Keimzelle), dass man eine exogene Nukleinsäure in eine geeignete eukaryontische Wirtszelle eingebracht hat.

[0140] Der Begriff "Stammzelle" wird vorliegend verwendet, um eine Zelle (z. B. eine pflanzliche Stammzelle, eine Vertebraten-Stammzelle) zu bezeichnen, die die Fähigkeit aufweist, sich sowohl selbst zu erneuern als auch einen differenzierten Zelltyp zu erzeugen (siehe Morrison et al. (1997) Cell 88: 287–298). Im Zusammenhang mit der Zell-Ontogenie ist das Adjektiv "differenziert" oder "differenzieren" ein relativer Begriff. Eine "differenzierte Zelle" ist eine Zelle, die im Vergleich mit einer anderen Zelle mehr Entwicklungsstufen durchlaufen hat. Somit können sich pluripotente Stammzellen (nachfolgend beschrieben) in durch die Abstammungslinie beschränkte Vorläuferzellen (z. B. mesodermale Stammzellen) differenzieren, welche sich wiederum in Zellen differenzieren können, die noch stärker eingeschränkt sind (z. B. Vorläufer-Neuronen), welche sich in die Zellen des Endstadiums differenzieren können (z. B. in terminal differenzierte Zellen, z. B. Neuronen, Kardiomyozyten, usw.), die eine charakteristische Rolle in einigen Arten von Geweben spielen, und die die Fähigkeit beibehalten haben, weiter zu proliferieren, oder diese Fähigkeit nicht beibehalten haben. Stammzellen können sowohl durch das Vorhandensein spezifischer Marker (z. B. Proteine, RNAs, usw.) als auch durch das Fehlen spezifischer Marker charakterisiert sein. Stammzellen können ferner durch Funktionsassays sowohl in vitro als auch in vivo identifiziert werden, insbesondere durch Assays, die auf die Fähigkeit von Stammzellen abstellen, mehrere differenzierte Nachkommen zu erzeugen.

[0141] Stammzellen von Interesse umfassen pluripotente Stammzellen (PSCs). Der Begriff "pluripotente Stammzellen" oder "PSC" wird vorliegend verwendet, um eine Stammzelle zu bezeichnen, die in der Lage ist, alle Zell-Typen des Organismus zu erzeugen. Daher kann eine PSC Zellen aller Keimblätter des Organismus erzeugen (z. B. das Endoderm, das Mesoderm und das Ektoderm eines Säugetiers). Pluripotente Zellen sind in der Lage, Teratome zu bilden und in einem lebenden Organismus zu Ektoderm-, Mesoderm- oder Endoderm-Gewebe beizutragen. Pluripotente Stammzellen von Pflanzen sind in der Lage, alle Zellarten der Pflanze zu erzeugen (z. B. Zellen der Wurzel, des Stamms, der Blätter, usw.).

[0142] PSCs von Tieren können auf verschiedenen Wegen erhalten werden. Beispielsweise werden embryonische Stammzellen (ESCs) von der inneren Zellmasse eines Embryos erhalten (Thomson et. al, Science. 1998 Nov 6; 282(5391): 1145–7), während induzierte pluripotente Stammzellen (iPSCs) von somatischen Zellen erhalten werden (Takahashi et. al, Cell. 2007 Nov 30; 131(5): 861–72; Takahashi et. al, Nat Protoc. 2007; 2(12): 3081–9; Vu et. al, Science. 2007 Dez 21; 318(5858): 1917–20. Epub 2007 Nov 20). Da der Begriff PSC pluripotente Stammzellen unabhängig von deren Herkunft bezeichnet, umfasst der Begriff PSC die Begriffe ESC und iPSC sowie auch den Begriff embryonische Keimstammzellen (embryonic germ stem cells, EGSC), welche ein weiteres Beispiel für PSC sind. PSCs könne in Form einer etablierten Zelllinie vorliegen, sie können direkt von einem primären embryonischen Gewebe erhalten werden, oder sie können von einer somatischen Zelle abgeleitet sein. PSCs können Zielzellen der vorliegend beschriebenen Verfahren sein.

[0143] Mit "embryonischer Stammzelle" (ESC) ist eine PSC gemeint, die von einem Embryo, üblicherweise von der inneren Zellmasse einer Blastocyste, isoliert wurde. ESC-Zelllinien werden in dem NIH-Register für humane embryonische Stammzellen geführt, z. B. hESBGN-01, hESBGN-02, hESBGN-03, hESBGN-04 (BresaGen, Inc.); HES-1, HES-2, HES-3, HES-4, HES-5, HES-6 (ES Cell International); Miz-hES1 (MizMedi Hospital-Nationale Universität Seoul); HSF-1, HSF-6 (Universität von Kalifornien in San Francisco); und H1, H7, H9, H13, H14 (Wisconsin Alumni Research Foundation (WiCell Research Institute)). Stammzellen von Interesse umfassen auch embryonische Stammzellen von anderen Primaten, wie beispielsweise Rhesus-Stammzellen und Stammzellen vom Pinseläffchen. Die Stammzellen können von beliebigen Säugetier-Spezies erhalten werden, z. B. von Mensch, Pferd, Rind, Schwein, Hund, Katze, Nagetier, z. B. Maus, Ratte, Hamster, Primat

usw. (Thomson et al. (1998) Science 282: 1145; Thomson et al. (1995) Proc. Natl. Acad. Sci. USA 92: 7844; Thomson et al. (1996) Biol. Reprod. 55: 254; Shambloott et al., Proc. Natl. Acad. Sci. USA 95: 13726, 1998). In Kultur wachsen ESCs üblicherweise als flache Kolonien mit großem Verhältnis von Kern zu Zytoplasma, mit definierten Grenzen und hervorgehobenen Kernkörperchen. Darüber hinaus exprimieren ESCs SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, und alkalische Phosphatase, nicht jedoch SSEA-1. Beispiele für Verfahren zur Herstellung und Charakterisierung von ESCs können z. B. in US Patent Nr. 7,029,913, US Patent Nr. 5,843,780 und US Patent Nr. 6,200,806 gefunden werden, wobei deren Offenbarung vorliegend durch Bezugnahme eingeschlossen ist. Verfahren zur Vermehrung von hESCs in undifferenzierter Form werden in WO 99/20741, WO 01/51616 und WO 03/020920 beschrieben.

[0144] Mit "embryonischer Keimstammzelle" (EGSC) oder "embryonischer Keimzelle" oder "EG-Zelle" ist eine PSC gemeint, die von Keimzellen und/oder Vorläufer-Keimzellen abgeleitet ist, z. B. von primordialen Keimzellen, d. h. solchen, aus denen Spermien und Eier entstehen würden. Von embryonischen Keimzellen (EG-Zellen) wird angenommen, dass die Eigenschaften haben, die denen der oben beschriebenen embryonischen Stammzellen ähneln. Beispiele für Verfahren zur Herstellung und Charakterisierung von EG-Zellen können beispielsweise in US Patent Nr. 7,153,684; Matsui, Y., et al., (1992) Cell 70: 841; Shambloott, M., et al. (2001) Proc. Natl. Acad. Sci. USA 98: 113; Shambloott, M., et al. (1998) Proc. Natl. Acad. Sci. USA, 95: 13726; und Koshimizu, U., et al. (1996) Development, 122: 1235, gefunden werden, wobei deren Offenbarung vorliegend durch Bezugnahme eingeschlossen ist.

[0145] Mit "induzierten pluripotenten Stammzellen" oder "iPSC" ist eine PSC gemeint, die von einer Zelle abgeleitet ist, bei der es sich nicht um eine PSC handelt (d. h. von einer Zelle, die im Vergleich zu einer PSC differenziert ist). iPSCs können von verschiedenen Zellarten erhalten werden, einschließlich terminal differenzierter Zellen. iPSCs haben eine ähnliche Morphologie wie ES-Zellen, wachsen als flache Kolonien mit großem Verhältnis von Kern zu Zytoplasma, mit definierten Grenzen und hervorgehobenem Kernkörperchen. Darüber hinaus exprimieren iPSCs einen oder mehrere Pluripotenz-Marker, die dem Fachmann bekannt sind, einschließlich (jedoch nicht beschränkt aus) alkalischer Phosphatase, SSEA3, SSEA4, Sox2, Oct3/4, Nanog, TRA160, TRA181, TDGF 1, Dnmt3b, FoxD3, GDF3, Cyp26a1, TERT, und zfp42. Beispiele für Verfahren zur Herstellung und Charakterisierung von iPSCs können beispielsweise in den US-Patent-Veröffentlichungsnummern US20090047263, US20090068742, US20090191159, US20090227032, US20090246875, und US20090304646 gefunden werden, wobei deren Offenbarung durch Bezugnahme vorliegend eingeschlossen ist. Im Allgemeinen werden zur Herstellung von iPSCs somatische Zellen mit Reprogrammierungsfaktoren (z. B. Oct4, SOX2, KLF4, MYC, Nanog, Lin28, usw.) in Kontakt gebracht, von denen bekannt ist, dass sie somatische Zellen in pluripotente Stammzellen reprogrammieren.

[0146] Mit "somatischer Zelle" ist eine beliebige Zelle in einem Organismus gemeint, die ohne experimentelle Manipulation normalerweise nicht alle Zellarten eines Organismus erzeugen kann. Mit anderen Worten sind somatische Zellen solche Zellen, die ausreichend differenziert sind, so dass sie nicht natürlicherweise Zellen aller drei Keimblätter des Körpers, d. h. Ektoderm, Mesoderm und Endoderm, erzeugen. Beispielsweise würden somatische Zelle sowohl Neuronen als auch neuronale Vorläufer umfassen, wobei letztere in der Lage sein könnten, natürlicherweise alle oder einige Zellarten des zentralen Nervensystems zu bilden, nicht jedoch Zellen der Mesoderm- oder Endoderm-Abstammungslinie.

[0147] Mit "mitotischer Zelle" ist eine Zelle gemeint, welche die Mitose durchläuft. Die Mitose ist ein Prozess, durch den eine eukaryontische Zelle die Chromosomen in ihrem Zellkern in zwei identische Gruppen in zwei getrennten Zellkernen auftrennt. Ihr folgt im Allgemeinen unmittelbar die Cytokinese, welche Zellkern, Zytoplasma, Organellen und Zellmembran in zwei Zellen trennt, die etwa gleiche Anteile dieser zellulären Komponenten enthalten.

[0148] Mit "post-mitotischer Zelle" ist eine Zelle gemeint, welche die Mitose verlassen hat, d. h. die "ruhend" ist, d. h. die keine Teilungen mehr durchläuft. Dieser ruhende Zustand kann temporär sein, d. h. reversibel, oder er kann permanent sein.

[0149] Mit "meiotischer Zelle" ist eine Zelle gemeint, welche die Meiose durchläuft. Meiose ist der Prozess, durch den eine Zelle ihr nukleares Material zum Zwecke der Herstellung von Gameten oder Sporen auftrennt. Anders als bei der Mitose durchlaufen die Chromosomen bei der Meiose einen Rekombinationsschritt, der genetisches Material zwischen Chromosomen mischt. Darüber hinaus stellt die Meiose als Ergebnis 4 (genetisch einzigartige) haploide Zellen bereit, im Vergleich zu 2 (genetisch identischen) diploiden Zellen, die bei der Mitose erzeugt werden.

[0150] Mit "Rekombination" ist ein Austauschprozess von genetischer Information zwischen zwei Polynukleotiden gemeint. Wie vorliegend verwendet bezeichnet eine "Homologie-gesteuerte Reparatur (homology-directed repair, HDR)" die spezialisierte Form einer DNA-Reparatur, welche beispielsweise während der Reparatur von Doppelstrangbrüchen in Zellen stattfindet. Dieser Prozess erfordert Nukleotidsequenz-Homologie, verwendet ein "Donor-Molekül" als Matrize für die Reparatur eines "Ziel-Moleküls" (d. h. des Moleküls, das den Doppelstrangbruch erlitten hat) und führt zum Transfer von genetischer Information vom Donor zum Ziel. Die Homologie-gesteuerte Reparatur kann zu einer Änderung der Sequenz des Ziel-Moleküls führen (z. B. Insertion, Deletion, Mutation), wenn sich das Donor-Polynukleotid vom Ziel-Molekül unterscheidet und ein Teil der Sequenz oder die gesamte Sequenz des Donor-Polynukleotids in die Ziel-DNA eingebaut wird. In einigen Ausführungsformen integriert das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA.

[0151] Mit "nicht-homologem End-Joining (NHEJ)" ist die Reparatur von Doppelstrangbrüchen durch direkte Ligation der Bruchenden miteinander ohne die Notwendigkeit einer homologen Matrize gemeint (im Gegensatz zur Homologie-gesteuerten Reparatur, die eine homologe Sequenz zur Steuerung der Reparatur erfordert). Das NHEJ führt oftmals zu einem Verlust (Deletion) von Nukleotidsequenz nahe der Stelle des Doppelstrangbruchs.

[0152] Die Begriffe "Behandlung", "behandeln" und ähnliche werden vorliegend verwendet und bezeichnen allgemein das Erreichen einer gewünschten pharmakologischen und/oder physiologischen Wirkung. Die Wirkung kann prophylaktisch sein, sodass sie vollständig oder teilweise eine Erkrankung oder Symptome derselben vermeidet, und/oder sie kann therapeutisch sein, sodass sie eine Erkrankung und/oder nachteilige Wirkungen, die mit dieser Erkrankung im Zusammenhang stehen, teilweise oder vollständig heilt. "Behandlung" deckt wie vorliegend verwendet eine beliebige Behandlung einer Erkrankung oder eines Symptoms in einem Säugetier ab und umfasst: (a) das Vermeiden des Auftretens der Erkrankung oder des Symptoms in einem Subjekt, welches prädisponiert sein kann, die Erkrankung oder das Symptom zu entwickeln, bei dem diese/dieses jedoch noch nicht diagnostiziert wurde; (b) das Inhibieren der Erkrankung oder des Symptoms, d. h. das Anhalten ihrer/seiner Entwicklung; oder (c) die Befreiung von der Erkrankung, d. h. das Verursachen der Rückentwicklung der Erkrankung. Das therapeutische Mittel kann vor, während oder nach Beginn der Erkrankung oder der Verletzung verabreicht werden. Die Behandlung einer vorhandenen Erkrankung, bei der die Behandlung die unerwünschten klinischen Symptome des Patienten stabilisiert oder verringert, ist von besonderem Interesse. Eine solche Behandlung wird vorzugsweise vor dem vollständigen Verlust der Funktion der betroffenen Gewebe durchgeführt. Die Therapie wird vorzugsweise während der symptomatischen Phase der Erkrankung verabreicht, und in einigen Fällen nach der symptomatischen Phase der Erkrankung.

[0153] Die Begriffe "Individuum", "Subjekt", "Wirt" und "Patient" werden vorliegend austauschbar verwendet und bezeichnen ein beliebiges Säugetier-Subjekt, für das eine Diagnose, Behandlung oder Therapie gewünscht ist, vorzugsweise einen Menschen.

[0154] Allgemeine Verfahren der molekularen und zellulären Biochemie können in Standard-Lehrbüchern gefunden werden, wie beispielsweise Molecular Cloning: A Laboratory Manual, 3. Aufl. (Sambrook et al., HaRBor Laboratory Press 2001); Short Protocols in Molecular Biology, 4. Aufl. (Ausubel et al. eds., John Wiley & Sons 1999); Protein Methods (Bollag et al., John Wiley & Sons 1996); Nonviral Vectors for Gene Therapy (Wagner et al. Hrsg., Academic Press 1999); Viral Vectors (Kaplift & Loewy Hrsg., Academic Press 1995); Immunology Methods Manual (I. Lefkovits Hrsg., Academic Press 1997); und Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998), wobei deren Offenbarung vorliegend durch Bezugnahme eingeschlossen ist.

[0155] Bevor die vorliegende Erfindung weiter beschrieben wird, sollte verstanden werden, dass die Erfindung nicht auf bestimmte beschriebene Ausführungsformen beschränkt ist, da diese selbstverständlich variieren können. Es sollte auch verstanden werden, dass die vorliegend verwendete Terminologie lediglich für die Beschreibung bestimmter Ausführungsformen verwendet wird und nicht einschränkend zu verstehen ist, da der Umfang der vorliegenden Erfindung lediglich durch die anliegenden Ansprüche beschränkt wird.

[0156] Wenn ein Bereich von Werten angegeben wird, sollte verstanden werden, dass jeder zwischenliegende Wert zwischen der oberen und unteren Grenze des Bereichs, sowie jeder beliebige andere genannte oder zwischenliegende Wert in dem genannten Bereich, soweit nicht anders angegeben, bis auf das Zehntel der Einheit der unteren Grenze von der vorliegenden Erfindung eingeschlossen wird. Die oberen und unteren Grenzen dieser kleineren Bereiche können unabhängig voneinander von diesen kleineren Bereichen umfasst sein, und sind ebenfalls von der Erfindung umfasst, es sei denn, eine Grenze in dem genannten Bereich

wird spezifisch ausgeschlossen. Wenn der genannte Bereich eine der Grenzen oder beide Grenzen umfasst, sind Bereiche, die entweder eine oder beide der eingeschlossenen Grenzen ausschließen, ebenfalls von der Erfindung umfasst.

[0157] Bestimmte Bereiche werden vorliegend als numerische Bereiche angegeben, denen der Begriff "etwa" vorausgeht. Der Begriff "etwa" wird vorliegend so verwendet, dass er eine wörtliche Basis für die exakte Zahl bietet, der er vorausgeht, sowie auch für eine Zahl, die der Zahl, der er vorausgeht, angenähert ist oder dieser etwa entspricht. Wenn man bestimmt, ob eine Zahl einer spezifisch genannten Zahl angenähert ist oder dieser etwa entspricht, kann es sich bei der angenäherten oder sich nähernden nicht genannten Zahl um eine Zahl handeln, die in dem Zusammenhang, in dem sie genannt wird, ein wesentliches Äquivalent der spezifisch genannten Zahl ist.

[0158] Soweit nicht anderweitig angegeben haben alle vorliegend verwendeten technischen und wissenschaftlichen Begriffe dieselbe Bedeutung, die normalerweise vom Fachmann, an den sich die Erfindung richtet, angenommen würde. Obgleich beliebige Verfahren und Materialien, die den vorliegend beschriebenen Verfahren und Materialien ähneln oder zu diesen äquivalent sind, bei der Ausführung oder der Überprüfung der vorliegenden Erfindung ebenfalls verwendet werden können, werden die bevorzugten Verfahren und Materialien nunmehr beschrieben. Alle genannten Veröffentlichungen sind vorliegend durch Bezugnahme eingeschlossen, so dass die Verfahren und/oder Materialien, in deren Zusammenhang die Veröffentlichungen zitiert werden, offenbart und beschrieben werden.

[0159] Alle in dieser Beschreibung zitierten Veröffentlichungen und Patente sind vorliegend durch Bezugnahme eingeschlossen, als wäre jede einzelne Veröffentlichung oder jedes einzelne Patent spezifisch und individuell als durch Bezugnahme eingeschlossen bezeichnet, und diese sind vorliegend durch Bezugnahme so eingeschlossen, dass die Verfahren und/oder Materialien, in deren Zusammenhang die Veröffentlichungen zitiert werden, offenbart und beschrieben werden. Die Nennung einer beliebigen Veröffentlichung erfolgt als Offenbarung vor dem Anmeldetag und sollte nicht als Eingeständnis verstanden werden, dass die vorliegende Erfindung nicht berechtigt ist, ein früheres Datum als diese Veröffentlichung im Sinne einer früheren Erfindung einzunehmen. Des Weiteren können sich die Daten der Veröffentlichung von den tatsächlichen Daten der Veröffentlichung unterscheiden und müssen im Einzelnen bestätigt werden.

[0160] Es ist anzumerken, dass wie vorliegend verwendet und in den anliegenden Ansprüchen die Singularform "ein/eine/einer" und "der/die/das" die jeweiligen Pluralformen umfassen, sofern der Zusammenhang nicht deutlich auf anderes verweist. Somit umfasst beispielsweise der Verweis auf "ein Polynukleotid" eine Vielzahl solcher Polynukleotide, und der Verweis auf "das Polypeptid" umfasst den Verweis auf eine oder mehrere Polypeptide und Äquivalente davon, die dem Fachmann bekannt sind, usw. Es wird ferner angemerkt, dass die Ansprüche so abgefasst sein können, dass ein optionales Element ausgeschlossen wird. Diese Aussage soll im Zusammenhang mit der Nennung von Elementen eines Anspruchs als vorhergehende Basis für die Verwendung einer ausschließenden Terminologie wie "einzig", "nur" und ähnlichem oder für die Verwendung einer "negativen" Einschränkung dienen.

[0161] Es wird anerkannt, dass bestimmte Merkmale der Erfindung, die aus Klarheitsgründen im Zusammenhang mit getrennten Ausführungsformen beschrieben werden, auch in Kombination in einer einzelnen Ausführungsform bereitgestellt werden können. Umgekehrt können verschiedene Merkmale der Erfindung, die aus Gründen der Übersichtlichkeit im Zusammenhang mit einer einzelnen Ausführungsform beschrieben sind, getrennt bereitgestellt werden oder in einer beliebigen geeigneten Unterkombinationen. Alle Kombinationen der Ausführungsformen, welche die Erfindung betreffen, sind spezifisch von der vorliegenden Erfindung umfasst und werden vorliegend offenbart, als ob jede Kombination einzeln und explizit offenbart würde. Darüber hinaus sind alle Unterkombinationen der verschiedenen Ausführungsformen und Elemente derselben ebenfalls spezifisch von der vorliegenden Erfindung umfasst und werden vorliegend offenbart, als ob jede solche Unterkombination einzeln und explizit offenbart würde.

[0162] Die vorliegend besprochenen Veröffentlichungen werden einzig zum Zwecke ihrer Offenbarung vor dem Anmeldetag der vorliegenden Erfindung bereitgestellt. Nichts sollte vorliegend als Eingeständnis gewertet werden, dass die vorliegende Erfindung nicht berechtigt wäre, ein früheres Datum als diese Publikationen im Sinne einer früheren Erfindung einzunehmen. Darüber hinaus können sich die bereitgestellten Daten der Veröffentlichung von den tatsächlichen Daten der Veröffentlichung unterscheiden, wobei es erforderlich sein könnte, diese unabhängig voneinander zu bestätigen.

DETAILLIERTE BESCHREIBUNG – TEIL I

[0163] Die vorliegende Offenbarung stellt eine auf RNA gerichtete DNA bereit, die eine Targeting-Sequenz umfasst und gemeinsam mit einem modifizierenden Polypeptid für eine ortsspezifische Modifikation einer Ziel-DNA und/oder eines Polypeptids, das mit der Ziel-DNA assoziiert ist, sorgt. Die vorliegende Offenbarung stellt ferner ortsspezifisch modifizierende Polypeptide bereit. Die vorliegende Offenbarung stellt ferner Verfahren für eine ortsspezifische Modifikation einer Ziel-DNA und/oder eines mit der Ziel-DNA assoziierten Polypeptids bereit. Die vorliegende Offenbarung stellt Verfahren zur Modulierung der Transkription einer Ziel-Nukleinsäure in einer Ziel-Zelle bereit, was im Allgemeinen das Inkontaktbringen einer Ziel-Nukleinsäure mit einem enzymatisch inaktiven Cas9-Polypeptid und einer auf DNA gerichteten RNA umfasst. Kits und Zusammensetzungen zur Ausführung der Verfahren werden ebenfalls bereitgestellt. Die vorliegende Offenbarung stellt genetisch modifizierte Zellen bereit, die Cas9 erzeugen; sowie Cas9-transgene, nicht-humane multizelluläre Organismen.

NUKLEINSÄUREN

Auf DNA gerichtete RNA

[0164] Die vorliegende Offenbarung stellt eine auf DNA gerichtete RNA bereit, die die Aktivitäten eines assoziierten Polypeptids (z. B. eines ortsspezifisch modifizierenden Polypeptids) zu einer spezifischen Ziel-Sequenz innerhalb einer Ziel-DNA leitet. Eine vorliegende auf DNA gerichtete RNA umfasst: ein erstes Segment (vorliegend auch als ein "auf DNA gerichtetes Segment" oder als eine "auf DNA gerichtete Sequenz" bezeichnet) und ein zweites Segment (vorliegend auch als "Protein-bindendes Segment" oder als eine "Protein-bindende Sequenz" bezeichnet).

Auf DNA gerichtetes Segment einer auf DNA gerichteten RNA

[0165] Das auf DNA gerichtete Segment einer vorliegenden auf DNA gerichteten RNA umfasst eine Nukleotidsequenz, die komplementär zu einer Sequenz in einer Ziel-DNA ist. Mit anderen Worten interagiert die auf DNA gerichtete Sequenz einer vorliegenden auf DNA gerichteten RNA mit einer Ziel-DNA in einer sequenzspezifischen Weise mittels Hybridisierung (d. h. Basenpaarung). Somit kann die Nukleotidsequenz des auf DNA gerichteten Segments variieren, und es bestimmt die Stelle innerhalb der Ziel-DNA, an der die auf DNA gerichtete RNA und die Ziel-DNA interagieren werden. Das auf DNA gerichtete Segment einer vorliegenden auf DNA gerichteten RNA kann so modifiziert werden (z. B. durch genetische Modifikation), dass dieses an eine beliebige Sequenz innerhalb der Ziel-DNA hybridisiert.

[0166] Das auf DNA gerichtete Segment kann eine Länge von etwa 12 Nukleotiden bis etwa 100 Nukleotiden aufweisen. Beispielsweise kann das auf DNA gerichtete Segment eine Länge von etwa 12 Nukleotiden (nt) bis etwa 80 Nukleotiden, von etwa 12 nt bis etwa 50 nt, von etwa 12 nt bis etwa 40 nt, von etwa 12 nt bis etwa 30 nt, von etwa 12 nt bis etwa 25 nt, von etwa 12 nt bis etwa 20 nt, oder von etwa 12 nt bis etwa 19 nt aufweisen. Beispielsweise kann das auf DNA gerichtete Segment eine Länge von etwa 19 nt bis etwa 20 nt, von etwa 19 nt bis etwa 25 nt, von etwa 19 nt bis etwa 30 nt, von etwa 19 nt bis etwa 35 nt, von etwa 19 nt bis etwa 40 nt, von etwa 19 nt bis etwa 45 nt, von etwa 19 nt bis etwa 50 nt, von etwa 19 nt bis etwa 60 nt, von etwa 19 nt bis etwa 70 nt, von etwa 19 nt bis etwa 80 nt, von etwa 19 nt bis etwa 90 nt, von etwa 19 nt bis etwa 100 nt, von etwa 20 nt bis etwa 25 nt, von etwa 20 nt bis etwa 30 nt, von etwa 20 nt bis etwa 35 nt, von etwa 20 nt bis etwa 40 nt, von etwa 20 nt bis etwa 45 nt, von etwa 20 nt bis etwa 50 nt, von etwa 20 nt bis etwa 60 nt, von etwa 20 nt bis etwa 70 nt, von etwa 20 nt bis etwa 80 nt, von etwa 20 nt bis etwa 90 nt, von etwa 20 nt bis etwa 100 nt aufweisen. Die Nukleotidsequenz (die auf DNA gerichtete Sequenz) des auf DNA gerichteten Segments, die komplementär zu einer Nukleotidsequenz (Ziel-Sequenz) der Ziel-DNA ist, kann eine Länge von mindestens etwa 12 nt haben. Beispielsweise kann die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, eine Länge von mindestens etwa 12 nt aufweisen, von mindestens etwa 15 nt, von mindestens etwa 18 nt, von mindestens etwa 19 nt, von mindestens etwa 20 nt, von mindestens etwa 25 nt, von mindestens etwa 30 nt, von mindestens etwa 35 nt oder von mindestens etwa 40 nt. Beispielsweise kann die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, eine Länge von etwa 12 Nukleotiden (nt) bis etwa 80 nt, von etwa 12 nt bis etwa 50 nt, von etwa 12 nt bis etwa 45 nt, von etwa 12 nt bis etwa 40 nt, von etwa 12 nt bis etwa 35 nt, von etwa 12 nt bis etwa 30 nt, von etwa 12 nt bis etwa 25 nt, von etwa 12 nt bis etwa 20 nt, von etwa 12 nt bis etwa 19 nt, von etwa 19 nt bis etwa 20 nt, von etwa 19 nt bis etwa 25 nt, von etwa 19 nt bis etwa 30 nt, von etwa 19 nt bis etwa 35 nt, von etwa 19 nt bis etwa 40 nt, von etwa 19 nt bis etwa 45 nt, von etwa 19 nt bis etwa 50 nt oder von etwa 19 nt bis etwa 60 nt, von etwa 20 nt bis etwa 25 nt, von etwa 20 nt bis etwa 30 nt, von etwa 20 nt bis etwa 35 nt, von etwa 20 nt bis etwa 40 nt, von etwa 20 nt bis etwa 45 nt,

von etwa 20 nt bis etwa 50 nt oder von etwa 20 nt bis etwa 60 nt aufweisen. Die Nukleotidsequenz (die auf die DNA gerichtete Sequenz) des auf DNA gerichteten Segments, die komplementär zu einer Nukleotidsequenz (Ziel-Sequenz) der Ziel-DNA ist, kann eine Länge von mindestens etwa 12 nt aufweisen.

[0167] In einigen Fällen ist die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, 20 Nukleotide lang. In einigen Fällen ist die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, 19 Nukleotide lang.

[0168] Die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA kann mindestens 60% betragen (z. B. mindestens 65%, mindestens 70%, mindestens 75%, mindestens 80%, mindestens 85%, mindestens 90%, mindestens 95%, mindestens 97%, mindestens 98%, mindestens 99% oder mindestens 100%). In einigen Fällen ist die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA 100% über die sieben am 5'-Ende gelegenen zusammenhängenden Nukleotide der Ziel-Sequenz des komplementären Strangs der Ziel-DNA. In einigen Fällen ist die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA mindestens 60% über 20 zusammenhängende Nukleotide. In einigen Fällen ist die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA 100% über die vierzehn am 5'-Ende gelegenen zusammenhängenden Nukleotide der Ziel-Sequenz des komplementären Strangs der Ziel-DNA und bis zu 0% über die restlichen. In einigen Fällen kann die auf DNA gerichtete Sequenz als 14 Nukleotide lang angesehen werden (siehe **Fig. 12D–E**). In einigen Fällen ist die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA 100% über die sieben am 5'-Ende gelegenen zusammenhängenden Nukleotide der Ziel-Sequenz des komplementären Strangs der Ziel-DNA und bis zu 0% über die restlichen. In einem solchen Fall kann die auf DNA gerichtete Sequenz als 7 Nukleotide lang angesehen werden.

Protein-bindendes Segment einer auf DNA gerichteten RNA

[0169] Das Protein-bindende Segment einer vorliegenden auf DNA gerichteten RNA interagiert mit einem ortsspezifisch modifizierenden Polypeptid. Die vorliegende auf DNA gerichtete RNA leitet das gebundene Polypeptid über das oben erwähnte auf DNA gerichtete Segment zu einer spezifischen Nukleotidsequenz innerhalb der Ziel-DNA. Das Protein-bindende Segment einer vorliegenden auf DNA gerichteten RNA umfasst zwei Abschnitte aus Nukleotiden, die komplementär zueinander sind. Die komplementären Nukleotide des Protein-bindenden Segments hybridisieren, wodurch ein doppelsträngiger RNA-Duplex (dsRNA) gebildet wird (siehe **Fig. 1A** und **Fig. 1B**).

[0170] Ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül umfasst zwei separate RNA-Moleküle. Jedes dieser zwei RNA-Moleküle eines vorliegenden auf DNA gerichteten RNA-Doppelmoleküls umfasst einen Abschnitt von Nukleotiden, die komplementär zueinander sind, so dass die komplementären Nukleotide der beiden RNA-Moleküle hybridisieren, wodurch der doppelsträngige RNA-Duplex des Protein-bindenden Segments gebildet wird (**Fig. 1A**).

[0171] In einigen Ausführungsformen ist das Duplex-bildende Segment der Aktivator-DNA mindestens etwa 60% identisch zu einem Segment der Aktivator-RNA(tracrRNA)-Moleküle, die in SEQ ID NOs: 431–562 aufgeführt werden, oder zu einem Komplement derselben, über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden. Beispielsweise ist das Duplex-bildende Segment der Aktivator-RNA (oder die DNA, die für das Duplex-bildende Segment der Aktivator-RNA kodiert) mindestens etwa 60% identisch, mindestens etwa 65% identisch, mindestens etwa 70% identisch, mindestens etwa 75% identisch, mindestens etwa 80% identisch, mindestens etwa 85% identisch, mindestens etwa 90% identisch, mindestens etwa 95% identisch, mindestens etwa 98% identisch, mindestens etwa 99% identisch oder mindestens etwa 100% identisch zu einer der tracrRNA-Sequenzen, die in SEQ ID NOs: 431–562 aufgeführt sind, oder zu einem Komplement derselben, über einen Abschnitt von mindestens etwa 8 zusammenhängenden Nukleotiden.

[0172] In einigen Ausführungsformen ist das Duplex-bildende Segment der Targeter-RNA mindestens etwa 60% identisch zu einer der Targeter-RNA(crRNA)-Sequenzen, die in SEQ ID NOs: 563–679 aufgeführt werden, oder zu einem Komplement derselben, über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden. Beispielsweise ist das Duplex-bildende Segment der Targeter-RNA (oder die DNA, die für das Duplex-bildende Segment der Targeter-RNA kodiert) mindestens etwa 65% identisch, mindestens etwa 70%

identisch, mindestens etwa 75% identisch, mindestens etwa 80% identisch, mindestens etwa 85% identisch, mindestens etwa 90% identisch, mindestens etwa 95% identisch, mindestens etwa 98% identisch, mindestens etwa 99% identisch oder mindestens etwa 100% identisch zu einer der crRNA-Sequenzen, die in SEQ ID NOs: 563–679 aufgeführt werden, oder zu einem Komplement derselben, über einen Abschnitt von mindestens etwa 8 zusammenhängenden Nukleotiden.

[0173] Eine auf DNA gerichtete RNA aus zwei Molekülen kann so erstellt werden, dass sie eine kontrollierte (d. h. eine konditionale) Bindung einer Targeter-RNA mit einer Aktivator-RNA ermöglicht. Da eine auf DNA gerichtete RNA aus zwei Molekülen nicht funktionsfähig ist, wenn nicht sowohl die Aktivator-RNA als auch die Targeter-RNA in einen funktionsfähigen Komplex mit dCas9 gebunden ist, kann eine auf DNA gerichtete RNA aus zwei Molekülen induzierbar sein (z. B. Wirkstoff-induzierbar), indem die Bindung zwischen der Aktivator-RNA und der Targeter-RNA induzierbar gemacht wird. Als nicht einschränkendes Beispiel können RNA-Aptamere verwendet werden, um die Bindung der Aktivator-RNA mit der Targeter-RNA zu regulieren (d. h. zu kontrollieren). Demgemäß kann die Aktivator-RNA und/oder die Targeter-RNA eine RNA-Aptamer-Sequenz umfassen.

[0174] RNA-Aptamere sind im Stand der Technik bekannt, und es handelt sich bei diesen im Allgemeinen um eine synthetische Version eines Riboswitches. Die Begriffe "RNA-Aptamer" und "Riboswitch" werden vorliegend austauschbar verwendet und umfassen sowohl synthetische wie auch natürliche Nukleinsäuresequenzen, die für eine induzierbare Regulation der Struktur (und damit der Verfügbarkeit von spezifischen Sequenzen) des DNA-Moleküls sorgen, von dem sie ein Teil sind. RNA-Aptamere umfassen üblicherweise eine Sequenz, die sich in eine bestimmte Struktur faltet (z. B. eine Haarnadelschleife), welche spezifisch an einen bestimmten Wirkstoff (z. B. an ein "Small Molecule") bindet. Die Bindung des Wirkstoffs verursacht eine strukturelle Änderung in der Faltung der RNA, was ein Merkmal der Nukleinsäure, von der das Aptamer ein Teil ist, verändert. Nicht limitierende Beispiele sind die Folgenden: (i) eine Aktivator-RNA mit einem Aptamer könnte nicht in der Lage sein, an die verwandte Targeter-DNA zu binden, wenn nicht das Aptamer durch den entsprechenden Wirkstoff gebunden ist; (ii) eine Targeter-RNA mit einem Aptamer könnte nicht in der Lage sein, an die verwandte Aktivator-RNA zu binden, wenn nicht das Aptamer durch den entsprechenden Wirkstoff gebunden ist; und (iii) eine Targeter-RNA und eine Aktivator-RNA, die jeweils ein unterschiedliches Aptamer umfassen, welches an unterschiedliche Wirkstoffe bindet, könnten nicht in der Lage sein, aneinander zu binden, wenn nicht beide Wirkstoffe vorhanden sind. Wie in diesen Beispielen veranschaulicht wird, kann ein auf DNA gerichtetes RNA-Doppelmolekül so erstellt werden, dass es induzierbar ist.

[0175] Beispiele für Aptamere und Riboswitch-Moleküle können beispielsweise in Nakamura et al., *Genes Cells*. 2012 Mai; 17(5): 344–64; Vavalle et al., *Future Cardiol*. 2012 Mai; 8(3): 371–82; Citartan et al., *Biosens Bioelectron*. 2012 Apr 15; 34(1): 1–11; und Liberman et al., *Wiley Interdiscip Rev RNA*. 2012 Mai–Juni; 3(3): 369–84; gefunden werden, von denen alle vorliegend durch Bezugnahme in ihrer Gesamtheit eingeschlossen sind.

[0176] Nicht-einschränkende Beispiele für Nukleotidsequenzen, die in ein auf DNA gerichtetes RNA-Doppelmolekül eingeschlossen sein können, umfassen eine beliebige der Sequenzen, die in SEQ ID NOs: 431–562 gezeigt sind, oder komplementäre Sequenzen dazu, die mit beliebigen der in SEQ ID NOs: 563–679 gezeigten Sequenzen, oder mit komplementären Sequenzen davon, Basenpaarungen ausbilden, wobei diese hybridisieren können, um ein Protein-bindendes Segment zu bilden.

[0177] Ein vorliegendes auf DNA gerichtetes RNA-Einzelmolekül umfasst zwei Nukleotid-Abschnitte (eine Targeter-RNA und eine Aktivator-RNA), die komplementär zueinander sind, kovalent miteinander über zwischenliegende Nukleotide ("Linker" oder "Linker-Nukleotide") verbunden sind, und hybridisieren, um den doppelsträngigen RNA-Duplex (dsRNA-Duplex) des Protein-bindenden Segments zu bilden, was zu einer Stamm-Schleifen-Struktur führt (**Fig. 1B**). Die Targeter-RNA und die Aktivator-RNA können kovalent über das 3'-Ende der Targeter-RNA und das 5'-Ende der Aktivator RNA miteinander verknüpft sein. Alternativ dazu können Targeter-RNA und die Aktivator-RNA kovalent über das 5'-Ende der Targeter-RNA und das 3'-Ende der Aktivator-RNA miteinander verknüpft sein.

[0178] Der Linker des auf DNA gerichteten RNA-Einzelmoleküls kann eine Länge von etwa 3 Nukleotiden bis etwa 100 Nukleotiden aufweisen. Beispielsweise kann der Linker eine Länge von etwa 3 Nukleotiden (nt) bis etwa 90 nt, von etwa 3 Nukleotiden (nt) bis etwa 80 nt, von etwa 3 Nukleotiden (nt) bis etwa 70 nt, von etwa 3 Nukleotiden (nt) bis etwa 60 nt, von etwa 3 Nukleotiden (nt) bis etwa 50 nt, von etwa 3 Nukleotiden (nt) bis etwa 40 nt, von etwa 3 Nukleotiden (nt) bis etwa 30 nt, von etwa 3 Nukleotiden (nt) bis etwa 20 nt, oder von etwa 3 Nukleotiden (nt) bis etwa 10 nt aufweisen. Beispielsweise kann der Linker eine Länge von etwa 3 nt

bis etwa 5 nt, von etwa 5 nt bis etwa 10 nt, von etwa 10 nt bis etwa 15 nt, von etwa 15 nt bis etwa 20 nt, von etwa 20 nt bis etwa 25 nt, von etwa 25 nt bis etwa 30 nt, von etwa 30 nt bis etwa 35 nt, von etwa 35 nt bis etwa 40 nt, von etwa 40 nt bis etwa 50 nt, von etwa 50 nt bis etwa 60 nt, von etwa 60 nt bis etwa 70 nt, von etwa 70 nt bis etwa 80 nt, von etwa 80 nt bis etwa 90 nt oder von etwa 90 nt bis etwa 100 nt aufweisen. In einigen Ausführungsformen ist der Linker eines auf DNA gerichteten RNA-Einzelmoleküls 4 nt lang.

[0179] Ein beispielhaftes auf DNA gerichtetes RNA-Einzelmolekül umfasst zwei komplementäre Abschnitte von Nukleotiden, die hybridisieren, wobei ein dsRNA-Duplex gebildet wird. In einigen Ausführungsformen ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder die DNA, die für diesen Abschnitt kodiert) mindestens etwa 60% identisch zu einem der Aktivator-RNA-(tracrRNA)-Moleküle, die in SEQ ID NOs: 431–562 aufgeführt werden, oder zu einem Komplement davon, über einen Bereich von mindestens 8 zusammenhängenden Nukleotiden. Beispielsweise ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder die DNA, die für diesen Abschnitt kodiert) mindestens etwa 65% identisch, mindestens etwa 70% identisch, mindestens etwa 75% identisch, mindestens etwa 80% identisch, mindestens etwa 85% identisch, mindestens etwa 90% identisch, mindestens etwa 95% identisch, mindestens etwa 98% identisch, mindestens etwa 99% identisch oder 100% identisch zu einer der tracrRNA-Sequenzen, die in SEQ ID NOs: 431–562 aufgeführt werden, oder zu einem Komplement davon, über einen Bereich von mindestens 8 zusammenhängenden Nukleotiden.

[0180] In einigen Ausführungsformen ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder die DNA, die für diesen Abschnitt kodiert) mindestens etwa 60% identisch zu einer der Targeter-RNA(crRNA)-Sequenzen, die in SEQ ID NOs: 563–679 aufgeführt werden, oder zu einem Komplement davon, über einen Bereich von mindestens 8 zusammenhängenden Nukleotiden. Beispielsweise ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder die DNA, die für diesen Abschnitt kodiert) mindestens etwa 65% identisch, mindestens etwa 70% identisch, mindestens etwa 75% identisch, mindestens etwa 80% identisch, mindestens etwa 85% identisch, mindestens etwa 90% identisch, mindestens etwa 95% identisch, mindestens etwa 98% identisch, mindestens etwa 99% identisch oder 100% identisch zu einer der crRNA-Sequenzen, die in SEQ ID NOs: 563–679 aufgeführt werden, oder zu einem Komplement davon, über einen Bereich von mindestens 8 zusammenhängenden Nukleotiden.

[0181] Geeignete, natürlich vorkommende, verwandte Paare aus crRNAs und tracrRNAs können routinemäßig für SEQ ID NOs: 431–679 bestimmt werden, indem man den Spezies-Namen und die Basenpaarung (für den dsRNA-Duplex der Protein-bindenden Domäne) bei der Bestimmung geeigneter, verwandter Paare berücksichtigt (siehe **Fig. 8** als nichteinschränkendes Beispiel).

[0182] In Bezug sowohl auf ein vorliegendes auf DNA gerichtetes RNA-Einzelmolekül als auch auf ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül zeigt **Fig. 57**, dass künstliche Sequenzen, die nur wenig (etwa 50% Identität) mit natürlich vorkommenden tracrRNAs und crRNAs zeigen, mit Cas9 zusammen wirken können, so dass eine Ziel-DNA gespalten wird, solange die Struktur der Protein-bindenden Domäne der auf DNA gerichteten RNA konserviert ist. Somit kann die RNA-Faltungsstruktur einer natürlich vorkommenden Protein-bindenden Domäne einer auf DNA gerichteten RNA berücksichtigt werden, um künstliche Protein-bindende Domänen zu erstellen (entweder Doppelmolekülversionen oder Einzelmolekülversionen). Als nicht einschränkendes Beispiel wurde die funktionsfähige, künstliche auf DNA gerichtete RNA von **Fig. 57** basierend auf der Struktur des Protein-bindenden Segments beim natürlich vorkommenden DNA-Targeting erstellt (umfasst z. B. die gleiche Anzahl von Basenpaaren im RNA-Duplex und dieselben "Beulen"-Region, die in der natürlich vorkommenden RNA vorhanden ist). Da Strukturen problemlos für beliebige, natürlich vorkommende crRNA:tracrRNA-Paare aus beliebigen Spezies vom Fachmann erzeugt werden können (siehe SEQ ID NOs: 431–679 für crRNA- und tracrRNA-Sequenzen von einer großen Vielzahl von verschiedenen Spezies), kann eine künstliche auf DNA gerichtete RNA erstellt werden, welche die natürliche Struktur für eine bestimmte Spezies nachahmt, wenn das Cas9 (oder ein verwandtes Cas9, siehe **Fig. 32A**) aus dieser Spezies verwendet wird (siehe **Fig. 24D** und verwandte Details in Beispiel 1). Somit kann eine geeignete, auf DNA gerichtete RNA eine künstlich erstellte RNA (nicht natürlich vorkommend) sein, die eine Protein-bindende Domäne umfasst, die so erstellt wurde, dass sie die Struktur einer Protein-bindenden Domäne einer natürlich vorkommenden auf DNA gerichteten RNA nachahmt (siehe SEQ ID NOs: 431–679, wobei der Spezies-Name berücksichtigt wird, wenn geeignete verwandte Paare bestimmt werden).

[0183] Das Protein-bindende Segment kann eine Länge von etwa 10 Nukleotiden bis etwa 100 Nukleotiden aufweisen. Beispielsweise kann das Protein-bindende Segment eine Länge von etwa 15 Nukleotiden (nt) bis

etwa 80 nt, von etwa 15 nt bis etwa 50 nt, von etwa 15 nt bis etwa 40 nt, von etwa 15 nt bis etwa 30 nt oder von etwa 15 nt bis etwa 25 nt aufweisen.

[0184] Auch in Bezug sowohl auf ein vorliegendes auf DNA gerichtetes RNA-Einzelmolekül als auch auf ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül kann der dsRNA-Duplex des Protein-bindenden Segments eine Länge von etwa 6 Basenpaaren (bp) bis etwa 50 bp aufweisen. Beispielsweise kann der dsRNA-Duplex des Protein-bindenden Segments eine Länge von etwa 6 bp bis etwa 40 bp, von etwa 6 bp bis etwa 30 bp, von etwa 6 bp bis etwa 25 bp, von etwa 6 bp bis etwa 20 bp, von etwa 6 bp bis etwa 15 bp, von etwa 8 bp bis etwa 40 bp, von etwa 8 bp bis etwa 30 bp, von etwa 8 bp bis etwa 25 bp, von etwa 8 bp bis etwa 20 bp oder von etwa 8 bp bis etwa 15 bp aufweisen. Beispielsweise kann der dsRNA-Duplex des Protein-bindenden Segments eine Länge von etwa 8 bp bis etwa 10 bp, von etwa 10 bp bis etwa 15 bp, von etwa 15 bp bis etwa 18 bp, von etwa 18 bp bis etwa 20 bp, von etwa 20 bis etwa 25 bp, von etwa 25 bp bis etwa 30 bp, von etwa 30 bp bis etwa 35 bp, von etwa 35 bp bis etwa 40 bp, oder von etwa 40 bp bis etwa 50 bp aufweisen. In einigen Ausführungsformen weist der dsRNA-Duplex des Protein-bindenden Segmentes eine Länge von 36 Basenpaaren auf. Die prozentuale Komplementarität zwischen der Nukleotidsequenz, die hybridisiert, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden, kann mindestens etwa 60% betragen. Beispielsweise kann die prozentuale Komplementarität zwischen der Nukleotidsequenz, die hybridisiert, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden, mindestens etwa 65%, mindestens etwa 70% mindestens etwa 75%, mindestens etwa 80% mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98% oder mindestens etwa 99% betragen. In einigen Fällen beträgt die prozentuale Komplementarität zwischen den Nukleotidsequenzen, die hybridisieren, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden, 100%.

Ortsspezifisch modifizierendes Polypeptid

[0185] Eine vorliegende auf DNA gerichtete RNA und ein vorliegendes ortsspezifisch modifizierendes Polypeptid bilden einen Komplex. Die auf DNA gerichtete RNA sorgt für eine Zielspezifität des Komplexes, indem sie eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz einer Ziel-DNA ist (wie oben beschrieben). Das ortsspezifisch modifizierende Polypeptid des Komplexes sorgt für die ortsspezifische Aktivität. Mit anderen Worten wird das ortsspezifisch modifizierende Polypeptid zu einer DNA-Sequenz (z. B. einer chromosomalen Sequenz oder einer extrachromosomalen Sequenz, z. B. einer episomalen Sequenz, einer kleinen zirkulären Sequenz (minicircle), einer mitochondrialen Sequenz, einer Chloroplasten-Sequenz, usw.) geführt, indem es sich mit mindestens dem Protein-bindenden Segment der auf DNA gerichteten RNA zusammenlagert (oben beschrieben).

[0186] Ein vorliegendes ortsspezifisch modifizierendes Polypeptid modifiziert eine Ziel-DNA (z. B. Spaltung oder Methylierung der Ziel-DNA) und/oder ein mit der Ziel-DNA assoziiertes Polypeptid (z. B. Methylierung oder Acetylierung eines Histon-Tails). Ein ortsspezifisch modifizierendes Polypeptid wird vorliegend auch als ein "ortsspezifisches Polypeptid" oder ein "RNA-bindendes, ortsspezifisch modifizierendes Polypeptid" bezeichnet.

[0187] In einigen Fällen ist das ortsspezifisch modifizierende Polypeptid ein natürlich vorkommendes modifizierendes Polypeptid. In anderen Fällen ist das ortsspezifisch modifizierende Polypeptid ein nicht natürlich vorkommendes Polypeptid (z. B. ein chimäres Polypeptid, wie nachstehend besprochen, oder ein natürlich vorkommendes Polypeptid, das modifiziert ist, z. B. durch Mutation, Deletion, Insertion).

[0188] Beispiele für natürlich vorkommende ortsspezifisch modifizierende Polypeptide werden in den SEQ ID NOs: 1–255 als nicht beschränkende und nicht erschöpfende Liste von natürlich vorkommenden Cas9/Csn1-Endonukleasen bereitgestellt. Diese natürlich vorkommenden Polypeptide binden, wie vorliegend beschrieben, eine auf DNA gerichtete RNA und sind somit auf eine spezifische Sequenz innerhalb einer Ziel-DNA gerichtet und spalten die Ziel-DNA, wobei ein Doppelstrangbruch erzeugt wird. Ein vorliegendes ortsspezifisch modifizierendes Polypeptid umfasst zwei Teile, einen RNA-bindenden Teil und einen Aktivitätsteil. In einigen Ausführungsformen umfasst ein vorliegendes ortsspezifisch modifizierendes Polypeptid Folgendes: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität zeigt (z. B. Aktivität für eine DNA-Methylierung, Aktivität für eine DNA-Spaltung, Aktivität für eine Histon-Acetylierung, Aktivität für eine Histon-Methylierung, usw.), wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0189] In anderen Ausführungsformen umfasst ein vorliegendes ortsspezifisch modifizierendes Polypeptid Folgendes: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf

DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist, und (ii) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert (z. B., um die Transkription zu erhöhen oder zu verringern), wobei die Stelle der veränderten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0190] In einigen Fällen weist ein vorliegendes ortsspezifisch modifizierendes Polypeptid eine enzymatische Aktivität auf, die Ziel-DNA modifiziert (z. B. Nuklease-Aktivität, Methyltransferase-Aktivität, Demethylase-Aktivität, DNA-Reparatur-Aktivität, DNA-Schädigungs-Aktivität, Deaminierungs-Aktivität, Dismutase-Aktivität, Alkylierungs-Aktivität, Depurinierungs-Aktivität, Oxidations-Aktivität, Pyrimidin-Dimer-bildende Aktivität, Integrase-Aktivität, Transposase-Aktivität, Rekombinase-Aktivität, Polymerase-Aktivität, Ligase-Aktivität, Helicase-Aktivität, Photolyase-Aktivität oder Glykosylase-Aktivität).

[0191] In einigen Fällen weist ein vorliegendes ortsspezifisch modifizierendes Polypeptid eine enzymatische Aktivität auf, die ein Polypeptid modifiziert (z. B. ein Histon), das mit Ziel-DNA assoziiert ist (z. B. Methyltransferase-Aktivität, Demethylase-Aktivität, Acetyltransferase-Aktivität, Deacetylase-Aktivität, Kinase-Aktivität, Phosphatase-Aktivität, Ubiquitin-Ligase-Aktivität, deubiquitinierende Aktivität, Adenylierungs-Aktivität, Deadenylierungs-Aktivität, SUMOylierende Aktivität, deSUMOylierende Aktivität, Ribosylierungs-Aktivität, Deribosylierungs-Aktivität, Myristoylierungs-Aktivität oder Demyristoylierungs-Aktivität).

Beispielhafte ortsspezifisch modifizierende Polypeptide

[0192] In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7-166 oder 73-1003 der in **Fig. 3** dargestellten Cas9/Csn1-Aminosäuresequenz oder zu den entsprechenden Teilen in einer der in SEQ ID NOS: 1-256 und 795-1346 gezeigten Aminosäuresequenzen aufweist.

Nukleinsäure-Modifikation

[0193] In einigen Ausführungsformen umfasst eine vorliegende Nukleinsäure (z. B. eine auf DNA gerichtete RNA) eine oder mehrere Modifikationen, z. B. eine Basen-Modifikation, eine Rückgrat-Modifikation, usw., um die Nukleinsäure mit einer neuen oder einer verbesserten Eigenschaft auszustatten (z. B. verbesserte Stabilität). Wie im Stand der Technik bekannt ist, ist ein Nukleosid eine Basen-Zucker-Kombination. Der Basen-Teil des Nukleosids ist normalerweise eine heterocyclische Base. Die zwei häufigsten Klassen solcher heterocyclischen Basen sind die Purine und die Pyrimidine. Nukleotide sind Nukleoside, die ferner eine Phosphatgruppe enthalten, die kovalent mit dem Zuckerteil des Nukleosids verknüpft ist. Für solche Nukleoside, die einen Pentofuranosyl-Zucker enthalten, kann die Phosphatgruppe an die 2', die 3' oder die 5'-Hydroxylgruppe des Zuckers gekoppelt sein. Bei der Bildung von Oligonukleotiden verknüpfen die Phosphatgruppen benachbarte Nukleoside kovalent miteinander, wodurch eine lineare, polymere Verbindung gebildet wird. Die entsprechenden Enden dieser linearen, polymeren Verbindung können wiederum verknüpft werden, um eine zirkuläre Verbindung zu bilden, wobei jedoch lineare Verbindungen im Allgemeinen geeignet sind. Darüber hinaus können lineare Verbindungen eine interne Nukleotid-Basen-Komplementarität aufweisen, und können sich daher in einer Weise falten, bei der eine vollständig oder partiell doppelsträngige Verbindung gebildet wird. Innerhalb von Oligonukleotiden wird von den Phosphatgruppen im Allgemeinen gesagt, dass sie das Internukleosid-Rückgrat des Oligonukleotids bilden. Die normale Verknüpfung oder das Rückgrat der RNA und DNA ist eine 3'- zu 5'-Phosphodiesterverknüpfung.

Modifiziertes Rückgrat und modifizierte Internukleosid-Verknüpfungen

[0194] Beispiele für geeignete Nukleinsäuren, die Modifikationen enthalten, umfassen Nukleinsäuren, die ein modifiziertes Rückgrat oder nicht natürliche Internukleosid-Verknüpfungen enthalten. Nukleinsäuren, die ein modifiziertes Rückgrat aufweisen, umfassen solche, die ein Phosphor-Atom im Rückgrat beibehalten und solche, die über kein Phosphor-Atom im Rückgrat verfügen.

[0195] Geeignete modifizierte Oligonukleotid-Rückgrat-Strukturen, die ein Phosphor-Atom enthalten, umfassen beispielsweise Phosphorthioate, chirale Phosphorthioate, Phosphordithioate, Phosphortriester, Aminoalkyl-Phosphotriester, Methyl- oder andere Alkyl-Phosphonate, einschließlich 3'-Alkyl-Phosphonat, 5'-Alkyl-Phosphonat und chirale Phosphonate, Phosphinate, Phosphoramidate, einschließlich 3'-Amino-Phosphoramidat und Aminoalkyl-Phosphoramidate, Phosphordiamidate, Thiono-Phosphoramidate, Thionoalkyl-Phosphonate, Thionoalkyl-Phosphortriester, Selenphosphate und Boranphosphate, die normale 3'-5'-Verknüpfungen

umfassen, 2'-5'-verknüpfte Analoge derselben und solche, die eine umgekehrte Polarität aufweisen, wobei eine oder mehrere Internukleotid-Verknüpfungen eine 3' zu 3', eine 5' zu 5' oder eine 2' zu 2'-Verknüpfung ist. Geeignete Oligonukleotide, die über eine umgekehrte Polarität verfügen, umfassen eine einzelne 3' zu 3'-Verknüpfung der am stärksten 3'-gelegenen Internukleotid-Verknüpfung, d. h. ein einzelner umgekehrter Nukleotid-Rest, der basisch sein kann (die Nukleobase fehlt oder es liegt stattdessen eine Hydroxylgruppe vor). Verschiedene Salze (z. B. Kalium oder Natrium), gemischte Salze und freie Säure-Formen sind ebenfalls eingeschlossen.

[0196] In einigen Ausführungsformen umfasst eine vorliegende Nukleinsäure ein oder mehrere Phosphorothioate und/oder Heteroatom-Internukleotid-Verknüpfungen, insbesondere $-\text{CH}_2\text{-NH-O-CH}_2-$, $-\text{CH}_2\text{-N}(\text{CH}_3)\text{-O-CH}_2-$ (als eine Methylen-(Methylimino-) oder MMI-Rückgrat bekannt), $-\text{CH}_2\text{-O-N}(\text{CH}_3)\text{-CH}_2-$, $-\text{CH}_2\text{-N}(\text{CH}_3)\text{-N}(\text{CH}_3)\text{-CH}_2-$ und $-\text{O-N}(\text{CH}_3)\text{-CH}_2\text{-CH}_2-$ (wobei die native Phosphordiester-Internukleotid-Verknüpfung als $-\text{O-P(=O)(OH)-O-CH}_2-$ dargestellt ist). Internukleotid-Verknüpfungen vom MMI-Typ werden im oben bezeichneten US Patent Nr. 5,489,677 offenbart. Geeignete Amid-Internukleosid-Verknüpfungen werden in US Patent Nr. 5,602,240 offenbart.

[0197] Ebenfalls geeignet sind Nukleinsäuren, die Morpholino-Rückgrat-Strukturen aufweisen, wie sie beispielsweise in US Patent Nr. 5,034,506 beschrieben werden. Beispielsweise umfasst in einigen Ausführungsformen eine vorliegende Nukleinsäure einen Morpholino-Ring mit 6 Mitgliedern anstelle eines Ribose-Rings. In einigen Ausführungsformen ersetzt eine Phosphordiamidat- oder eine andere Nicht-Phosphordiester-Internukleosid-Verknüpfung eine Phosphordiester-Verknüpfung.

[0198] Geeignete modifizierte Polynukleotid-Rückgrat-Strukturen, die nicht über ein Phosphor-Atom verfügen, weisen Rückgrat-Strukturen auf, die durch kurzkettenige Alkyl- oder Cycloalkyl-Internukleotid-Verknüpfungen, durch gemischte Heteroatom- und Alkyl- oder Cycloalkyl-Internukleotid-Verknüpfungen oder durch ein oder mehrere kurzkettenige Heteroatom- oder heterocyclische Internukleotid-Verknüpfungen gebildet werden. Diese umfassen solche mit Morpholino-Verknüpfungen (teilweise aus dem Zuckerteil eines Nukleosids gebildet); Siloxan-Rückgrat-Strukturen; Sulfid-, Sulfoxid- oder Sulfon-Rückgrat-Strukturen; Formacetyl- und Thioformacetyl-Rückgrat-Strukturen; Methylen-, Formacetyl- und Thioformacetyl-Rückgrat-Strukturen; Ribocacetyl-Rückgrat-Strukturen; Alken-haltige Rückgrat-Strukturen; Sulfamat-Rückgrat-Strukturen; Methylenimino- und Methylenhydrazin-Rückgrat-Strukturen; Sulfonat- und Sulfonamid-Rückgrat-Strukturen; Amid-Rückgrat-Strukturen; und andere, die gemischte N-, O-, S- oder CH_2 -Bestandteile umfassen.

Mimetika

[0199] Eine vorliegende Nukleinsäure kann ein Nukleinsäure-Mimetikum sein. Der Begriff "Mimetikum", sofern er auf Polynukleotide bezogen ist, soll Polynukleotide umfassen, bei denen lediglich der Furanose-Ring oder sowohl der Furanose-Ring als auch die Internukleotid-Verknüpfung durch Nicht-Furanose-Gruppen ersetzt werden, wobei das Ersetzen lediglich des Furanose-Rings im Stand der Technik auch als Zucker-Surrogat bezeichnet wird. Der heterocyclische Basenrest oder ein modifizierter heterocyclischer Basenrest wird für die Hybridisierung mit einer geeigneten Ziel-Nukleinsäure beibehalten. Eine solche Nukleinsäure, ein Polynukleotid-Mimetikum, das hervorragende Hybridisierungseigenschaften gezeigt hat, wird als Peptid-Nukleinsäure (peptide nucleic acid, PNA) bezeichnet. In einem PNA wird das Zucker-Rückgrat eines Polynukleotids durch ein Amid-haltiges Rückgrat ersetzt, insbesondere durch ein Aminoethylglycin-Rückgrat. Die Nukleotide werden beibehalten, und diese werden entweder direkt oder indirekt an Aza-Stickstoff-Atome des Amid-Teils des Rückgrats gebunden.

[0200] Ein Polynukleotid-Mimetikum, von dem hervorragende Hybridisierungseigenschaften berichtet wurden, ist eine Peptid-Nukleinsäure (PNA). Das Rückgrat in PNA-Verbindungen besteht aus zwei oder mehreren verknüpften Aminoethylglycin-Einheiten, was der PNA ein Amid-haltiges Rückgrat verleiht. Die heterocyclischen Basenreste sind entweder direkt oder indirekt an Aza-Stickstoff-Atome des Amid-Teils des Rückgrats gebunden. Beispielshafte US-Patente, welche die Herstellung von PNA-Verbindungen beschreiben umfassen (sind jedoch nicht beschränkt auf): US-Patent-Nm. 5,539,082; 5,714,331; und 5,719,262.

[0201] Eine weitere Klasse von Polynukleotid-Mimetika, die untersucht wurde, basiert auf verknüpften Morpholino-Einheiten (Morpholino-Nukleinsäure), die heterocyclische Basen aufweisen, die an den Morpholino-Ring gekoppelt sind. Es wurde eine Vielzahl von Verknüpfungsgruppen beschrieben, welche die monomeren Morpholino-Einheiten in einer Morpholino-Nukleinsäure verknüpfen. Eine Klasse von Verknüpfungsgruppen wurde ausgewählt, um eine nicht-ionische, oligomere Verbindung zu bilden. Bei den nicht-ionischen, Morpholino-basierten oligomeren Verbindungen ist es weniger wahrscheinlich, dass sie unerwünschte Interaktionen mit

zellulären Proteinen zeigen. Morpholino-basierte Polynukleotide sind nicht-ionische Mimetika von Oligonukleotiden, bei denen es weniger wahrscheinlich ist, dass sie unerwünschte Interaktionen mit zellulären Proteinen zeigen (Dwayne A. Braasch und David R. Corey, *Biochemistry*, 2002, 41(14), 4503–4510). Morpholino-basierte Polynukleotide werden in US-Patent Nr. 5,034,506 offenbart. Eine Vielzahl von Verbindungen innerhalb der Morpholino-Klasse von Polynukleotiden ist hergestellt worden, wobei diese eine Vielzahl von verschiedenen Verknüpfungsgruppen aufweisen, welche die monomeren Untereinheiten verbinden.

[0202] Eine weitere Klasse von Polynukleotid-Mimetika wird als Cyclohexenyl-Nukleinsäuren (cyclohexenyl nucleic acids, CeNA) bezeichnet. Der Furanose-Ring, der normalerweise in einem DNA/RNA-Molekül vorhanden ist, wird gegen einen Cyclohexenyl-Ring ersetzt. CeNA-DMT-geschützte Phosphoramidit-Monomere wurden hergestellt und zur Synthese oligomerer Verbindungen verwendet, wobei herkömmliche Phosphoramidit-Chemie angewendet wurde. Vollständig modifizierte oligomere CeNA-Verbindungen und Oligonukleotide, bei denen spezifische Positionen mit CeNA modifiziert wurden, wurden hergestellt und untersucht (siehe Wang et al., *J. Am. Chem. Soc.*, 2000, 122, 8595–8602). Im Allgemeinen erhöht der Einbau von CeNA-Monomeren in eine DNA-Kette deren Stabilität in einem DNA-RNA-Hybrid. CeNA-Oligoadenylate bildeten Komplexe mit RNA- und DNA-Komplementen mit einer Stabilität, die ähnlich zu der von nativen Komplexen war. Es konnte durch NMR und zirkulären Dichroismus gezeigt werden, dass die Untersuchung des Einbaus von CeNA-Strukturen in natürliche Nukleinsäure-Strukturen mit einfachen Konformationsanpassungen verlief.

[0203] Eine weitere Modifikation umfasst die geschlossenen Nukleinsäuren (locked nucleic acids, LNAs), bei denen die 2'-Hydroxylgruppe an das 4'-Kohlenstoffatom des Zuckerrings gebunden ist, wodurch eine 2'-C, 4'-C-Oxymethylen-Verknüpfung gebildet wird, wodurch ein bicyclischer Zuckerrest gebildet wird. Bei der Verknüpfung kann es sich um ein Methylen (-CH₂-) handeln, eine Gruppe, die das 2'-Sauerstoffatom und das 4'-Kohlenstoffatom miteinander verbindet, wobei n 1 oder 2 ist (Singh et al., *Chem. Commun.*, 1998, 4, 455–456). LNA und LNA-Analoga zeigen sehr hohe Duplex-Temperaturstabilität mit komplementärer DNA und RNA (T_m = +3 bis +10°C), sowie eine Stabilität gegen 3'-exonukleolytischen Abbau und gute Löslichkeitseigenschaften. Starke und nicht-toxische Antisense-Oligonukleotide, welche LNAs enthalten, sind beschrieben worden (Wahlestedt et al., *Proc. Natl. Acad. Sci. U.S.A.*, 2000, 97, 5633–5638).

[0204] Die Synthese und Herstellung der RNA-Monomere Adenin, Cytosin, Guanin, 5'-Methylcytosin, Thymin und Uracil sind gemeinsam mit ihrer Oligomerisierung und ihren Nukleinsäure-Erkennungseigenschaften beschrieben worden (Koshkin et al., *Tetrahedron*, 1998, 54, 3607–3630). LNAs und deren Herstellung sind ebenfalls in WO 98/39352 und WO 99/14226 beschrieben.

Modifizierte Zuckerreste

[0205] Eine vorliegende Nukleinsäure kann ferner ein oder mehrere substituierte Zuckerreste umfassen. Geeignete Polynukleotide umfassen eine Zucker-Substituenten-Gruppe, die ausgewählt ist aus: OH; F; O-, S- oder N-Alkyl; O-, S- oder N-Alkenyl; O-, S- oder N-Alkynyl; oder O-Alkyl-O-Alkyl, wobei das Alkyl, Alkenyl, und Alkynyl substituiertes oder unsubstituiertes C₁ bis C₁₀ Alkyl oder C₂ bis C₁₀ Alkenyl oder Alkynyl sein kann. Besonders geeignet sind O((CH₂)_nO)_mCH₃, O(CH₂)_nOCH₃, O(CH₂)_nNH₂, O(CH₂)_nCH₃, O(CH₂)₂ONH₂ und O(CH₂)_nON((CH₂)_nCH₃)₂, wobei n und m von 1 bis etwa 10 sind. Andere geeignete Polynukleotide umfassen eine Zucker-Substituenten-Gruppe, die ausgewählt ist aus C₁ bis C₁₀ Niederalkyl, substituiertes Niederalkyl, Alkenyl, Alkynyl, Alkaryl, Aralkyl, O-Alkaryl oder O-Aralkyl, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, Heterocycloalkyl, Heterocycloalkaryl, Aminoalkylamino, Polyalkylamino, substituiertes Silyl, eine RNA-Spaltungsgruppe, eine Reportergruppe, einen Interkalator, eine Gruppe zur Verbesserung der pharmakokinetischen Eigenschaften eines Oligonukleotids, oder eine Gruppe zur Verbesserung der pharmakodynamischen Eigenschaften eines Oligonukleotids und andere Substituenten mit ähnlichen Eigenschaften. Eine geeignete Modifikation umfasst 2'-Methoxyethoxy (2'-O-CH₂CH₂OCH₃, auch bekannt als 2'-O-(2-Methoxyethyl) oder 2'-MOE) (Martin et al., *Helv. Chim. Acta*, 1995, 78, 486–504), d. h. eine Alkoxyalkoxy-Gruppe. Eine weitere geeignete Modifikation umfasst 2'-Dimethylaminoethoxyethoxy, d. h. eine O(CH₂)₂ON(CH₃)₂-Gruppe, die auch als 2'-DMAOE bekannt ist, wie in den nachfolgenden Beispielen beschrieben, und 2'-Dimethylaminoethoxyethoxy (im Stand der Technik auch als 2'-O-Dimethyl-Amino-Ethoxy-Ethyl oder 2'-DMAEOE bekannt), d. h. 2'-O-CH₂-O-CH₂-N(CH₃)₂.

[0206] Andere geeignete Zucker-Substituenten-Gruppen umfassen Methoxy (-O-CH₃), Aminopropoxy (--OCH₂CH₂CH₂NH₂), Allyl (-CH₂-CH=CH₂), -O-Allyl (--O-CH₂-CH=CH₂) und Fluor (F). 2'-Zucker-Substituenten-Gruppen können in der Arabino (oberen) Position oder in der Ribo (unteren) Position vorliegen. Eine geeignete 2'-Arabino-Modifikation ist 2'-F. Ähnliche Modifikationen können auch an anderen Positionen an der oligomeren Verbindung gemacht werden, insbesondere an der 3'-Position des Zuckers am 3'-terminalen Nukleosid

oder in 2'-5'-verknüpften Oligonukleotiden an der 5'-Position des 5'-terminalen Nukleotids. Oligomere Verbindungen können auch Zucker-Mimetika aufweisen, wie beispielsweise Cyclobutylreste anstelle der Pentofuranosyl-Zucker.

Basen Modifikation und Substitutionen

[0207] Eine vorliegende Nukleinsäure kann Modifikationen oder Substitutionen der Nukleobase (im Stand der Technik oftmals einfach als "Base" bezeichnet) umfassen. Wie vorliegend verwendet umfassen "unmodifizierte" oder "natürliche" Nukleobasen die Purinbasen Adenin (A) und Guanin (G) und die Pyrimidinbasen Thymin (T), Cytosin (C) und Uracil (U). Modifizierte Nukleobasen umfassen andere synthetische und natürliche Nukleobasen, wie beispielsweise Methylcytosin (5-me-C), 5-Hydroxymethyl-Cytosin, Xanthin, Hypoxanthin, 2-Aminoadenin, 6-Methyl- und andere Alkyl-Derivate von Adenin und Guanin, 2-Propyl und andere Alkyl-Derivate von Adenin und Guanin, 2-Thiouracil, 2-Thiothymin und 2-Thiocytosin, 5-Halo-Uracil und -Cytosin, 5-Propynyl (-C≡C-CH₃)-Uracil und -Cytosin und andere Alkynyl-Derivate von Pyrimidin-Basen, 6-Azo-Uracil, -Cytosin und -Thymin, 5-Uracil (Pseudo-Uracil), 4-Thio-Uracil, 8-Halo, 8-Amino, 8-Thiol, 8-Thioalkyl, 8-Hydroxyl und andere 8-substituierte Adenin und Guanin, 5-Halo-, insbesondere 5-Brom-, 5-Trifluormethyl und andere 5-substituierte Uracile und Cytosin, 7-Methyl-Guanin und 7-Methyl-Adenin, 2-F-Adenin, 2-Amino-Adenin, 8-Aza-Guanin und 8-Aza-Adenin, 7-Deaza-Guanin und 7-Deaza-Adenin und 3-Deaza-Guanin und 3-Deaza-Adenine. Weitere modifizierte Nukleobasen umfassen tricyclische Pyrimidine, wie beispielsweise Phenoxazin-Cytidin (1H-Pyrimido(5,4-b)(1,4)benzoxazin-2(3H)-on), Phenothiazin-Cytidin (1H-Pyrimido(5,4-b)(1,4)benzothiazin-2(3H)-on), G-Clamps, wie beispielsweise substituiertes Phenoxazin-Cytidin (z. B. 9-(2-Aminoethoxy)-H-pyrimido(5,4-b)(1,4)benzoxazin-2(3H)-on), Carbazol-Cytidin (2H-Pyrimido(4,5-b)indol-2-on), Pyridoindol-Cytidine (H-Pyrido(3',2':4,5)pyrrolo(2,3-d)Pyrimidin-2-on).

[0208] Heterocyclische Basenreste können ferner auch solche umfassen, in denen die Purin- oder Pyrimidin-Base gegen andere heterocyclische Verbindungen ersetzt wurde, beispielsweise 7-Deaza-Adenin, 7-Deaza-Guanosine, 2-Amino-Pyridin und 2-Pyridon. Weitere Nukleobasen umfassen solche, die in US Patent Nr. 3, 687,808 offenbart sind, solche, die in The Concise Encyclopedia Of Polymer Science And Engineering, Seiten 858–859, Kroschwitz, J. I., Hrsg. John Wiley & Sons, 1990 offenbart sind, solche, die von English et al., Angewandte Chemie, International Edition, 1991, 30, 613 offenbart sind, und solche, die in Sanghvi, Y. S., Kapitel 15, Antisense Research and Applications, Seiten 289–302, Crooke, S. T. und Lebleu, B., Hrsg., CRC Press, 1993 offenbart sind. Bestimmte dieser Nukleobasen sind nützlich zur Erhöhung der Bindungsaffinität einer oligomeren Verbindung. Diese umfassen 5'-substituierte Pyrimidine, 6'-Aza-Pyrimidine und N-2-, N-6- und O-6-substituierte Purine, einschließlich 2-Aminopropyl-Adenin, 5-Propinyl-Uracil und 5-Propinyl-Cytosin. Es wurde gezeigt, dass 5-Methyl-Cytosin-Substitutionen die Stabilität eines Nukleinsäure-Duplex um 0,6–1,2° C erhöht (Sanghvi et al., Hrsg., Antisense Research and Applications, CRC Press, Boca Raton, 1993, Seiten 276–278) und geeignete Basen-Substitutionen sind, z. B. in Kombination mit 2'-O-Methoxyethyl-Zucker-Modifikationen.

Konjugate

[0209] Eine weitere mögliche Modifikation einer vorliegenden Nukleinsäure umfasst die chemische Anbindung von einem oder von mehreren Resten oder Konjugaten, welche die Aktivität, die zelluläre Verteilung oder die zelluläre Aufnahme von Oligonukleotiden verstärken, an das Polynukleotid. Diese Reste oder Konjugate können Konjugat-Gruppen umfassen, die kovalent an funktionelle Gruppen gebunden sind, wie beispielsweise an primäre oder sekundäre Hydroxyl-Gruppen. Konjugat-Gruppen umfassen, sind jedoch nicht beschränkt auf, Interkalatoren, Reportermoleküle, Polyamine, Polyamide, Polyethylenglykole, Polyether, Gruppen, die die pharmakodynamischen Eigenschaften von Oligomeren verstärken, und Gruppen, die die pharmakokinetischen Eigenschaften von Oligomeren verstärken. Geeignete Konjugat-Gruppen umfassen (sind jedoch nicht beschränkt auf) Cholesterole, Lipide, Phospholipide, Biotin, Phenazin, Folat, Phenanthridin, Anthrachinon, Acridin, Fluoreszine, Rhodamine, Coumarine und Farbstoffe. Gruppen, welche die pharmakodynamischen Eigenschaften verstärken, umfassen Gruppen, die die Aufnahme verbessern, die Resistenz gegen Abbau verstärken und/oder die Sequenz-spezifische Hybridisierung mit der Ziel-Nukleinsäure verstärken. Gruppen, die die pharmakokinetischen Eigenschaften verstärken, umfassen Gruppen, welche die Aufnahme, die Verteilung, den Metabolismus oder die Ausscheidung einer vorliegenden Nukleinsäure verstärken.

[0210] Konjugat-Reste umfassen (sind jedoch nicht beschränkt auf) Lipid-Reste, wie beispielsweise einen Cholesterol-Rest (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553–6556), Cholsäure (Manoharan et al., Bioorg. Med. Chem. Let., 1994, 4, 1053–1060), einen Thioether, z. B. Hexyl-S-tritylthiol (Manoharan et al., Ann. N. Y. Acad. Sci., 1992, 660, 306–309; Manoharan et al., Bioorg. Med. Chem. Let., 1993, 3, 2765–2770), einen Thiocholesterol (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533–538), eine aliphatische Kette,

z. B. Dodecandiol- oder Undecyl-Reste (Saison-Behmoaras et al., EMBO J., 1991, 10, 1111–1118; Kabanov et al., FEBS Lett., 1990, 259, 327–330; Svinarchuk et al., Biochimie, 1993, 75, 49–54), ein Phospholipid, z. B. Di-Hexadecyl-rac-glycerol oder Triethylammonium-1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonat (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651–3654; Shea et al., Nucl. Acids Res., 1990, 18, 3777–3783), eine Polyamin- oder Polyethylenglycol-Kette (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969–973) oder Adamantan-Essigsäure (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651–3654), einen Palmityl-Rest (Mishra et al., Biochim. Biophys. Acta, 1995, 1264, 229–237) oder einen Octadecylamin- oder Hexylaminocarboxy-cholesterol-Rest (Crooke et al., J. Pharmacol. Exp. Ther., 1996, 277, 923–937.)

[0211] Ein Konjugat kann eine "Proteintransduktions-Domäne" oder PTD (auch als CPP bekannt; Zell-penetrierendes Peptid) umfassen, die sich auf ein Polypeptid, ein Polynukleotid, ein Kohlenhydrat oder eine organische oder anorganische Verbindung beziehen kann, welche die Durchquerung einer Doppellipidschicht, einer Mizelle, einer Zellmembran, einer Organellen-Membran oder einer Vesikelmembran ermöglicht. Eine PDD, die an ein anderes Molekül gebunden ist, welches von einem kleinen, polaren Molekül bis zu einem großen Makromolekül und/oder einem Nanopartikel reichen kann, ermöglicht dem Molekül die Durchquerung der Membran, beispielsweise ausgehend vom extrazellulären Raum in den intrazellulären Raum oder vom Cytosol in eine Organelle. In einigen Ausführungsformen ist eine PDD kovalent an den Amino-Terminus eines exogenen Polypeptids gekoppelt (z. B. eines ortsspezifisch modifizierenden Polypeptids). In einigen Ausführungsformen ist eine PDD kovalent an den Carboxy-Terminus eines exogenen Polypeptids gekoppelt (z. B. eines ortsspezifisch modifizierenden Polypeptids). In einigen Ausführungsformen ist eine PDD kovalent an eine Nukleinsäure gekoppelt (z. B. an eine auf DNA gerichtete RNA, an ein Polynukleotid, das für eine auf DNA gerichtete RNA kodiert, an ein Polynukleotid, das für ein ortsspezifisch modifizierendes Polypeptid kodiert, usw.) Beispielhafte PDDs umfassen (sind jedoch nicht beschränkt auf) eine minimale Undecapeptid-Proteintransduktionsdomäne (entsprechend den Resten 47–57 von HIV-1-TAT, umfassend YGRKKRRQRRR; SEQ ID NO: 264); eine Polyarginin-Sequenz, welche mehrere Arginine umfasst, die ausreichen, um den Eintritt in die Zelle zu gewährleisten (z. B. 3, 4, 5, 6, 7, 8, 9, 10 oder 10–50 Arginine); eine VP22-Domäne (Inder et al. (2002) Cancer Gene Ther. 9(6): 489–96); eine Drosophila-Antennapedia-Proteintransduktionsdomäne (Noguchi et al. (2003) Diabetes 52(7): 1732–1737); ein trunkiertes humanes Calcitoninpeptid (Trehin et al. (2004) Pharm. Research 21: 1248–1256); Polylysin (Wender et al. (2000) Proc. Natl. Acad. Sci. USA 97: 13003–13008); RRQRRTSKLMKR (SEQ ID NO: 265); Transportan GWTLNSAGYLLGKINLKALAAALAKKIL (SEQ ID NO: 266); KALAWKALAKALAKALAKHLAKALAKALKCEA (SEQ ID NO: 267); and RQIKIWFQNRRMKWKK (SEQ ID NO: 268). Beispielhafte PTDs umfassen (sind jedoch nicht beschränkt auf) YGRKKRRQRRR (SEQ ID NO: 264), RKKRRQRRR (SEQ ID NO: 269); ein Arginin-Homopolymer aus 3 Arginin-Resten bis 50 Arginin-Resten. Aminosäuresequenzen beispielhafter PTD-Domänen umfassen (sind jedoch nicht beschränkt auf) beliebige der folgenden Sequenzen: YGRKKRRQRRR (SEQ ID NO: 264); RKKRRQRR (SEQ ID NO: 270); YARAAARQARA (SEQ ID NO: 271); THRLPRRRRRR (SEQ ID NO: 272); und GGRRARRRRRR (SEQ ID NO: 273). In einigen Ausführungsformen ist die PTD eine aktivierbare CPP (ACPP) (Aguilera et al. (2009) Integr Biol (Camb) June; 1(5–6): 371–381). ACPPs umfassen ein polykationisches CPP (z. B. Arg9 oder "R9"), das über einen spaltbaren Linker an ein passendes Polyanion (z. B. Glu9 oder "E9") gebunden ist, welches die Nettoladung auf annähernd 0 verringert und somit die Adhäsion und die Aufnahme in Zellen inhibiert. Bei Spaltung des Linkers wird das Polyanion freigesetzt, was lokal die Polyarginine und ihre inhärente Adhäsionsfähigkeit demaskiert, wodurch das ACPP "aktiviert" wird, die Membran zu durchqueren.

Beispielhafte auf DNA-gerichtete RNAs

[0212] In einigen Ausführungsformen umfasst eine geeignete auf DNA gerichtete RNA zwei getrennte Polynukleotid-Moleküle. Das erste der beiden getrennten RNA-Polynukleotid-Moleküle (die Aktivator-RNA) umfasst eine Nukleotidsequenz, die mindestens etwa 60%, mindestens etwa 65%, mindestens etwa 70%, mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99%, oder 100% Nukleinsäure-Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden mit einer der in SEQ ID NOs: 431–562 gezeigten Nukleinsäuresequenzen oder deren Komplementen zeigt. Das zweite der beiden getrennten RNA-Polynukleotid-Moleküle (die Targeter-RNA) umfasst eine Nukleotidsequenz, die mindestens etwa 60%, mindestens etwa 65%, mindestens etwa 70%, mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99%, oder 100% Nukleinsäure-Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden mit einer der in SEQ ID NOs: 563–679 gezeigten Nukleinsäuresequenzen oder deren Komplementen zeigt.

[0213] In einigen Ausführungsformen ist eine geeignete auf DNA gerichtete RNA ein einzelnes RNA-Polynukleotid und umfasst eine erste Nukleotidsequenz mit mindestens etwa 60%, mindestens etwa 65%, minde-

tens etwa 70%, mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder 100% Nukleinsäure-Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden mit einer der in SEQ ID NOs: 431–562 gezeigten Nukleotidsequenzen und eine zweite Nukleotidsequenz mit mindestens etwa 60%, mindestens etwa 65%, mindestens etwa 70%, mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder 100% Nukleotidsequenz-Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden mit einer der in SEQ ID NOs: 463–679 gezeigten Nukleotidsequenzen.

[0214] In einigen Ausführungsformen ist die auf DNA gerichtete RNA ein auf DNA gerichtetes RNA-Doppelmolekül, und die Targeter-RNA umfasst die Sequenz 5'GUUUUAGAGCUA-3' (SEQ ID NO: 679), welche an ihren 5'-Ende an einen Abschnitt von Nukleotiden gekoppelt ist, der komplementär zu einer Ziel-DNA ist. In einigen Ausführungsformen ist die auf DNA gerichtete RNA ein auf DNA gerichtetes RNA-Doppelmolekül und die Aktivator-RNA umfasst die Sequenz 5' UAGCAAGUUAAAAUAAGGCUAGUCCG-3' (SEQ ID NO: //).

[0215] In einigen Ausführungsformen ist die auf DNA gerichtete RNA ein auf DNA gerichtetes RNA-Einzelmolekül, und umfasst die Sequenz 5'-GUUUUAGAGCUA-Linker-UAGCAAGUUAAAAUAAGGCUAGUCCG-3', die über ihr 5'-Ende an einen Abschnitt von Nukleotiden gekoppelt ist, der komplementär zu einer Ziel-DNA ist (wobei "Linker" eine beliebige Linker-Nukleotidsequenz bezeichnet, die eine beliebige Nukleotidsequenz umfassen kann) (SEQ ID NO: //). Andere exemplarische, auf DNA gerichtete RNA-Einzelmoleküle umfassen solche, die in SEQ ID NOs: 680–682 gezeigt sind.

Nukleinsäuren, die für eine vorliegende auf DNA gerichtete RNA und/oder für ein vorliegendes ortsspezifisch modifizierendes Polypeptid kodieren

[0216] Die vorliegende Offenbarung stellt eine Nukleinsäure bereit, die eine Nukleotidsequenz umfasst, welche für eine vorliegende auf DNA gerichtete RNA und/oder ein vorliegendes ortsspezifisch modifizierendes Polypeptid kodieren. In einigen Ausführungsformen ist eine Nukleinsäure, die für eine vorliegende auf DNA gerichtete RNA kodiert, ein Expressionsvektor, z. B. ein rekombinanter Vektor.

[0217] In einigen Ausführungsformen umfasst ein betreffendes Verfahren das Inkontaktbringen einer oder mehrerer Nukleinsäure, die Nukleotidsequenzen umfassen, welche für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodieren, mit einer Ziel-DNA oder die Einbringung derselben in eine Zelle (oder in eine Population von Zellen). In einigen Ausführungsformen liegt eine Zelle, die eine Ziel-DNA umfasst, in vitro vor. In einigen Ausführungsformen liegt eine Zelle, die eine Ziel-DNA umfasst, in vivo vor. Geeignete Nukleinsäuren, die Nukleotidsequenzen umfassen, welche für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodieren, umfassen Expressionsvektoren, wobei ein Expressionsvektor eine Nukleotidsequenz umfasst, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, ein "rekombinanter Expressionsvektor" ist.

[0218] In einigen Ausführungsformen ist der rekombinante Expressionsvektor ein virales Konstrukt, z. B. ein Konstrukt aus einem rekombinanten Adeno-assoziierten Virus (siehe beispielsweise US Patent Nr. 7,078,387), ein rekombinantes adenovirales Konstrukt, ein rekombinantes lentivirales Konstrukt, ein rekombinantes retrovirales Konstrukt, usw.

[0219] Geeignete Expressionsvektoren umfassen (sind jedoch nicht beschränkt auf) virale Vektoren (z. B. virale Vektoren, die auf einem Vaccinia-Virus basieren; Poliovirus; Adenovirus (siehe beispielsweise Lee et al., Invest Ophthalmol Vis Sci 35: 2543 2549, 1994; Borrás et al., Gene Ther 6: 515 524, 1999; Li and Davidson, PNAS 92: 7700 7704, 1995; Sakamoto et al., H Gene Ther 5: 1088 1097, 1999; WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 und WO 95/00655); Adeno-assoziiertes Virus (siehe z. B., Ali et al., Hum Gene Ther 9: 81 86, 1998; Flannery et al., PNAS 94: 6916 6921, 1997; Bennett et al., Invest Ophthalmol Vis Sci 38: 2857 2863, 1997; Jomary et al., Gene Ther 4: 683 690, 1997; Rolling et al., Hum Gene Ther 10: 641 648, 1999; Ali et al., Hum Mol Genet 5: 591 594, 1996; Srivastava in WO 93/09239, Samulski et al., J. Vir. (1989) 63: 3822–3828; Mendelson et al., Virol. (1988) 166: 154–165; und Flotte et al., PNAS (1993) 90: 10613–10617); SV40; Herpes simplex-Virus; humanes Immundefizienzvirus (siehe z. B. Miyoshi et al., PNAS 94: 10319 23, 1997; Takahashi et al., J Virol 73: 7812 7816, 1999); einen retroviralen Vektor (z. B. murines Leukämievirus, Milz-Nekrose-Virus, und Vektoren, die von Retroviren abgeleitet sind, wie beispielsweise Rous-Sarkom-Virus, Harvey-Sarkom-Virus, Vogel-Leukose-Virus, einem Lentivirus, humanes Immundefizienzvirus, myeloproliferatives Sarkom-Virus, und Brusttumor-Virus); und ähnliche.

[0220] Zahlreiche geeignete Expressionsvektoren sind dem Fachmann bekannt und viele davon sind kommerziell verfügbar. Die folgenden Vektoren werden beispielhaft bereitgestellt. Für eukaryontische Wirtszellen: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, und pSVLSV40 (Pharmacia). Beliebige andere Vektoren können jedoch ebenfalls verwendet werden, solange diese mit der Wirtszelle kompatibel sind.

[0221] Abhängig von dem Wirt-Vektor-System, das verwendet wird, können eine beliebige Vielzahl von geeigneten Transkriptions- und Translations-Kontrollelementen in dem Expressionsvektor verwendet werden, einschließlich konstitutiver und induzierbarer Promotoren Transkriptions-Verstärkungselemente, Transkriptions-Terminatoren, usw. (siehe z. B. Bitter et al. (1987) *Methods in Enzymology*, 153: 516–544).

[0222] In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit einem Kontrollelement verknüpft, z. B. mit einem transkriptionalen Kontrollelement, wie beispielsweise mit einem Promotor. Das transkriptionale Kontrollelement kann entweder in einer eukaryontischen Zelle, z. B. in einer Säugetierzelle, oder in einer prokaryontischen Zelle, z. B. in einer bakteriellen Zelle oder in einer Archaeen-Zelle funktionsfähig sein. In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise an mehrere Kontrollelemente gekoppelt, die eine Expression der Nukleotidsequenz, welche für die auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, sowohl in prokaryontischen als auch eukaryontischen Zellen erlauben.

[0223] Nicht einschränkende Beispiele für geeignete eukaryontische Promotoren (Promotoren, die in einer eukaryontischen Zelle funktionsfähig sind) umfassen solche von Cytomegalovirus (CMV) immediate early, Herpes simplex Virus (HSV), Thymidinkinase, früher und später SV40, Long Terminal Repeats (LTRs) von Retroviren und Maus-Metallothionein-I. Die Auswahl des geeigneten Vektors und des geeigneten Promotors liegt im Bereich des Könnens des Durchschnittsfachmanns. Der Expressionsvektor kann ferner eine Ribosomenbindestelle für die Initiation der Translation sowie einen Transkriptions-Terminator umfassen. Der Expressionsvektor kann ferner geeignete Sequenzen für die Verstärkung der Expression umfassen. Der Expressionsvektor kann ferner eine Nukleotidsequenz umfassen, die für Protein-Tags kodiert (z. B. 6xHis-Tag, Hämagglutinin-Tag, grünes fluoreszierendes Protein, usw.), welche an das ortsspezifisch modifizierende Polypeptid fusioniert werden, was zu einem chimären Polypeptid führt.

[0224] In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise an einen induzierbaren Promotor gekoppelt. In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise an einen konstitutiven Promotor gekoppelt.

[0225] Verfahren zur Einbringung von Nukleinsäure in eine Wirtszelle sind im Stand der Technik bekannt und jedes beliebige bekannte Verfahren kann verwendet werden, um eine Nukleinsäure (z. B. ein Expressionskonstrukt) in eine Zelle einzubringen. Geeignete Verfahren umfassen z. B. virale Infektion oder Bakteriophagen-Infektion, Transfektion, Konjugation, Protoplasten-Fusion, Lipofektion, Elektroporation, Kalziumphosphat-Präzipitation, Polyethylenimin(PEI)-vermittelte Transfektion, DEAE-Dextran-vermittelte Transfektion, Liposomen-vermittelte Transfektion, "Particle Gun"-Technik, Kalziumphosphat-Präzipitation, direkte Mikroinjektion, Nanopartikel-vermittelte Nukleinsäure-Bereitstellung (z. B., Panyam et., al *Adv Drug Deliv Rev.* 2012 Sep 13. pii: 50169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023) und ähnliche.

CHIMÄRE POLYPEPTIDE

[0226] Die vorliegende Offenbarung stellt ein chimäres ortsspezifisch modifizierendes Polypeptid bereit. Ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid interagiert mit (z. B. bindet an) einer vorliegenden auf DNA gerichteten RNA (oben beschrieben). Die auf DNA gerichtete RNA leitet das chimäre ortsspezifisch modifizierende Polypeptid zu einer Ziel-Sequenz innerhalb einer Ziel-DNA (z. B. einer chromosomalen Sequenz oder einer extrachromosomalen Sequenz, z. B. einer episomalen Sequenz, einer kleinen zirkulären Sequenz (minicircle), einer mitochondrialen Sequenz, einer Chloroplasten-Sequenz, usw.). Ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid modifiziert Ziel-DNA (z. B. Spaltung oder Methylierung von Ziel-DNA) und/oder ein Polypeptid, das mit einer Ziel-DNA assoziiert ist (z. B. Methylierung oder Acetylierung eines Histon-Tails).

[0227] Ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid modifiziert Ziel-DNA (z. B. Spaltung oder Methylierung von Ziel-DNA) und/oder ein Polypeptid, das mit einer Ziel-DNA assoziiert ist (z. B. Methylierung oder Acetylierung eines Histon-Tails). Ein chimäres ortsspezifisch modifizierendes Polypeptid wird vorliegend auch als ein "chimäres ortsspezifisches Polypeptid" oder als "chimäres RNA-bindendes ortsspezifisch modifizierendes Polypeptid" bezeichnet.

[0228] Ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid umfasst zwei Teile, einen RNA-bindenden Teil und einen Aktivitätsteil. Ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid umfasst Aminosäuresequenzen, die von mindestens zwei unterschiedlichen Polypeptiden abgeleitet sind. Ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid kann modifizierte und/oder natürlich vorkommende Polypeptidsequenzen umfassen (z. B. eine erste Aminosäuresequenz von einem modifizierten oder unmodifizierten Cas9/Csn1-Protein; und eine zweite Aminosäuresequenz, bei der es sich nicht um das Cas9/Csn1-Protein handelt).

RNA-bindender Teil

[0229] In einigen Fällen ist der RNA-bindende Teil eines vorliegenden ortsspezifisch modifizierenden Polypeptids ein natürlich vorkommendes Polypeptid. In anderen Fällen ist der RNA-bindende Teil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids kein natürlich vorkommendes Molekül (modifiziert, z. B. durch Mutation, Deletion, Insertion). Natürlich vorkommende RNA-bindende Teile von Interesse werden von im Stand der Technik bekannten ortsspezifisch modifizierenden Polypeptiden abgeleitet. Beispielsweise stellen die SEQ ID NOs: 1–256 und 795–1346 eine nicht einschränkende und nicht erschöpfende Liste von natürlich vorkommenden Cas9/Csn1-Endonukleasen bereit, die als ortsspezifisch modifizierende Polypeptide verwendet werden können. In einigen Fällen umfasst der RNA-bindende Teil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu dem RNA-bindenden Teil eines Polypeptids mit einer der in den Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346 gezeigten Polypeptide aufweist.

[0230] In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99%, oder 100% Aminosäuresequenz-Identität zu den Aminosäuresequenzen von SEQ ID NOs: 7–166 oder 731–1003 der in **Fig. 3** gezeigten Cas9/Csn1-Aminosäuresequenz aufweist oder zu entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen.

Aktivitätsteil

[0231] Zusätzlich zu dem RNA-bindenden Teil umfasst das chimäre ortsspezifisch modifizierende Polypeptid einen "Aktivitätsteil". In einigen Ausführungsformen umfasst der Aktivitätsteil eines vorliegenden ortsspezifisch modifizierenden Polypeptids den natürlich vorkommenden Aktivitätsteil eines ortsspezifisch modifizierenden Polypeptids (z. B. Cas9/Csn1-Endonuklease). In anderen Ausführungsformen umfasst der Aktivitätsteil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids eine modifizierte Aminosäuresequenz (z. B. Substitution, Deletion, Insertion) eines natürlich vorkommenden Aktivitätsteils eines ortsspezifisch modifizierenden Polypeptids. Natürlich vorkommende Aktivitätsteile von Interesse werden von im Stand der Technik bekannten ortsspezifisch modifizierenden Polypeptiden abgeleitet. Beispielsweise stellen die SEQ ID NOs: 1–246 und 795–1346 eine nicht einschränkende und nicht erschöpfende Liste von natürlich vorkommenden Cas9/Csn1-Endonukleasen bereit, die als ortsspezifisch modifizierende Polypeptide verwendet werden können. Der Aktivitätsteil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids ist variabel und kann beliebige heterologe Polypeptidsequenzen umfassen, die im Rahmen der vorliegend offenbarten Verfahren nützlich sein können.

[0232] In einigen Ausführungsformen umfasst ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid das Folgende: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) einen Aktivitätsteil, der eine ortsspezifisch enzymatische Aktivität zeigt (z. B. Aktivität für DNA-Methylierung, Aktivität für DNA-Spaltung, Aktivität für Histon-Acetylierung, Aktivität für Histon-Methylierung, usw.), wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0233] In anderen Ausführungsformen umfasst ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid das Folgende: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert (z. B. die Transkription erhöht oder verringert), wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0234] In einigen Fällen weist der Aktivitätsteil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids eine enzymatische Aktivität auf, welche Ziel-DNA modifiziert (z. B. Nuklease-Aktivität, Methyltransferase-Aktivität, Demethylase-Aktivität, DNA-Reparatur-Aktivität, DNA-Schädigungsaktivität, Deaminierungsaktivität, Dismutase-Aktivität, Alkylierungs-Aktivität, Depurinierungs-Aktivität, Oxidations-Aktivität, Pyrimidin-Dimer-bildende Aktivität, Integrase-Aktivität, Transposase-Aktivität, Rekombinase-Aktivität, Polymerase-Aktivität, Ligase-Aktivität, Helicase-Aktivität, Photolyase-Aktivität oder Glykosylase-Aktivität).

[0235] In anderen Fällen weist der Aktivitätsteil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids eine enzymatische Aktivität auf (z. B. Methyltransferase-Aktivität, Demethylase-Aktivität, Acetyltransferase-Aktivität, Deacetylase-Aktivität, Kinase-Aktivität, Phosphatase-Aktivität, Ubiquitin-Ligase-Aktivität, deubiquitinierende Aktivität, Adenylierungs-Aktivität, Deadenylierungs-Aktivität, SUMOylierende Aktivität, deSUMOylierende Aktivität, Ribosylierungs-Aktivität, Deribosylierungs-Aktivität, Myristoylierungs-Aktivität oder Demyristoylierungs-Aktivität), die ein mit der Ziel-DNA assoziiertes Polypeptid modifiziert (z. B. ein Histon).

[0236] In einigen Fällen weist der Aktivitätsteil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids eine enzymatische Aktivität auf (oben beschrieben). In anderen Fällen moduliert der Aktivitätsteil eines vorliegenden chimären ortsspezifisch modifizierenden Polypeptids die Transkription der Ziel-DNA (oben beschrieben). Der Aktivitätsteil eines vorliegenden ortsspezifisch modifizierenden Polypeptids ist variabel und kann eine heterologe Polypeptidsequenz umfassen, die in den vorliegend beschriebenen Verfahren nützlich ist.

Beispielhafte chimäre ortsspezifisch modifizierende Polypeptide

[0237] In einigen Ausführungsformen umfasst der Aktivitätsteil des chimären ortsspezifisch modifizierenden Polypeptids eine modifizierte Form des Cas9/Csn1-Proteins. In einigen Fällen umfasst die modifizierte Form des Cas9/Csn1-Proteins eine Aminosäure-Änderung (z. B. Deletion, Insertion oder Substitution), welche die natürlich vorkommende Nuklease-Aktivität des Cas9/Csn1-Proteins reduziert. Beispielsweise weist die modifizierte Form des Cas9/Csn1-Proteins in einigen Fällen weniger als 50%, weniger als 40%, weniger als 30%, weniger als 20%, weniger als 10%, weniger als 5% oder weniger als 1% der Nuklease-Aktivität des entsprechenden Wildtyp-Cas9/Csn1-Polypeptids auf. In einigen Fällen hat die modifizierte Form des Cas9/Csn1-Polypeptids keine wesentliche Nuklease-Aktivität.

[0238] In einigen Ausführungsformen handelt es sich bei der modifizierten Form des Cas9/Csn1-Polypeptids um eine D10A(Asparat zu Alanin an Aminosäure-Position 10 von SEQ ID NO: 8)-Mutation (oder die entsprechende Mutation eines der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine), welches den komplementären Strang der Ziel-DNA spalten kann, jedoch eine verringerte Fähigkeit aufweist, den nicht-komplementären Strang der Ziel-DNA zu spalten (siehe **Fig. 11**). In einigen Ausführungsformen ist die modifizierte Form des Cas9/Csn1-Polypeptids eine H840A(Histidin zu Alanin an Aminosäure-Position 840)-Mutation (oder die entsprechende Mutation eines der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine), welche den nicht-komplementären Strang der Ziel-DNA spalten kann, jedoch eine verringerte Fähigkeit zur Spaltung des komplementären Strangs der Ziel-DNA aufweist (siehe **Fig. 11**). In einigen Ausführungsformen umfasst die modifizierte Form des Cas9/Csn1-Polypeptids sowohl die D10A- als auch die H840A-Mutation (oder die entsprechenden Mutationen eines der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine), so dass das Polypeptid eine verringerte Fähigkeit zur Spaltung sowohl des komplementären als auch des nicht-komplementären Strangs der Ziel-DNA aufweist. Andere Reste können mutiert werden, um die obigen Wirkungen zu erreichen (d. h. den einen oder anderen Nuklease-Teil zu inaktivieren). Als nicht einschränkende Beispiele können die Reste D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, und/oder A987 (oder die entsprechenden Mutationen eines der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine) verändert werden (d. h. substituiert) (siehe **Fig. 3**, **Fig. 5**, **Fig. 11A** und Tabelle 1 für weitere Informationen bezüglich der Konservierung von Aminosäure-Resten in Cas9). Ferner sind auch andere Mutationen geeignet, bei denen es sich nicht um Alanin-Substitutionen handelt.

Tabelle 1. Tabelle 1 führt 4 Motive auf, die in Cas9-Sequenzen von verschiedenen Spezies vorhanden sind (siehe auch Fig. 3 und Fig. 5). Die hier aufgeführten Aminosäuren stammen vom Cas9 aus *S. pyogenes* (SEQ ID NO: 8).

Motiv#	Motiv	Aminosäuren (Rest #s)	Hochkonserviert
1	RuvC-ähnlich I	IGLDIGTNSVGWAVI (7-21) (SEQ ID NO:260)	D10, G12, G17
2	RuvC-ähnlich II	IVIMARE (759-766) (SEQ ID NO:261)	E762
3	HNH-Motiv	DVDHIVPQSFLKDDSIDNKVLTRSDKN (837-863) (SEQ ID NO:262)	H840, N854, N863
4	RuvC-ähnlich II	HHAHDAYL (982-989) (SEQ ID NO:263)	H982, H983, A984, D986, A987

[0239] In einigen Fällen umfasst das chimäre ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Cas9/Csn1-Aminosäuresequenz aufweist oder zu den entsprechenden Teilen einer der in SEQ ID NOs: 1–246 und 795–1346 gezeigten Aminosäuresequenzen. In einigen Fällen umfasst das chimäre ortsspezifisch modifizierende Polypeptid vier Motive (wie in Tabelle 4 aufgeführt und in **Fig. 3A** und **Fig. 5** gezeigt), jedes mit Aminosäuresequenzen, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu jedem der in Tabelle 1 aufgeführten vier Motiven (SEQ ID NOs: 260–63) oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 aufgeführten Aminosäuresequenzen. In einigen Fällen umfasst das chimäre ortsspezifisch modifizierende Polypeptid Aminosäuresequenzen, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Cas9/Csn1-Aminosäuresequenz aufweisen oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen.

[0240] In einigen Ausführungsformen umfasst der Aktivitätsteil des ortsspezifisch modifizierenden Polypeptids ein heterologes Polypeptid, das eine DNA-modifizierende Aktivität und/oder eine Transkriptionsfaktor-Aktivität und/oder eine Aktivität, die ein mit der DNA assoziiertes Polypeptid modifiziert, aufweist. In einigen Fällen ersetzt ein heterologes Polypeptid einen Teil des Cas9/Csn1-Polypeptids, der Nuklease-Aktivität bereitstellt. In anderen Ausführungsformen umfasst ein vorliegendes ortsspezifisch modifizierendes Polypeptid sowohl einen Teil des Cas9/Csn1-Polypeptids, der normalerweise die Nuklease-Aktivität bereitstellt (und dieser Teil kann vollständig aktiv sein oder stattdessen so modifiziert sein, dass er weniger als 100% der entsprechenden Wildtyp-Aktivität aufweist), als auch ein heterologes Polypeptid. Mit anderen Worten ist ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid in einigen Fällen ein Fusions-Polypeptid, das sowohl den Teil des Cas9/Csn1-Polypeptids umfasst, der normalerweise die Nuklease-Aktivität bereitstellt, als auch das heterologe Polypeptid. In anderen Fällen ist ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid ein Fusions-Polypeptid, das eine modifizierte Variante des Aktivitätsteils des Cas9/Csn1-Polypeptids umfasst (z. B. Aminosäure-Veränderung, Deletion, Insertion) und ein heterologes Polypeptid. In noch anderen Fällen ist ein vorliegendes ortsspezifisch modifizierendes Polypeptid ein Fusions-Polypeptid, das ein heterologes Polypeptid und den RNA-bindenden Teil eines natürlich vorkommenden oder eines modifizierten ortsspezifisch modifizierenden Polypeptids umfasst.

[0241] Beispielsweise kann in einem chimären Cas9/Csn1-Protein ein natürlich vorkommendes (oder ein modifiziertes, z. B. Mutation, Deletion, Insertion) bakterielles Cas9/Csn1-Polypeptid an eine heterologe Polypeptidsequenz (d. h. eine Polypeptidsequenz von einem Protein, bei dem es sich nicht um Cas9/Csn1 handelt, oder eine Polypeptidsequenz von einem anderen Organismus) fusioniert werden. Die heterologe Polypeptidsequenz kann eine Aktivität zeigen (z. B. enzymatische Aktivität), die auch durch das chimäre Cas9/Csn1-Protein gezeigt werden wird (z. B. Methyltransferase-Aktivität, Acetyltransferase-Aktivität, Kinase-Aktivität, ubiquitinierende Aktivität, usw.). Eine heterologe Nukleinsäuresequenz kann mit einer anderen Nukleinsäuresequenz verknüpft sein (z. B. durch genetische Modifikation), wodurch eine chimäre Nukleotidsequenz erzeugt wird, die für ein chimäres Polypeptid kodiert. In einigen Ausführungsformen wird ein chimäres Cas9/Csn1-Polypeptid erzeugt, indem man ein Cas9/Csn1-Polypeptid (z. B. Wildtyp-Cas9 oder eine Cas9-Variante, z. B. ein Cas9 mit verringerter oder inaktiverter Nuklease-Aktivität) mit einer heterologen Sequenz fusioniert, die für eine subzelluläre Ortsbestimmung sorgt (z. B. ein nukleares Ortsbestimmungssignal (nuclear localization signal; NLS) für die gezielte Bereitstellung im Zellkern; ein mitochondriales Ortsbestimmungssignal für die gezielte Bereitstellung in den Mitochondrien; ein Chloroplasten-Ortsbestimmungssignal für die gezielte Bereitstellung in einem

Chloroplasten; ein ER-Retentionssignal, und ähnliches). In einigen Ausführungsformen kann die heterologe Sequenz einen Tag bereitstellen, der die Verfolgung oder Aufreinigung erleichtert (z. B. ein fluoreszentes Protein, z. B. grünes fluoreszierendes Protein (GFP), YFP, RFP, CFP, mCherry, tdTomato, und ähnliche; einen HIS-Tag, z. B. einen 6xHis-Tag; einen Hämagglutinin(HA)-Tag; einen FLAG-Tag; einen Myc-Tag; und ähnliches). In einigen Ausführungsformen kann die heterologe Sequenz für eine erhöhte oder verringerte Stabilität sorgen. In einigen Ausführungsformen kann die heterologe Sequenz eine Bindungsdomäne bereitstellen (z. B., um einem chimären Cas9-Polypeptid die Fähigkeit zu verleihen, an ein anderes Protein von Interesse zu binden, z. B. an ein DNA- oder Histon-modifizierendes Protein, an einen Transkriptionsfaktor oder an einen Transkriptions-Repressor, an ein rekrutierendes Protein usw.).

[0242] Beispiele für verschiedene, zusätzlich geeignete Fusionspartner (oder Fragmente davon) für eine vorliegende Variante eines ortsspezifischen Cas9-Polypeptids umfassen (sind jedoch nicht beschränkt auf) solche, die in **Fig. 54** aufgeführt sind.

Nukleinsäure, die für ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid kodiert

[0243] Die vorliegende Offenbarung stellt eine Nukleinsäure bereit, die eine Nukleotidsequenz umfasst, welche für ein vorliegendes ortsspezifisch modifizierendes Polypeptid kodiert. In einigen Ausführungsformen ist die Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid kodiert, ein Expressionsvektor, z. B. ein rekombinanter Expressionsvektor.

[0244] In einigen Ausführungsformen umfasst ein vorliegendes Verfahren Schritte, bei denen man eine oder mehrere Nukleinsäuren, die ein chimäres ortsspezifisch modifizierendes Polypeptid umfassen, mit einer Ziel-DNA in Kontakt bringt oder diese in eine Zelle (oder in eine Population von Zellen) einbringt. Geeignete Nukleinsäuren, die Nukleotidsequenzen umfassen, welche für ein chimäres ortsspezifisch modifizierendes Polypeptid kodieren, umfassen Expressionsvektoren, wobei ein Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, ein "rekombinanter Expressionsvektor" ist.

[0245] In einigen Ausführungsformen ist der rekombinante Expressionsvektor ein virales Konstrukt, z. B. ein Konstrukt aus einem rekombinanten Adeno-assoziierten Virus (z. B. US Patent Nr. 7,078,387), ein rekombinantes adenovirales Konstrukt, ein rekombinantes lentivirales Konstrukt, usw.

[0246] Geeignete Expressionsvektoren umfassen, sind jedoch nicht beschränkt auf, virale Vektoren (z. B. virale Vektoren, die auf Folgendem basieren: Vaccinia-Virus; Poliovirus; Adenovirus (siehe z. B. Li et al., Invest Ophthalmol Vis Sci 35: 2543–2549, 1994; Borrás et al., Gene Ther 6: 515–524, 1999; Li und Davidson, PNAS 92: 7700–7704, 1995; Sakamoto et al., Hum Gene Ther 5: 1088–1097, 1999; WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 und WO 95/00655); Adeno-assoziiertes Virus (siehe z. B. Ali et al., Hum Gene Ther 9: 81–86, 1998; Flannery et al., PNAS 94: 6916–6921, 1997; Bennett et al., Invest Ophthalmol Vis Sci 38: 2857–2863, 1997; Jomary et al., Gene Ther 4: 683–690, 1997; Rolling et al., Hum Gene Ther 10: 641–648, 1999; Ali et al., Hum Mol Genet 5: 591–594, 1996; Srivastava in WO 93/09239, Samulski et al., J. Vir. (1989) 63: 3822–3828; Mendelson et al., Virol. (1988) 166: 154–165; und Flotte et al., PNAS (1993) 90: 10613–10617); SV40; Herpes-simplex-Virus; humanem Immundefizienzvirus (siehe z. B. Miyoshi et al., PNAS 94: 10319–10323, 1997; Takahashi et al., J Virol 73: 7812–7816, 1999); einem retroviralen Vektor (z. B. murinem Leukämievirus, Milz-Nekrose-Virus, und Vektoren, die von Retroviren abgeleitet sind, wie beispielsweise dem Rous-Sarkom-Virus, dem Harvey-Sarkom-Virus, dem Vogel-Leukose-Virus, einem Lentivirus, humanem Immundefizienzvirus, myeloproliferativem Sarkom-Virus und Brusttumor-Virus); und ähnlichen.

[0247] Mehrere geeignete Expressionsvektoren sind dem Fachmann bekannt und zahlreiche sind kommerziell verfügbar. Die folgenden Vektoren werden beispielhaft genannt; für eukaryotische Wirtszellen: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSGq, und pSVLSV40 (Pharmacia). Es kann jedoch auch ein beliebiger anderer Vektor verwendet werden, solange dieser mit der Wirtszelle kompatibel ist.

[0248] Abhängig vom Wirt/Vektor-System, das verwendet wird, kann eine beliebige Zahl geeigneter Transkriptions- und Translations-Kontrollelemente in dem Expressionsvektor verwendet werden, einschließlich konstitutiver und induzierbarer Promotoren, Transkriptions-Verstärker-Elemente, Transkriptions-Terminatoren, usw. (siehe z. B. Bitter et al. (1987) Methods in Enzymology, 153: 516–544).

[0249] In einigen Ausführungsformen ist eine Nukleotidsequenz, die für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit einem Kontrollelement verbunden, z. B. mit einem

transkriptionalen Kontrollelement, wie beispielsweise einem Promotor. Das transkriptionale Kontrollelement kann entweder in einer eukaryontischen Zelle, z. B. einer Säugetierzelle, oder in einer prokaryontischen Zelle (z. B. in einer bakteriellen Zelle oder in einer Archaeen-Zelle) funktionsfähig sein. In einigen Ausführungsformen ist die Nukleotidsequenz, die für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit mehreren Kontrollelementen verbunden, die eine Expression der Nukleotidsequenz, welche für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, sowohl in prokaryontischen als auch in eukaryontischen Zellen erlauben.

[0250] Nicht einschränkende Beispiele für geeignete eukaryontische Promotoren (Promotoren, die in eukaryontischen Zellen funktionsfähig sind) umfassen solche von Cytomegalovirus (CMV) Immediate Early, Herpes-simplex-Virus (HSV), Thymidinkinase, Early und Late SV40, Long Terminal Repeats (LTRs) von Retrovirus und Maus-Metallothionein-I. Die Auswahl eines geeigneten Vektors und Promotors liegt im Bereich des Könnens des Durchschnittsfachmanns. Der Expressionsvektor kann ferner eine Ribosomenbindungsstelle für die Initiation der Translation umfassen, sowie einen Transkriptions-Terminator. Der Expressionsvektor kann ferner für die Verstärkung der Expression geeignete Sequenzen umfassen. Der Expressionsvektor kann ferner Nukleotidsequenzen umfassen, die für Protein-Anhänge (z. B. 6xHis-Tag, Hämagglutinin(HA)-Tag, ein fluoreszierendes Protein (z. B. ein grün fluoreszierendes Protein; ein gelb fluoreszierendes Protein, usw.) usw.), kodieren, welche an das chimäre ortsspezifisch modifizierende Polypeptid fusioniert werden.

[0251] In einigen Ausführungsformen ist eine Nukleotidsequenz, die für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit einem induzierbaren Promotor verknüpft (z. B. Hit-zeschock-Promotor, Tetracyclin-regulierter Promotor, Steroid-regulierter Promotor, Metall-regulierter Promotor, Östrogenrezeptor-regulierter Promotor, usw.). In einigen Ausführungsformen ist eine Nukleotidsequenz, die für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit einem räumlich beschränkten und/oder einem zeitlich beschränkten Promotor verknüpft (z. B. einem Gewebe-spezifischen Promotor, einem Zelltyp-spezifischen Promotor, usw.). In einigen Ausführungsformen ist eine Nukleotidsequenz, die für ein chimäres ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit einem konstitutiven Promotor verknüpft.

[0252] Verfahren zur Einbringung einer Nukleinsäure in eine Wirtszelle sind im Stand der Technik bekannt, und jedes bekannte Verfahren kann verwendet werden, um eine Nukleinsäure (z. B. ein Expressions-Konstrukt) in eine Stammzelle oder in eine Vorläuferzelle einzubringen. Geeignete Verfahren umfassen, z. B. virale Infektion oder Bakteriophagen-Infektion, Transfektion, Konjugation, Protoplasten-Fusion, Lipofektion, Elektroporation, Kalzium-Phosphat-Präzipitation, Polyethylenimin(PEI)-vermittelte Transfektion, DEAE-Dextran-vermittelte Transfektion, Liposomen-vermittelte Transfektion, "Particle Gun"-Technologie, Kalzium-Phosphat-Präzipitation, direkte Mikroinjektion, Nanopartikel-vermittelte Nukleinsäure-Bereitstellung (siehe z. B. Panyam et., al Adv Drug Deliv Rev. 2012 Sep 13. pii: 50169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023), und ähnliches.

VERFAHREN

[0253] Die vorliegende Offenbarung stellt Verfahren zur Modifizierung einer Ziel-DNA und/oder eines mit einer Ziel-DNA assoziierten Polypeptids bereit. Im Allgemeinen umfasst ein vorliegendes Verfahren das Inkontaktbringen einer Ziel-DNA mit einem Komplex (einem "Targeting-Komplex"), wobei der Komplex eine auf DNA gerichtete RNA und ein ortsspezifisch modifizierendes Polypeptid umfasst.

[0254] Wie oben diskutiert wurde, bilden eine vorliegende auf DNA gerichtete RNA und ein vorliegendes ortsspezifisch modifizierendes Polypeptid einen Komplex. Die auf DNA gerichtete RNA sorgt für die Ziel-Spezifität des Komplexes, indem sie eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz einer Ziel-DNA ist. Das ortsspezifisch modifizierende Polypeptid des Komplexes stellt die ortsspezifische Aktivität bereit. In einigen Ausführungsformen modifiziert ein vorliegender Komplex eine Ziel-DNA, was beispielsweise zu DNA-Spaltung, DNA-Methylierung, DNA-Schädigung, DNA-Reparatur, usw., führt. In anderen Ausführungsformen modifiziert ein vorliegender Komplex ein Ziel-Polypeptid, das mit einer Ziel-DNA assoziiert ist (z. B. ein Histon, ein DNA-bindendes Protein, usw.), was beispielsweise zu einer Histon-Methylierung, Histon-Acetylierung, Histon-Ubiquitinierung und ähnlichem führt. Die Ziel-DNA kann beispielsweise eine nackte DNA in vitro, chromosomale DNA in Zellen in vitro, chromosomale DNA in Zellen in vivo, usw., sein.

[0255] In einigen Fällen zeigt das ortsspezifisch modifizierende Polypeptid eine Nuklease-Aktivität, die Ziel-DNA an einer Ziel-DNA-Sequenz, die durch die Region der Komplementarität zwischen der auf DNA gerichteten RNA und der Ziel-DNA definiert ist, spaltet. Wenn das ortsspezifisch modifizierende Polypeptid ein Cas9- oder ein Cas9-verwandtes Polypeptid ist, so tritt die ortsspezifische Spaltung der Ziel-DNA in einigen Fällen an

Stellen auf, die sowohl durch (i) Basenpaarungs-Komplementarität zwischen der auf DNA gerichteten RNA und der Ziel-DNA; und (ii) einem kurzen Motiv [als das Protospacer-benachbarte Motiv (protospacer adjacent motif (PAM)) bezeichnet] in der Ziel-DNA bestimmt wird. In einigen Ausführungsformen (z. B. wenn Cas9 aus *S. pyogenes* oder ein eng damit verwandtes Cas9 verwendet wird (siehe SEQ ID NOs: 1–256 und 795–1346)), ist die PAM-Sequenz des nicht-komplementären Strangs 5'-XGG-3', wobei X ein beliebiges DNA-Nukleotid ist und X unmittelbar 3' von der Ziel-Sequenz des nicht komplementären Stranges der Ziel-DNA liegt (siehe **Fig. 10**). Somit ist die PAM-Sequenz des komplementären Strangs 5'-CCY-3', wobei Y ein beliebiges Nukleotid ist und Y unmittelbar 5' von der Ziel-Sequenz des komplementären Stranges der Ziel-DNA liegt (siehe **Fig. 10**, bei der das PAM des nicht komplementären Strangs 5'-GGG-3' ist, und das PAM des komplementären Strangs 5'-CCC-3'). In einigen solcher Ausführungsformen können X und Y komplementär sein, und die Basenpaarung X-Y kann ein beliebiges Basenpaar sein (z. B. X=C und Y=G; X=G und Y=C; X=A und Y=T, X=T und Y=A).

[0256] In einigen Fällen kann es vorteilhaft sein, verschiedene Cas9-Proteine (d. h. Cas9-Proteine von verschiedenen Spezies) in den verschiedenen vorliegend bereitgestellten Verfahren zu verwenden, um verschiedene enzymatische Eigenschaften der verschiedenen Cas9-Proteine auszunutzen (z. B. Präferenzen für unterschiedliche PAM-Sequenzen; erhöhte oder verringerte enzymatische Aktivität; erhöhtes oder verringertes Ausmaß von zellulärer Toxizität; Veränderung der Balance zwischen NHEJ, Homologie-gesteuerter Reparatur, Einzelstrangbrüchen, Doppelstrangbrüchen, usw.). Cas9-Proteine von verschiedenen Spezies (SEQ ID NOs: 1–246 und 795–1346) können verschiedene PAM-Sequenzen in der Ziel-DNA erfordern. Folglich kann sich für ein bestimmtes ausgewähltes Cas9-Protein die erforderliche PAM-Sequenz von der Sequenz 5'-XGG-3', die oben beschrieben wurde, unterscheiden.

[0257] Zahlreiche Cas9-Orthologe von einer großen Vielzahl von Spezies wurden vorliegend identifiziert, und den Proteinen sind lediglich wenige Aminosäuren gemein. Alle identifizierten Cas9-Orthologe haben denselben Domänen-Aufbau mit einer zentralen HNH-Endonuklease-Domäne und einer geteilten RuvC/RNaseH-Domäne (siehe **Fig. 3A**, **Fig. 3B**, **Fig. 5** und Tabelle 1). Cas9-Proteine teilen vier Schlüsselmotive mit konserviertem Aufbau. Die Motive 1, 2 und 4 sind RuvC-ähnliche Motive, während das Motiv 3 ein HNH-Motiv ist. In einigen Fällen umfasst ein geeignetes ortsspezifisch modifizierendes Polypeptid eine Aminosäuresequenz, die über vier Motive verfügt, wobei jedes der Motive 1–4 mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Motiven 1–4 der in **Fig. 3A** dargestellten Cas9/Csn1-Aminosäuresequenz (SEQ ID NOs: 260–263, wie in Tabelle 1 gezeigt) oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist (siehe **Fig. 5** für ein Alignment der Motive 1–4 aus unterschiedlichen Cas9-Sequenzen). In einigen Fällen umfasst ein geeignetes ortsspezifisch modifizierendes Polypeptid eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Cas9/Csn1-Aminosäuresequenz aufweist oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen. Ein beliebiges wie oben definiertes Cas9-Protein kann als ein ortsspezifisch modifizierendes Polypeptid oder als Teil eines chimären ortsspezifisch modifizierenden Polypeptids für die vorliegenden Verfahren verwendet werden.

[0258] Die Nuklease-Aktivität spaltet die Ziel-DNA um Doppelstrangbrüche zu erzeugen. Diese Brüche werden anschließend durch die Zelle auf einem von zwei Wegen repariert: nicht-homologes End Joining und Homologie-gesteuerte Reparatur (**Fig. 2**). Bei dem nicht-homologen End Joining (NHEJ) werden die Doppelstrangbrüche durch direkte Ligation der Bruchenden miteinander repariert. Somit wird kein neues Nukleinsäure-Material in die Stelle inseriert, obwohl einiges Nukleinsäure-Material verloren gehen kann, was zu einer Deletion führt. Bei der Homologie-gesteuerten Reparatur wird ein Donor-Polynukleotid mit Homologie zu der gespaltenen Ziel-DNA-Sequenz als Matrize für die Reparatur der Ziel-DNA-Sequenz verwendet, was zu einem Transfer von genetischer Information vom Donor-Polynukleotid zur Ziel-DNA führt. Somit kann neues Nukleinsäure-Material in die Stelle inseriert/kopiert werden. In einigen Fällen wird eine Ziel-DNA mit einem vorliegenden Donor-Polynukleotid in Kontakt gebracht. In einigen Fällen wird ein vorliegendes Donor-Polynukleotid in eine vorliegende Zelle eingebracht. Die Modifikationen der Ziel-DNA führen aufgrund des NHEJ und/oder der Homologie-gesteuerten Reparatur beispielsweise zu Gen-Korrektur, Gen-Ersatz, Gen-Tagging, Transgen-Insertion, Nukleotid-Deletion, Gen-Zerstörung, Gen-Mutation, usw.

[0259] Somit kann eine Spaltung von DNA durch ein ortsspezifisch modifizierendes Polypeptid verwendet werden, um Nukleinsäure-Material aus einer Ziel-DNA-Sequenz zu deletieren (z. B., um ein Gen zu zerstören, das Zellen empfänglich für eine Infektion macht (z. B. die Gene CCR5 oder CXCR4, welche T-Zellen empfänglich für eine Infektion mit HIV machen), um krankheitsverursachende Trinukleotid-Repeat-Sequenzen in

Neuronen zu entfernen, um Gen-Knockouts und Mutationen als Krankheitsmodell für die Forschung zu erzeugen, usw.), indem man die Ziel-DNA-Sequenz spaltet und es der Zelle erlaubt, die Sequenz in Abwesenheit eines exogen bereitgestellten Donor-Polynukleotids zu reparieren. Somit können die vorliegenden Verfahren verwendet werden, um einen Gen-Knockout zu erzeugen (was zum vollständigen Erliegen einer Transkription oder zu einer veränderten Transkription führt) oder um genetisches Material durch einen Knockin in einen ausgewählten Locus in der Ziel-DNA einzuführen.

[0260] Wenn eine auf DNA gerichtet RNA und ein ortsspezifisch modifizierendes Polypeptid gemeinsam mit einer Donor-Polynukleotidsequenz, welche mindestens ein Segment mit Homologie zur Ziel-DNA-Sequenz umfasst, verabreicht werden, können die vorliegenden Verfahren alternativ dazu verwendet werden, um Nukleinsäure-Material einer Ziel-DNA-Sequenz zuzufügen, d. h. insertieren oder ersetzen, (z. B., um Nukleinsäure, die für ein Protein, eine siRNA, eine miRNA, usw. kodiert, durch "Knockin" einzuführen), um einen Anhang zuzufügen (z. B. 6xHis, ein fluoreszierendes Protein (z. B. ein grün fluoreszierendes Protein, ein gelb fluoreszierendes Protein, usw.), Hämagglutinin (HA), FLAG, usw.), um eine regulatorische Sequenz an ein Gen anzufügen (z. B. Promotor, Polyadenylierungssignal, interne Ribosomen-Eintrittsstelle (IRES), 2A-Peptid, Startcodon, Stoppcodon, Spleiß-Signal, Ortsbestimmungssignal, usw.), um eine Nukleinsäuresequenz zu modifizieren (z. B., um eine Mutation einzubringen, usw.), und ähnliches. Somit ist ein Komplex, der eine auf DNA gerichtete RNA und ein ortsspezifisch modifizierendes Polypeptid umfasst, in einer beliebigen Anwendung in vitro und in vivo nützlich, bei der es gewünscht ist, die DNA in einer ortsspezifischen, d. h. "gezielten" Weise zu modifizieren, z. B. durch Gen-Knockout, Gen-Knockin, Gen-Bearbeitung, Gen-Tagging, usw., wie es beispielsweise bei der Gentherapie verwendet wird, z. B. um eine Krankheit zu behandeln oder als ein anti-virales, anti-pathogenes, oder anti-Krebs-Therapeutikum, bei der Erzeugung von genetisch modifizierten Organismen in der Landwirtschaft, bei der großtechnischen Herstellung von Proteinen durch Zellen für therapeutische, diagnostische oder Forschungszwecke, für die Induktion von iPS-Zellen, für die biologische Forschung, für das Targeting von Genen von Pathogenen für die Deletion oder für das Ersetzen, usw.

[0261] In einigen Ausführungsformen umfasst das ortsspezifisch modifizierende Polypeptid eine modifizierte Form des Cas9/Csn1-Proteins. In einigen Fällen umfasst die modifizierte Form des Cas9/Csn1-Proteins eine Aminosäure-Änderung (z. B. eine Deletion, Insertion oder Substitution), welche die natürlich vorkommende Nuklease-Aktivität des Cas9/Csn1-Proteins verringert. Beispielsweise weist die modifizierte Form des Cas9/Csn1-Proteins in einigen Fällen weniger als 50%, weniger als 40%, weniger als 30%, weniger als 20%, weniger als 10%, weniger als 5% oder weniger als 1% der Nuklease-Aktivität des entsprechenden Wildtyp-Cas9/Csn1-Polypeptids auf. In einigen Fällen weist die modifizierte Form des Cas9/Csn1-Polypeptids keine wesentliche Nuklease-Aktivität auf. Wenn ein vorliegendes ortsspezifisch modifizierendes Polypeptid eine modifizierte Form des Cas9/Csn1-Polypeptids ist, das im Wesentlichen keine Nuklease-Aktivität aufweist, so kann es als "dCas9" bezeichnet werden.

[0262] In einigen Ausführungsformen ist die modifizierte Form des Cas9/Csn1-Polypeptids eine D10A(Asparat zu Alanin in Aminosäure-Position 10 von SEQ ID NO: 8)-Mutation (oder die entsprechende Mutation in einem der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine), die den komplementären Strang der Ziel-DNA spalten kann, jedoch eine verringerte Fähigkeit zur Spaltung des nicht komplementären Strangs der Ziel-DNA aufweist (was zu einem Einzelstrangbruch (single strand break, SSB) anstelle eines DSB führt; siehe **Fig. 11**). In einigen Ausführungsformen ist die modifizierte Form des Cas9/Csn1-Polypeptids eine H840A (Histidin zu Alanin in Aminosäure-Position 840 von SEQ ID NO: 8)-Mutation (oder die entsprechende Mutation in einem der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine), die den nicht-komplementären Strang der Ziel-DNA spalten kann, jedoch eine verringerte Fähigkeit zur Spaltung des komplementären Strangs der Ziel-DNA aufweist (was zu einem Einzelstrangbruch (SSB) anstelle eines DSB führt; siehe **Fig. 11**). Die Verwendung der D10A- oder H840A-Variante von Cas9 (oder den entsprechenden Mutationen in einem der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine) kann das erwartete biologische Ergebnis verändern, da ein nicht-homologes End Joining (NHEJ) mit höherer Wahrscheinlichkeit auftreten wird, wenn DSBs vorhanden sind und keine SSBs. Wenn es daher in einigen Fällen gewünscht ist, die Wahrscheinlichkeit eines DSB zu verringern (und somit die Wahrscheinlichkeit von NHEJ zu verringern), kann eine D10A- oder eine H840A-Variante von Cas9 verwendet werden. Andere Reste können mutiert werden, um die gleiche Wirkung zu erzielen (d. h. den einen oder den anderen Teil der Nuklease zu inaktivieren). Als nicht einschränkende Beispiele können die Reste D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, und/oder A987 (oder die entsprechenden Mutationen in einem der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine) verändert werden (d. h. substituiert) (siehe **Fig. 3**, **Fig. 5**, **Fig. 11A** und Tabelle 1 für weitere Informationen bezüglich der Konservierung von Cas9-Aminosäure-Resten). Darüber hinaus sind auch andere Mutationen, bei denen es sich nicht um Alanin-Substitutionen handelt, geeignet. In einigen Ausführungsformen, bei denen ein ortsspezifisches Polypeptid (z. B. ein ortsspezifisch modifizierendes Polypeptid) eine verringerte katalytische

Aktivität aufweist (z. B., wenn ein Cas9-Protein eine D10-, G12-, G17-, E762-, H840-, N854-, N863-, H982-, H983-, A984-, D986- und/oder eine A987-Mutation aufweist, z. B. D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A und/oder D986A), kann das Polypeptid weiterhin an Ziel-DNA in spezifischer Weise binden (weil es weiterhin durch eine auf DNA gerichtete RNA zu der Ziel-DNA-Sequenz geführt wird), solange es die Fähigkeit beibehält, mit der auf DNA gerichteten RNA zu interagieren.

[0263] In einigen Ausführungsformen umfasst die modifizierte Form des Cas9/Csn1-Polypeptids sowohl die D10A- als auch die H840A-Mutation (oder die entsprechenden Mutationen in einem der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine), sodass das Polypeptid eine verringerte Fähigkeit zur Spaltung sowohl des komplementären als auch des nicht-komplementären Strangs der Ziel-DNA aufweist (d. h. die Variante kann keine wesentliche Nuklease-Aktivität besitzen). Andere Reste können mutiert werden, um die gleiche Wirkung zu erzielen (d. h. den einen oder anderen Teil der Nuklease zu inaktivieren). Als nicht einschränkende Beispiele können die Reste D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 und/oder A987 (oder die entsprechenden Mutationen in einem der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine) verändert werden (d. h. substituiert) (siehe **Fig. 3**, **Fig. 5**, **Fig. 11A** und Tabelle 1 für weitere Informationen bezüglich der Konservierung von Cas9-Aminosäure-Resten). Darüber hinaus sind auch Mutationen, bei denen es sich nicht um Alanin-Substitutionen handelt, geeignet.

[0264] In einigen Ausführungsformen umfasst das ortsspezifisch modifizierende Polypeptid eine heterologe Sequenz (z. B. eine Fusion). In einigen Ausführungsformen kann eine heterologe Sequenz für eine subzelluläre Ortsbestimmung des ortsspezifisch modifizierenden Polypeptids sorgen (z. B. ein Signal für eine nukleare Ortsbestimmung (nuclear localization signal, NLS) für die gezielte Bereitstellung im Zellkern; ein Signal für eine mitochondriale Ortsbestimmung für die gezielte Bereitstellung in den Mitochondrien; ein Signal für eine Chloroplasten-Ortsbestimmung für die gezielte Bereitstellung in einem Chloroplasten; ein ER-Retentionssignal, und ähnliches). In einigen Ausführungsformen kann die heterologe Sequenz einen Tag bereitstellen, der die Verfolgung oder Aufreinigung erleichtert (z. B. ein fluoreszierendes Protein, z. B. grünes fluoreszierendes Protein (GFP), YFP, RFP, CFP, mCherry, tdTomato, und ähnliche; einen HIS-Tag, z. B. einen 6xHis-Tag; einen Hämagglutinin(HA)-Tag; einen FLAG-Tag; einen Myc-Tag; und ähnliches). In einigen Ausführungsformen kann die heterologe Sequenz für eine erhöhte oder verringerte Stabilität sorgen.

[0265] In einigen Ausführungsformen kann ein vorliegendes ortsspezifisch modifizierendes Polypeptid Codon-optimiert sein. Die Art der Optimierung ist im Stand der Technik bekannt und umfasst die Mutation von Fremd-abgeleiteter DNA in der Weise, dass die Codon-Präferenz des gewünschten Wirtsorganismus oder der gewünschten Wirtszelle nachgeahmt wird, wobei das gleiche Protein kodiert wird. Somit werden die Codons verändert, wobei das kodierte Protein unverändert bleibt. Wenn beispielsweise die gewünschte Ziel-Zelle eine humane Zelle ist, wäre ein humanes Codon-optimiertes Cas9 (oder eine Variante, z. B. eine enzymatisch inaktive Variante) ein geeignetes ortsspezifisch modifizierendes Polypeptid (siehe z. B. SEQ ID NOs: 256). Jedes geeignete ortsspezifisch modifizierende Polypeptid (z. B. jedes Cas9, wie beispielsweise eine der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Sequenzen) kann Codon-optimiert werden. Als weiteres nicht einschränkendes Beispiel wäre in dem Fall, dass die gewünschte Wirtszelle eine Mauszelle ist, ein Maus-Codon-optimiertes Cas9 (oder eine Variante, z. B. eine enzymatisch inaktive Variante) ein geeignetes ortsspezifisch modifizierendes Polypeptid. Während Codon-Optimierung nicht erforderlich ist, ist diese akzeptabel und kann in einigen Fällen bevorzugt sein.

[0266] In einigen Ausführungsformen werden eine vorliegende auf DNA gerichtete RNA und ein vorliegendes ortsspezifisch modifizierendes Polypeptid als induzierbares System verwendet, um die Gen-Expression in bakteriellen Zellen abzuschalten. In einigen Fällen werden Nukleinsäuren, die für geeignete auf DNA gerichtete RNA und/oder ein geeignetes ortsspezifisch modifizierendes Polypeptid kodieren, in ein Chromosom einer Ziel-Zelle eingebaut und stehen unter Kontrolle eines induzierbaren Promotors. Wenn die auf DNA gerichtete RNA und/oder das ortsspezifische Polypeptid induziert werden, wird die Ziel-DNA an der Stelle von Interesse (z. B. an einem Ziel-Gen auf einem separaten Plasmid) gespalten (oder anderweitig modifiziert), wenn sowohl die auf DNA gerichtete RNA als auch das ortsspezifisch modifizierende Polypeptid vorhanden sind und einen Komplex bilden. Somit werden in einigen Fällen bakterielle Expressionsstämme so erstellt, dass sie Nukleinsäuresequenzen umfassen, welche für ein geeignetes ortsspezifisch modifizierendes Polypeptid in dem bakteriellen Genom und/oder für eine geeignete auf DNA gerichtete RNA auf einem Plasmid (z. B. unter Kontrolle eines induzierbaren Promotors) kodieren, was Experimente ermöglicht, bei denen die Expression eines beliebigen Ziel-Gens (exprimiert ausgehend von einem getrennten Plasmid, was in den Stamm eingebracht wird) kontrolliert werden kann, indem die Expression der auf DNA gerichteten RNA und des ortsspezifischen Polypeptids induziert werden.

[0267] In einigen Fällen weist das ortsspezifisch modifizierende Polypeptid eine enzymatische Aktivität auf, welche die Ziel-DNA in einer Weise modifiziert, bei der keine Doppelstrangbrüche eingeführt werden. Die enzymatische Aktivität von Interesse, die verwendet werden kann, um Ziel-DNA zu modifizieren (z. B. durch Fusion eines heterologen Polypeptids mit enzymatischer Aktivität an ein ortsspezifisch modifizierendes Polypeptid, wodurch ein chimäres ortsspezifisch modifizierendes Polypeptid erzeugt wird) umfasst, ist jedoch nicht beschränkt auf, Methyltransferase-Aktivität, Demethylase-Aktivität, DNA-Reparatur-Aktivität, DNA-Schädigungs-Aktivität, Deaminierungs-Aktivität, Dismutase-Aktivität, Alkylierungs-Aktivität, Depurinierungs-Aktivität, Oxidations-Aktivität, Pyrimidin-Dimer-bildende Aktivität, Integrase-Aktivität, Transposase-Aktivität, Rekombinase-Aktivität, Polymerase-Aktivität, Ligase-Aktivität, Helicase-Aktivität, Photolyase-Aktivität oder Glykosylase-Aktivität. Methylierung und Demethylierung werden im Stand der Technik als wichtige Arten epigenetischer Genregulation angesehen, während DNA-Schädigung und DNA-Reparaturaktivität für das Überleben der Zelle und für eine angemessene Aufrechterhaltung des Genoms in Reaktion auf Umweltstress essentiell sind.

[0268] Die vorliegenden Verfahren können bei der epigenetischen Modifikation von Ziel-DNA angewendet werden, und sie können dazu verwendet werden, die epigenetische Modifikation von Ziel-DNA an einer beliebigen Stelle in einer Ziel-DNA zu kontrollieren, indem die gewünschte komplementäre Nukleinsäuresequenz durch genetische Modifikation in das auf DNA gerichtete Segment einer auf DNA gerichteten RNA eingebaut wird. Die vorliegenden Verfahren können auch bei der gewollten und kontrollierten Schädigung von DNA an einer beliebigen Stelle innerhalb der Ziel-DNA verwendet werden. Die vorliegenden Verfahren können auch bei der sequenzspezifischen und kontrollierten Reparatur von DNA an einer beliebigen Stelle innerhalb der Ziel-DNA verwendet werden. Verfahren, um DNA-modifizierende, enzymatische Aktivitäten zu bestimmten Stellen einer Ziel-DNA zu bringen, können sowohl für die Forschung als auch bei klinischen Applikationen verwendet werden.

[0269] In einigen Fällen weist das ortsspezifisch modifizierende Polypeptid eine Aktivität auf, die die Transkription von Ziel-DNA moduliert (z. B. im Falle eines chimären ortsspezifisch modifizierenden Polypeptids, usw.). In einigen Fällen wird ein chimäres ortsspezifisch modifizierendes Polypeptid, das ein heterologes Polypeptid umfasst, welches die Fähigkeit zur Erhöhung oder Verringerung der Transkription aufweist (z. B. transkriptionales Aktivator-Polypeptid oder transkriptionales Repressor-Polypeptid), verwendet, um die Transkription einer Ziel-DNA an einer bestimmten Stelle in einer Ziel-DNA zu erhöhen oder zu verringern, was durch das auf DNA gerichtete Segment der auf DNA gerichteten RNA gesteuert wird. Beispiele für Ausgangs-Polypeptide für die Bereitstellung eines chimären ortsspezifisch modifizierenden Polypeptids mit transkriptionsmodulierenden Eigenschaften umfassen, sind jedoch nicht beschränkt auf, die leicht-induzierbaren Transkriptionsregulatoren, auf Small Molecule/Wirkstoff ansprechende Transkriptionsregulatoren, Transkriptionsfaktoren, Transkriptionsrepressoren, usw. In einigen Fällen wird das vorliegende Verfahren verwendet, um die Expression einer als Ziel bestimmten kodierenden RNA (Protein-kodierendes Gen) und/oder einer als Ziel bestimmten nicht-kodierenden RNA (z. B. tRNA, rRNA, snoRNA, siRNA, miRNA, long ncRNA, usw.) zu kontrollieren.

[0270] In einigen Fällen weist das ortsspezifisch modifizierende Polypeptid eine enzymatische Aktivität auf, die ein mit DNA assoziiertes Polypeptid modifiziert (z. B. ein Histon). In einigen Ausführungsformen ist die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität (d. h. eine ubiquitinierende Aktivität), eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität, eine Demyristoylierungs-Aktivität, eine Glycosylierungs-Aktivität (z. B. von einer O-GlcNAc-Transferase) oder eine Deglycosylierungs-Aktivität. Die vorliegend aufgeführten enzymatischen Aktivitäten katalysieren kovalente Modifikationen an Proteinen. Von solchen Modifikationen ist im Stand der Technik bekannt, dass sie die Stabilität oder Aktivität des Zielproteins verändern (z. B. kann eine Phosphorylierung aufgrund von Kinase-Aktivität eine Proteinaktivität in Abhängigkeit vom Zielprotein stimulieren oder herunterregeln). Von besonderem Interesse als Protein-Ziele sind Histone. Von Histon-Proteinen ist im Stand der Technik bekannt, dass sie DNA binden und Komplexe bilden, die als Nukleosomen bekannt sind. Histone können modifiziert sein (z. B. durch Methylierung, Acetylierung, Ubiquitinierung, Phosphorylierung), um strukturelle Veränderungen in der umgebenden DNA auszulösen, und auf diese Weise die Zugänglichkeit von potentiell großen Teilen der DNA zu interagierenden Faktoren, wie beispielsweise Transkriptionsfaktoren, Polymerasen, und ähnlichem, zu kontrollieren. Ein einzelnes Histon kann auf vielfältige Weise modifiziert werden und in zahlreichen verschiedenen Kombinationen (z. B. ist die Trimethylierung von Lysin 27 von Histon 3, H3K27, mit DNA-Regionen von reprimierter Transkription assoziiert, während die Trimethylierung von Lysin 4 von Histon 3, H3K4, mit DNA-Regionen aktiver Transkription assoziiert ist). Somit kann ein ortsspezifisch modifizierendes Polypeptid mit Histon-modifizierender Aktivität bei der ortsspezifischen Kontrolle von DNA-Strukturen verwendet werden, und es kann dazu verwendet werden,

die Histon-Modifikationsmuster in einem ausgewählten Bereich der Ziel-DNA zu verändern. Solche Verfahren finden sowohl in der Forschung als auch bei klinischen Applikationen Anwendung.

[0271] In einigen Ausführungsformen werden mehrere auf DNA gerichtete RNAs gleichzeitig verwendet, um gleichzeitig unterschiedliche Stellen der gleichen Ziel-DNA oder von verschiedenen Ziel-DNAs zu modifizieren. In einigen Ausführungsformen sind zwei oder mehr auf DNA gerichtete RNAs auf das gleiche Gen oder das gleiche Transkript oder den gleichen Locus gerichtet. In einigen Ausführungsformen sind zwei oder mehr auf DNA gerichtete RNAs auf unterschiedliche, nicht miteinander verwandte Loci gerichtet. In einigen Ausführungsformen sind zwei oder mehr auf DNA gerichtete RNAs auf verschiedene, jedoch verwandte Loci gerichtet.

[0272] In einigen Fällen wird das ortsspezifisch modifizierende Polypeptid direkt als Protein bereitgestellt. Als ein nicht einschränkendes Beispiel können Pilze (z. B. Hefe) mit exogenem Protein und/oder Nukleinsäure transformiert werden, indem eine Sphäroplasten-Transformation verwendet wird (siehe Kawai et al., *Bioeng Bugs*. 2010 Nov–Dec; 1(6): 395–403: "Transformation of *Saccharomyces cerevisiae* and other fungi: methods and possible underlying mechanism"; und Tanka et al., *Nature*. 2004 Mar 18; 428(6980): 323–8: "Conformational variations in an infectious protein determine prion strain differences"); wobei beide vorliegend durch Bezugnahme in ihrer Gesamtheit eingeschlossen sind. Somit kann ein ortsspezifisch modifizierendes Polypeptid (z. B. Cas9) in einen Sphäroplasten eingebracht werden (mit oder ohne Nukleinsäure, die für eine auf DNA gerichtete RNA kodiert, und mit oder ohne Donor-Polynukleotid), und der Sphäroplast kann verwendet werden, um den Inhalt in eine Hefezelle einzubringen. Ein ortsspezifisch modifizierendes Polypeptid kann durch jedes zweckdienliche Verfahren in eine Zelle eingebracht werden (einer Zelle bereitgestellt werden); solche Verfahren sind dem Fachmann bekannt. Als weiteres nicht einschränkendes Beispiel kann ein ortsspezifisch modifizierendes Polypeptid direkt in eine Zelle indiziert werden (z. B. mit oder ohne Nukleinsäure, die für eine auf DNA gerichtete RNA kodiert, und mit oder ohne Donor-Polynukleotid), z. B. in eine Zelle eines Zebrafisch-Embryos, den Pronukleus einer befruchteten Maus-Oozyte, usw. Ziel-Zellen von Interesse

[0273] In einigen der obigen Anwendungen können die vorliegenden Verfahren verwendet werden, um eine DNA-Spaltung, eine DNA-Modifikation und/oder eine transkriptionale Modulation in mitotischen oder post-mitotischen Zellen in vivo und/oder ex vivo und/oder in vitro zu induzieren (z. B., um genetisch modifizierte Zellen zu erzeugen, die wieder in ein Individuum eingebracht werden können). Weil die auf DNA gerichtete RNA Spezifität bereitstellt, indem sie an die Ziel-DNA hybridisiert, kann eine mitotische und/oder eine post-mitotische Zelle von Interesse in den offenbarten Verfahren eine Zelle von einem beliebigen Organismus umfassen (z. B. eine bakterielle Zelle, eine Archaeen-Zelle, eine Zelle eines einzelligen eukaryontischen Organismus, eine Pflanzenzelle, eine Algenzelle, z. B. *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens* C. Agardh, und ähnliche, eine Pilz-Zelle (z. B. eine Hefezelle), eine tierische Zelle, eine Invertebraten-Zelle (z. B. Fruchtfliege, Nesseltier, Stachelhäuter, Fadenwurm, usw.), eine Zelle von einem Wirbeltier (z. B. Fisch, Amphibie, Reptil, Vogel, Säugetier), eine Zelle von einem Säugetier, eine Zelle von einem Nagetier, eine Zelle von einem Menschen, usw.).

[0274] Jede Art von Zelle kann von Interesse sein (z. B. eine Stammzelle, z. B. eine embryonische Stammzelle (ES), eine induzierte pluripotente Stammzelle (iPS), eine Keimzelle; eine somatische Zelle, z. B. ein Fibroblast, eine hämatopoietische Zelle, ein Neuron, eine Muskelzelle, eine Knochenzelle, eine Hepatozyte, eine Bauchspeicheldrüsen-Zelle; eine in vitro oder in vivo vorliegende embryonische Zelle eines Embryos in einem beliebigen Stadium, z. B. das 1-Zell-, 2-Zell-, 4-Zell-, 8-Zell-, usw., Stadium eines Zebrafisch-Embryos; usw.). Die Zellen können von etablierten Zelllinien stammen, oder es kann sich um primäre Zellen handeln, wobei „primäre Zellen“, „primäre Zelllinien“ und „primäre Kulturen“ vorliegend austauschbar verwendet werden und sich auf Zellen und Zellkulturen beziehen, die von einem Subjekt abgeleitet wurden, und denen es erlaubt wurde, in vitro für eine begrenzte Anzahl von Passagen zu wachsen, d. h. Aufspaltungen der Kultur. Beispielsweise handelt es sich bei primären Kulturen um Kulturen, die kein Mal, einmal, zweimal, viermal, fünfmal, zehnmal oder fünfzehnmal passagiert wurden, jedoch nicht häufig genug, um die Krisenphase zu durchlaufen. Üblicherweise werden die primären Zelllinien der Erfindung für weniger als zehn Passagen in vitro gehalten. Die Ziel-Zellen sind in vielen Ausführungsformen einzellige Organismen oder werden in Kultur angezüchtet.

[0275] Wenn die Zellen primäre Zellen sind, können sie von einem Individuum durch zweckmäßige Verfahren erhalten werden. Beispielsweise können Leukozyten in praktischer Weise durch Apherese, Leukozytapherese, Dichtegradientenzentrifugation, usw. geerntet werden, während Zellen von Geweben wie Haut, Muskel, Knochenmark, Milz, Leber, Bauchspeicheldrüse, Lunge, Darm, Magen, usw. praktischerweise durch Biopsie erhalten werden. Eine geeignete Lösung kann für die Dispersion oder Suspension der geernteten Zellen verwendet werden. Eine solche Lösung wird üblicherweise eine ausgewogene Salzlösung sein, z. B. normale Salzlösung, Phosphat-gepufferte Salzlösung (PBS), ausgewogene Salzlösung nach Hanks, usw., wobei die-

se in praktischer Weise mit fötalem Kälberserum oder mit anderen natürlich vorkommenden Faktoren supplementiert sein kann, in Verbindung mit einem akzeptablen Puffer in geringer Konzentration, üblicherweise von 5–25 mM. Praktische Puffer umfassen HEPES, Phosphatpuffer, Lactatpuffer, usw. Die Zellen können sofort verwendet werden oder sie können für längere Zeiträume gelagert oder eingefroren werden, wobei sie beim Auftauen wieder verwendet werden können. In solchen Fällen werden die Zellen üblicherweise in 10% DMSO, 50% Serum, 40% gepuffertes Medium eingefroren, oder in anderen Lösungen, die üblicherweise im Stand der Technik verwendet werden, um Zellen bei solchen Einfriertemperaturen aufzubewahren, und sie können nach einem Verfahren, das im Stand der Technik herkömmlicherweise zum Auftauen von gefrorenen kultivierten Zellen verwendet werden, aufgetaut werden. Nukleinsäuren, die für eine vorliegende auf DNA gerichtete RNA und/oder für ein vorliegendes ortsspezifisch modifizierendes Polypeptid kodieren

[0276] In einigen Ausführungsformen umfasst ein vorliegendes Verfahren Schritte, bei denen man ein oder mehrere Nukleinsäuren, die Nukleotidsequenzen umfassen, welche für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid und/oder eine Donor-Polynukleotid kodieren, mit einer Ziel-DNA in Kontakt bringt oder diese in eine Zelle (oder in eine Population von Zellen) einbringt. Geeignete Nukleinsäuren, die Nukleotidsequenzen umfassen, welche für eine auf DNA gerichtete RNA oder für ein ortsspezifisch modifizierendes Polypeptid kodieren, umfassen Expressionsvektoren, wobei ein Expressionsvektor, der eine Nukleotidsequenz umfasst, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, ein „rekombinanter Expressionsvektor“ ist.

[0277] In einigen Ausführungsformen ist der rekombinante Expressionsvektor ein virales Konstrukt, z. B. ein rekombinantes Adeno-assoziiertes Virus-Konstrukt (siehe z. B. US Patent Nr. 7,078,387), ein rekombinantes adenovirales Konstrukt, ein rekombinantes lentivirales Konstrukt, usw.

[0278] Geeignete Expressionsvektoren umfassen (sind jedoch nicht beschränkt auf) virale Vektoren (z. B. virale Vektoren, die auf einem Vaccinia-Virus basieren; Poliovirus; Adenovirus (siehe beispielsweise Lee et al., Invest Ophthalmol Vis Sci 35: 2543 2549, 1994; Borrás et al., Gene Ther 6: 515 524, 1999; Li and Davidson, PNAS 92: 7700 7704, 1995; Sakamoto et al., H Gene Ther 5: 1088 1097, 1999; WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 und WO 95/00655); Adeno-assoziiertes Virus (siehe z. B., Ali et al., Hum Gene Ther 9: 81 86, 1998; Flannery et al., PNAS 94: 6916 6921, 1997; Bennett et al., Invest Ophthalmol Vis Sci 38: 2857 2863, 1997; Jomary et al., Gene Ther 4: 683 690, 1997, Rolling et al., Hum Gene Ther 10: 641 648, 1999; Ali et al., Hum Mol Genet 5: 591 594, 1996; Srivastava in WO 93/09239, Samulski et al., J. Vir. (1989) 63: 3822–3828; Mendelson et al., Virol. (1988) 166: 154–165; und Flotte et al., PNAS (1993) 90: 10613–10617); SV40; Herpes simplex-Virus; humanes Immundefizienzvirus (siehe z. B. Miyoshi et al., PNAS 94: 10319 23, 1997; Takahashi et al., J Virol 73: 7812 7816, 1999); einen retroviralen Vektor (z. B. murines Leukämievirus, Milz-Nekrose-Virus, und Vektoren, die von Retroviren abgeleitet sind, wie beispielsweise Rous-Sarkom-Virus, Harvey-Sarkom-Virus, Vogel-Leukose-Virus, einem Lentivirus, humanes Immundefizienzvirus, myeloproliferatives Sarkom-Virus, und Brusttumor-Virus); und ähnliche.

[0279] Zahlreiche geeignete Expressionsvektoren sind den Fachleuten bekannt, und zahlreiche sind kommerziell verfügbar. Die folgenden Vektoren werden lediglich beispielhaft genannt; für eukaryotische Wirtszellen: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, und pSVLSV40 (Pharmacia). Es kann jedoch ein beliebiger anderer Vektor verwendet werden, solange dieser mit der Wirtszelle kompatibel ist.

[0280] In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit einem Kontrollelement verbunden, z. B. einem transkriptionalen Kontrollelement, wie beispielsweise einem Promotor. Das transkriptionale Kontrollelement kann entweder in einer eukaryotischen Zelle, z. B. einer Säugetierzelle, funktionsfähig sein, oder in einer prokaryotischen Zelle (z. B. in einer bakteriellen Zelle oder einer Archaeen-Zelle). In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise mit mehreren Kontrollelementen verbunden, welche die Expression einer Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid kodiert, sowohl in prokaryontischen als auch in eukaryontischen Zellen ermöglichen.

[0281] Abhängig von dem verwendeten Wirt-/Vektor-System können verschiedene geeignete Transkriptions- und Translations-Kontrollelementen verwendet werden, einschließlich konstitutiver und induzierbarer Promotoren, Transkriptions-Verstärkungselemente, Transkriptions-Terminatoren, usw., in dem Expressionsvektor verwendet werden (z. B. U6-Promoter, H1-Promoter, usw.; siehe oben) (siehe z. B. Bitter et al. (1987) Methods in Enzymology, 153: 516–544).

[0282] In einigen Ausführungsformen kann eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid als RNA bereitgestellt werden. In solchen Fällen kann die auf DNA gerichtete RNA und/oder die für das ortsspezifisch modifizierende Polypeptid kodierende RNA durch direkte chemische Synthese hergestellt werden, oder sie kann in vitro von einer DNA transkribiert werden, welche für die auf DNA gerichtete RNA kodiert. Verfahren zur Synthese von RNA ausgehend von einer DNA-Matrize sind im Stand der Technik hinreichend bekannt. In einigen Fällen wird die auf DNA gerichtete RNA und/oder die für das ortsspezifisch modifizierende Polypeptid kodierende RNA in vitro unter Verwendung eines RNA-Polymerase-Enzyms synthetisiert (z. B. T7-Polymerase, T3-Polymerase, SP6-Polymerase, usw.). Nach der Synthese kann die RNA direkt mit einer Ziel-DNA in Kontakt gebracht werden oder sie kann in eine Zelle eingebracht werden, wobei ein beliebiges gut bekanntes Verfahren zur Einbringung von Nukleinsäuren in Zellen verwendet werden kann (z. B. Mikroinjektion, Elektroporation, Transfektion, usw.).

[0283] Nukleotide, die für eine auf DNA gerichtete RNA kodieren (entweder als DNA oder RNA eingebracht) und/oder für ein ortsspezifisch modifizierendes Polypeptid (als DNA oder RNA eingebracht) und/oder ein Donor-Polynukleotid können den Zellen unter Verwendung hinreichend bekannter Transfektionsverfahren bereitgestellt werden; siehe z. B. Angel and Yanik (2010) PLoS ONE 5(7): e11756, und die kommerziell verfügbaren TransMessenger®-Reagenzien von Qiagen, das Stemfect™-RNA-Transfektions-Kit von Stemgent, und das TransIT®-mRNA-Transfektions-Kit von Mirus Bio LLC. Siehe auch Beumer et al. (2008) "Efficient gene targeting in Drosophila by direkt embryo injection with zinc-finger nucleases", PNAS 105(50): 19821–19826. Alternativ dazu können Nukleinsäuren, die für eine auf DNA gerichtete RNA und/oder für ein ortsspezifisch modifizierendes Polypeptid und/oder für ein chimäres ortsspezifisch modifizierendes Polypeptid und/oder für ein Donor-Polynukleotid kodieren, auf DNA-Vektoren bereitgestellt werden. Zahlreiche Vektoren, z. B. Plasmide, Cosmide, kleine zirkuläre Sequenzen (minicircles), Phagen, Viren, usw., die für die Übertragung von Nukleinsäuren in Zellen nützlich sind, sind verfügbar. Die Vektoren, welche die Nukleinsäure(n) umfassen, können episomal beibehalten werden, z. B. als Plasmide, kleine zirkuläre Sequenzen (minicircles), DNAs, Viren, wie beispielsweise Cytomegalovirus, Adenovirus, usw., oder sie können in das Genom der Ziel-Zelle durch homologe Rekombination oder zufällige Integration, z. B. Retrovirus-vermittelte Vektoren, wie beispielsweise MMLV, HIV-1, ALV, usw., integriert werden.

[0284] Vektoren können den vorliegenden Zellen direkt bereitgestellt werden. Mit anderen Worten werden die Zellen mit Vektoren in Kontakt gebracht, welche die Nukleinsäure umfassen, die für die auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein chimäres ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid kodieren, sodass die Vektoren von den Zellen aufgenommen werden. Verfahren zum Inkontaktbringen von Zellen mit Nukleinsäure-Vektoren, bei denen es sich um Plasmide handelt, sind im Stand der Technik bekannt, einschließlich Elektroporation, Kalziumchlorid-Transfektion, Mikroinjektion und Lipofektion. Für die Bereitstellung von viralem Vektor werden die Zellen mit viralen Partikeln in Kontakt gebracht, welche die Nukleinsäure umfassen, die für eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein chimäres ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid kodieren. Retroviren, z. B. Lentiviren, sind für die Verfahren der Erfindung besonders geeignet. Üblicherweise verwendete retrovirale Vektoren sind "defekt", d. h. sie sind nicht in der Lage, die viralen Proteine, welche für eine produktive Infektion erforderlich sind, zu erzeugen. Vielmehr erfordert die Replikation des Vektors das Wachstum in einer Verpackungszelllinie. Um virale Partikel zu erzeugen, die die Nukleinsäuren von Interesse umfassen, werden die retroviralen Nukleinsäuren, welche die Nukleinsäure umfassen, durch die Verpackungszelllinie in virale Kapside verpackt. Verschiedene Verpackungszelllinien stellen ein unterschiedliches Hüllprotein bereit (ekotroph, amphotroph oder xenotroph), welches in das Kapsid eingebaut wird, wobei das Hüllprotein die Spezifität des viralen Partikels für die Zellen bestimmt (ekotroph für Maus und Ratte; amphotroph für die meisten Säugetier-Zelltypen, einschließlich Mensch, Hund und Maus; und xenotroph für die meisten Säugetier-Zelltypen, außer murinen Zellen). Die geeignete Verpackungszelllinie kann verwendet werden, um sicher zu stellen, dass die Zellen durch die verpackten Viruspartikel angesteuert werden. Verfahren zur Einbringung der retroviralen Vektoren, welche die Nukleinsäuren umfassen, die für die Reprogrammierungsfaktoren kodieren, in Verpackungszelllinien und Verfahren zum Ernten der viralen Partikel, welche durch die Verpackungszelllinien hergestellt werden, sind im Stand der Technik hinreichend bekannt. Nukleinsäuren können darüber hinaus durch direkte Mikroinjektion eingebracht werden (z. B. Injektion von RNA in einen Zebrafisch-Embryo).

[0285] Vektoren, die verwendet werden, um die Nukleinsäuren, welche für auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein chimäres ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid kodieren, in den vorliegenden Zellen bereitzustellen, werden üblicherweise geeignete Promotoren zur Steuerung der Expression, d. h. zur transkriptionalen Aktivierung, der Nukleinsäure von Interesse umfassen. Mit anderen Worten wird die Nukleinsäure von Interesse in funktionsfähiger Weise mit

einem Promotor verbunden sein. Dies kann universell wirkende Promotoren umfassen, z. B. den CMV- β -Actin-Promotor, oder induzierbare Promotoren, wie beispielsweise Promotoren, die in bestimmten Zellpopulationen aktiv sind, oder die auf die Gegenwart von Wirkstoffen, wie beispielsweise Tetracyclin reagieren. Mit transkriptionaler Aktivierung ist gemeint, dass die Transkription in den Zielzellen mindestens etwa 10-fach, mindestens etwa 100-fach, insbesondere mindestens etwa 1000-fach über das Basisniveau erhöht wird. Darüber hinaus können Vektoren, die dazu verwendet werden, eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein chimäres ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid an die vorliegenden Zellen bereit zu stellen, Nukleinsäuresequenzen umfassen, die für in den Ziel-Zellen selektierbare Marker kodieren, so dass Zellen identifiziert werden können, welche die auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein chimäres ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid aufgenommen haben.

[0286] Eine vorliegende auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein chimäres ortsspezifisch modifizierendes Polypeptid kann stattdessen verwendet werden, um DNA zu kontaktieren, oder kann als RNA in Zellen eingebracht werden. Verfahren zur Einbringung von RNA in Zellen sind im Stand der Technik bekannt und umfassen, beispielsweise, direkte Injektion, Transfektion, oder andere Verfahren, die zur Einbringung von DNA verwendet werden.

[0287] Ein vorliegendes ortsspezifisch modifizierendes Polypeptid kann stattdessen Zellen als Polypeptid bereitgestellt werden. Ein solches Polypeptid kann wahlweise an eine Polypeptid-Domäne fusioniert sein, welche die Löslichkeit des Produkts erhöht. Die Domäne kann über eine definierte Protease-Schnittstelle an ein Polypeptid gebunden sein, z. B. eine TEV-Sequenz, die von einer TEV-Protease gespalten wird. Der Linker kann ferner ein oder mehrere flexible Sequenzen umfassen, z. B. von 1 bis 10 Glycin-Reste. In einigen Ausführungsformen wird die Spaltung des Fusionsproteins in einem Puffer durchgeführt, der die Löslichkeit des Produkts aufrecht erhält, z. B. in Gegenwart von 0,5 bis 2 M Harnstoff, in Gegenwart von Polypeptiden und/oder Polynukleotiden, welche die Löslichkeit erhöhen, und ähnlichem. Domänen von Interesse umfassen endosomolytische Domänen, z. B. die Influenza-HA-Domäne; und andere Polypeptide, die bei der Erzeugung z. B. einer IF2-Domäne, einer GST-Domäne, einer GRPE-Domäne und ähnlichem helfen. Das Polypeptid kann im Hinblick auf eine verbesserte Stabilität formuliert sein. Beispielsweise können die Peptide PEGyliert sein, wobei die Polyethylenoxy-Gruppe eine erhöhte Halbwertszeit im Blutstrom ermöglicht.

[0288] Zusätzlich oder alternativ dazu kann das vorliegende ortsspezifisch modifizierende Polypeptid an eine Durchdringungs-Polypeptid-Domäne fusioniert sein, um die Aufnahme durch die Zelle zu fördern. Verschiedene Durchdringungs-Domänen sind im Stand der Technik bekannt, und diese können bei den nicht-integrierenden Polypeptiden der vorliegenden Erfindung angewendet werden, einschließlich Peptiden, Peptidomimetika und Nicht-Peptid-Trägern. Beispielsweise kann ein Durchdringungs-Peptid von der dritten Alpha-Helix von des Transkriptionsfaktors Antennapedia von *Drosophila melanogaster*, der als Penetratin bezeichnet wird und die Aminosäuresequenz RQIKIWFQNRRMKWKK (SEQ ID NO:1) umfasst, abgeleitet sein. In einem weiteren Beispiel umfasst das Durchdringungs-Peptid die HIV-1 tat-Aminosäuresequenz der basischen Region, die beispielsweise die Aminosäuren 49–57 des natürlich vorkommenden tat-Proteins umfassen kann. Andere Durchdringungs-Domänen umfassen Poly-Arginin-Motive, beispielsweise die Region der Aminosäuren 34–56 des Rev-Proteins von HIV-1, Nona-Arginin, Octa-Arginin, und ähnliche (siehe beispielsweise Futaki et al. (2003) Curr Protein Pept Sci. 2003 Apr; 4(2): 87–9 und 446; und Wender et al. (2000) Proc. Natl. Acad. Sci. U.S.A 2000 Nov. 21; 97(24): 13003–8; veröffentlichte US-Patentanmeldungen 20030220334; 20030083256; 20030032593 und 20030022831, die vorliegend spezifisch durch Bezugnahme auf die Ausführungen zur Translokation von Peptiden und Peptiden eingeschlossen sind. Die Nona-Arginin(R9)-Sequenz ist eine der effizienteren PTDs, die charakterisiert worden sind (Wender et al. 2000; Uemura et al. 2002). Die Stelle, an der die Fusion durchgeführt wird, kann ausgewählt werden, um die biologische Aktivität, die Sekretions- oder die Bindungseigenschaften des Polypeptids zu optimieren. Die optimale Stelle wird durch Routineversuche ermittelt werden.

[0289] Ein vorliegendes ortsspezifisch modifizierendes Polypeptid kann in vitro erzeugt werden oder durch eukaryontische Zellen oder durch prokaryontische Zellen, und es kann durch Entfaltung, z. B. Hitze-Denaturierung, DTT-Reduktion, usw., weiter bearbeitet werden, und es kann wieder zurückgefaltet werden, wobei im Stand der Technik bekannte Verfahren verwendet werden.

[0290] Modifikationen von Interesse, welche die Primär-Sequenz nicht verändern, umfassen chemische Derivatisierung von Polypeptiden, z. B. Acylierung, Acetylierung, Carboxylierung, Amidierung, usw. Ebenfalls umfasst sind Modifikationen der Glykosylierung, z. B. solche, die durch Modifizieren der Glykosylierungsmuster eines Polypeptids während seiner Synthese und seiner Prozessierung oder in weiteren Prozessierungsschritten erzeugt wurden, z. B. indem man das Polypeptid Enzymen aussetzt, welche die Glykosylierung beeinflusst.

sen, wie beispielsweise glykosylierende oder deglykosylierende Enzyme aus einem Säugetier. Ebenfalls umfasst sind Sequenzen, die phosphorylierte Aminosäure-Reste aufweisen, z. B. Phosphotyrosin, Phosphoserin, oder Phosphothreonin.

[0291] Ebenfalls umfasst von der Erfindung sind auf DNA gerichtete RNAs und ortsspezifisch modifizierende Polypeptide, die unter Verwendung herkömmlicher molekularbiologischer Verfahren und synthetischer Chemie so verändert wurden, dass ihre Resistenz gegen proteolytischen Abbau verbessert wurde, die Spezifität für die Ziel-Sequenz verändert wurde, die Lösungseigenschaften optimiert wurden, die Protein-Aktivität verändert wurde (z. B. Transkriptions-modulierende Aktivität, enzymatische Aktivität, usw.) oder um diese als therapeutisches Mittel geeigneter zu machen. Analoga solcher Polypeptide umfassen solche, die Reste umfassen, bei denen es sich nicht um natürlich vorkommende L-Aminosäuren handelt, z. B. D-Aminosäuren oder nicht-natürlich vorkommende synthetische Aminosäuren. D-Aminosäuren können einige oder alle Aminosäure-Reste ersetzen.

[0292] Die ortsspezifisch modifizierenden Polypeptide können durch in vitro-Synthese unter Verwendung herkömmlicher, im Stand der Technik bekannter Verfahren erzeugt werden. Verschiedene kommerzielle synthetische Vorrichtungen sind verfügbar, z. B. automatisierte Synthese-Geräte von Applied Biosystems, Inc., Beckman, usw. Durch die Verwendung von Synthese-Geräten können natürlich vorkommende Aminosäuren durch unnatürliche Aminosäuren ersetzt werden. Die bestimmte Sequenz und die Art der Herstellung werden nach praktischen Erwägungen, ökonomischen Erwägungen, der erforderlichen Reinheit und ähnlichem bestimmt werden.

[0293] Sofern dies gewünscht ist, können verschiedene Gruppen während der Synthese oder während der Expression in das Peptid eingebracht werden, was die Anbindung an andere Moleküle oder an eine Oberfläche ermöglicht. Daher können Cysteine verwendet werden, um Thioether herzustellen, Histidine für die Verknüpfung an einen Metall-Ionen-Komplex, Carboxylgruppen für die Bildung von Amid oder Ester, Aminogruppen für die Bildung von Amid und ähnliches.

[0294] Die ortsspezifisch modifizierenden Polypeptide können auch durch herkömmliche Verfahren der rekombinanten Synthese isoliert und aufgereinigt werden. Es kann ein Lysat des Expressions-Wirts hergestellt werden, und das Lysat kann durch Verwendung von HPLC, Ausschluss-Chromatographie, Gel-Elektrophorese, Affinitäts-Chromatographie oder anderer Reinigungsverfahren aufgereinigt werden. In den meisten Fällen werden die verwendeten Zusammensetzungen mindestens 20 Gew.-% des gewünschten Produkts, üblicherweise mindestens etwa 75 Gew.-%, vorzugsweise mindestens etwa 95 Gew.-% und für therapeutische Verfahren üblicherweise mindestens etwa 99,5 Gew.-% umfassen, im Hinblick auf Kontaminationen, die mit dem Verfahren der Herstellung des Produkts und seiner Aufreinigung in Verbindung stehen.

[0295] Um die DNA-Spaltung und die Rekombination oder beliebige gewünschte Modifikationen bei einer Ziel-DNA oder beliebige gewünschte Modifikationen bei einem Polypeptid, das mit einer Ziel-DNA assoziiert ist, zu induzieren, wird die auf DNA gerichtete RNA und/oder das ortsspezifisch modifizierende Polypeptid und/oder das Donor-Polynukleotid, unabhängig davon, ob sie als Nukleinsäuren oder Polypeptide eingebracht werden, der Zelle für etwa 30 Minuten bis etwa 24 Stunden bereitgestellt, z. B. 1 Stunde, 1,5 Stunden, 2 Stunden, 2,5 Stunden, 3 Stunden, 3,5 Stunden 4 Stunden, 5 Stunden, 6 Stunden, 7 Stunden, 8 Stunden, 12 Stunden, 16 Stunden, 18 Stunden, 20 Stunden, oder für einen beliebigen anderen Zeitraum von etwa 30 Minuten bis etwa 24 Stunden, wobei dies mit einer Häufigkeit von etwa täglich bis etwa alle vier Tage wiederholt werden kann, z. B. alle 1,5 Tage, alle 2 Tage, alle 3 Tage oder mit einer anderen Häufigkeit von etwa täglich bis etwa alle vier Tage. Das Mittel/die Mittel kann/können den vorliegenden Zellen ein oder mehrfach bereitgestellt werden, z. B. einmal, zweimal, dreimal oder mehr als dreimal, wobei es den Zellen ermöglicht wird, mit dem Mittel/den Mitteln für eine gewisse Zeit nach jedem Inkontaktbringen zu verweilen, z. B. für 16–24 Stunden, wobei das Medium anschließend durch frisches Medium ersetzt wird und die Zellen weiter kultiviert werden.

[0296] In Fällen, bei denen der Zelle zwei oder mehrere verschiedene Targeting-Komplexe bereitgestellt werden (z. B. zwei unterschiedliche auf DNA gerichtete RNAs, die komplementär zu unterschiedlichen Sequenzen innerhalb der gleichen oder in unterschiedlichen Ziel-DNAs sind), wobei die Komplexe gleichzeitig bereitgestellt werden können (z. B. in Form von zwei Polypeptiden und/oder Nukleinsäuren) oder gleichzeitig verabreicht werden können. Alternativ dazu können sie nacheinander bereitgestellt werden (z. B. der Targeting-Komplex wird als erstes bereitgestellt, gefolgt von dem zweiten Targeting-Komplex, usw., oder umgekehrt).

[0297] Üblicherweise wird eine wirksame Menge der auf DNA gerichteten RNA und/oder des ortsspezifisch modifizierenden Polypeptids und/oder des Donor-Polynukleotids an die Ziel-DNA oder an Zellen bereitgestellt,

um die Spaltung zu induzieren. Eine wirksame Menge der auf DNA gerichteten RNA und/oder des ortsspezifisch modifizierenden Polypeptids und/oder des Donor-Polynukleotids ist die Menge, die eine zweifache oder stärkere Erhöhung der Menge der Ziel-Modifikation induziert, welche zwischen zwei homologen Sequenzen relativ zu einer Negativ-Kontrolle, z. B. einer Zelle, die mit einem leeren Vektor oder einem irrelevanten Polypeptid in Kontakt gebracht wird, beobachtet wird. Dies bedeutet, dass eine wirksame Menge oder Dosis der auf DNA gerichteten RNA und/oder des ortsspezifisch modifizierenden Polypeptids und/oder des Donor-Polynukleotids eine 2-fache, eine 3-fache, eine 4-fache, oder stärkere Erhöhung der Menge der Ziel-Modifikation, welche in einer Ziel-DNA-Region beobachtet wird, induzieren wird, in einigen Fällen eine 5-fache Erhöhung, eine 6-fache Erhöhung oder mehr, manchmal eine 7-fache oder 8-fache Erhöhung oder mehr, der Menge der beobachteten Rekombination, z. B. eine Erhöhung um das 10-fache, 50-fache oder 100-fache und mehr, in manchen Fällen eine Erhöhung um das 200-fache, 500-fache, 700-fache oder 1000-fache oder mehr, z. B. ein 5000-fache oder 10000-fache Erhöhung der Menge der beobachteten Rekombination. Die Menge der Ziel-Modifikation kann durch ein beliebiges handhabbares Verfahren gemessen werden.

[0298] Beispielsweise wird ein stilles Reporter-Konstrukt, das eine zu dem Targeting-Segment (Targeting-Sequenz) der auf DNA gerichteten RNA komplementäre Sequenz umfasst, die durch Repeat-Sequenzen flankiert ist, bei Rekombination eine Nukleinsäure hervorbringen, die für einen aktiven Reporter kodiert, und dieses kann in die Zellen co-transfiziert werden, und die Menge des Reporter-Proteins wird nach Inkontaktbringen mit der auf DNA gerichteten RNA und/oder dem ortsspezifisch modifizierenden Polypeptid und/oder dem Donor-Polynukleotid bestimmt, z. B. 2 Stunden, 4 Stunden, 8 Stunden, 12 Stunden, 24 Stunden, 36 Stunden, 48 Stunden, 72 Stunden oder mehr nach Inkontaktbringen mit der auf DNA gerichteten RNA und/oder dem ortsspezifisch modifizierenden Polypeptid und/oder dem Donor-Polynukleotid. Als weiterer, sensitiverer Assay kann beispielsweise das Ausmaß der Rekombination an einer Region einer genomischen DNA von Interesse, welche Ziel-DNA-Sequenzen umfasst, durch PCR oder Southern-Hybridisierung der Region nach Kontakt mit einer auf DNA gerichteten RNA und/oder einem ortsspezifisch modifizierenden Polypeptid und/oder eines Donor-Polynukleotids, z. B. 2 Stunden, 4 Stunden, 8 Stunden, 12 Stunden, 24 Stunden, 36 Stunden, 48 Stunden, 72 Stunden oder mehr nach Inkontaktbringen mit der auf DNA gerichteten RNA und/oder dem ortsspezifisch modifizierenden Polypeptid und/oder dem Donor-Polynukleotid festgestellt werden.

[0299] Das Inkontaktbringen der Zellen mit einer auf DNA gerichteten RNA und/oder einem ortsspezifisch modifizierenden Polypeptid und/oder einem Donor-Polynukleotid kann in einem beliebigen Kulturmedium und unter beliebigen Kulturbedingungen, welche das Überleben der Zellen fördern, durchgeführt werden. Beispielsweise können die Zellen in einem beliebigen geeigneten Nährmedium suspendiert sein, welches praktisch anwendbar ist, wie beispielsweise Iscove-modifiziertes DMEM oder RPMI 1640, das mit fötalem Kälberserum oder Hitze-inaktiviertem Ziegen-Serum (etwa 5–10%), L-Glutamin, einem Thiol, insbesondere 2-Mercaptoethanol, und mit Antibiotika, z. B. Penicillin und Streptomycin, supplementiert ist. Die Kultur kann Wachstumsfaktoren umfassen, auf die die Zellen reagieren. Wachstumsfaktoren sind wie vorliegend definiert Moleküle, die in der Lage sind, das Überleben, das Wachstum und/oder die Differenzierung von Zellen, entweder in Kultur oder im intakten Gewebe, durch spezifische Effekte an einem Transmembranen-Rezeptor zu fördern. Wachstumsfaktoren umfassen Polypeptide und Faktoren, bei denen es sich nicht um Polypeptide handelt. Bedingungen, die das Überleben von Zellen fördern, erlauben üblicherweise nicht-homologes End Joining und Homologie-gesteuerte Reparatur.

[0300] Bei Anwendungen, bei denen es wünschenswert ist, eine Polynukleotidsequenz in eine Ziel-DNA-Sequenz zu insertieren, wird ferner der Zelle ein Polynukleotid bereitgestellt, welches eine Donor-Sequenz umfasst, die insertiert werden soll. Mit einer "Donor-Sequenz" oder einem "Donor-Polynukleotid" ist eine Nukleinsäuresequenz gemeint, die an einer Spaltungsstelle, welche durch ein ortsspezifisch modifizierendes Polypeptid induziert wurde, eingebracht werden soll. Das Donor-Polynukleotid wird ausreichend Homologie zu einer genomischen Sequenz an der Spaltungsstelle aufweisen, z. B. 70%, 80%, 85%, 90%, 95% oder 100% Homologie zu den Nukleotidsequenzen, welche die Spaltungsstelle flankieren, z. B. innerhalb von etwa 50 Basen oder weniger von der Spaltungsstelle, z. B. innerhalb von etwa 30 Basen, innerhalb von etwa 15 Basen, innerhalb von etwa 10 Basen, innerhalb von etwa 5 Basen, oder unmittelbar die Spaltungsstelle flankierend, um die Homologie-gesteuerte Reparatur zwischen diesem und der genomischen Sequenz, zu dem es Homologie zeigt, zu fördern. Ungefähr 25, 50, 100 oder 200 Nukleotide oder mehr als 200 Nukleotide der Sequenz-Homologie zwischen einer Donor-Sequenz und einer genomischen Sequenz (oder einem beliebigen ganzzahligen Wert zwischen 10 und 200 Nukleotiden oder mehr) werden die Homologie-gesteuerte Reparatur unterstützen. Donor-Sequenzen können eine beliebige Länge aufweisen, z. B. 10 Nukleotide oder mehr, 50 Nukleotide oder mehr, 100 Nukleotide oder mehr, 250 Nukleotide oder mehr, 500 Nukleotide oder mehr, 1000 Nukleotide oder mehr, 5000 Nukleotide oder mehr, usw.

[0301] Die Donor-Sequenz ist üblicherweise nicht identisch zu der genomischen Sequenz, die sie ersetzt. Vielmehr kann die Donor-Sequenz mindestens einen oder mehrere Einzelbasen-Austausche, Insertionen, Deletionen, Inversionen oder Umlagerungen in Bezug auf die genomische Sequenz enthalten, solange ausreichend Homologie vorhanden ist, um die Homologie-gesteuerte Reparatur zu unterstützen. In einigen Ausführungsformen umfasst die Donor-Sequenz eine nicht-homologe Sequenz, die von zwei homologen Regionen flankiert wird, so dass eine Homologie-gesteuerte Reparatur zwischen der Ziel-DNA-Region und den zwei flankierenden Sequenzen zur Insertion einer nicht-homologen Sequenz an der Ziel-Region führt. Donor-Sequenzen können ferner ein Vektor-Rückgrat umfassen, welches Sequenzen enthält, die nicht homolog zur DNA-Region von Interesse sind, und die nicht für die Insertion in die DNA-Region von Interesse gedacht sind. Üblicherweise wird (werden) die homologe(n) Region(en) einer Donor-Sequenz mindestens 50% Sequenz-Identität zu einer genomischen Sequenz, mit der eine Rekombination gewünscht ist, aufweisen. In einigen Ausführungsformen ist eine Sequenz-Identität von 60%, 70%, 80%, 90%, 95%, 98%, 99% oder 99,9% vorhanden. Jeder Wert zwischen 1% und 100% Sequenz-Identität kann vorhanden sein, abhängig von der Länge des Donor-Polynukleotids.

[0302] Die Donor-Sequenz kann bestimmte Sequenz-Unterschiede im Vergleich zu der genomischen Sequenz umfassen, z. B. Restriktionsstellen, Nukleotid-Polymorphismen, selektierbare Marker (z. B. Wirkstoff-Resistenz-Gene, fluoreszente Proteine, Enzyme, usw.), usw., die verwendet werden können, um eine erfolgreiche Insertion der Donor-Sequenz an der Spaltstelle festzustellen, oder sie können in einigen Fällen für andere Zwecke verwendet werden (z. B. um die Expression am beabsichtigten genomischen Locus anzuzeigen). In einigen Fällen werden solche Nukleotidsequenzen, wenn sie in einer kodierenden Region lokalisiert sind, die Aminosäuresequenz nicht ändern, oder sie werden für stille Aminosäure-Änderungen sorgen (d. h. Änderungen, welche die Struktur oder die Funktion des Proteins nicht verändern). Alternativ dazu können diese Sequenz-Unterschiede flankierende Rekombinations-Sequenzen, wie beispielsweise FLPs, 1oxP-Sequenzen oder ähnliches umfassen, die zu einem späteren Zeitpunkt zur Entfernung der Marker-Sequenz aktiviert werden können.

[0303] Die Donor-Sequenz kann der Zelle als einzelsträngige DNA, einzelsträngige RNA, doppelsträngige DNA oder doppelsträngige RNA bereitgestellt werden. Sie kann in linearer oder zirkulärer Form in eine Zelle eingebracht werden. Sofern sie in linearer Form eingebracht wird, können die Enden der Donor-Sequenz geschützt werden (z. B. vor exonukleolytischem Abbau), wobei Verfahren angewendet werden, die dem Fachmann bekannt sind. Beispielsweise können ein oder mehrere Didesoxynukleotid-Reste an den 3'-Terminus eines linearen Moleküls angefügt werden, und/oder es können selbst-komplementäre Oligonukleotide an eines der Enden oder an beide Enden ligiert werden. Siehe beispielsweise Chang et al. (1987) Proc. Natl. Acad. Sci. USA 84: 4959–4963; Nehls et al. (1996) Science 272: 886–889. Zusätzliche Verfahren zum Schutz von exogenen Polynukleotiden vor Abbau umfassen (sind jedoch nicht beschränkt auf) die Zufügung von einer terminalen Aminogruppe oder von terminalen Aminogruppen sowie die Verwendung von modifizierten Internukleotid-Verknüpfungen, wie beispielsweise Phosphorthioaten, Phosphoramidaten und O-Methyl-Ribose oder Desoxyribose-Resten. Als eine Alternative zum Schutz der Termini einer linearen Donor-Sequenz können zusätzliche Sequenz-Abschnitte außerhalb der homologen Regionen eingebracht werden, welche abgebaut werden können, ohne die Rekombination zu beeinflussen. Eine Donor-Sequenz kann als Teil eines Vektor-Moleküls, das zusätzliche Sequenzen umfasst, z. B. Replikationsursprünge, Promotoren und Gene, die für antibiotische Resistenz kodieren, in eine Zelle eingebracht werden. Darüber hinaus können Donor-Sequenzen als nackte Nukleinsäure eingebracht werden, als Nukleinsäure, die mit einem Mittel, wie beispielsweise einem Liposom oder einem Poloxamer komplexiert ist, oder sie können durch Viren bereitgestellt werden (z. B. Adenovirus, AAV), wie es oben für Nukleinsäuren, die für eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid kodieren, beschrieben wurde.

[0304] Nach den oben beschriebenen Verfahren kann eine DNA-Region von Interesse gespalten und modifiziert werden, d. h. "genetisch modifiziert", ex vivo. In einigen Ausführungsformen, wenn z. B. ein selektierbarer Marker in die DNA-Region von Interesse inseriert wurde, kann die Population von Zellen im Hinblick auf solche Zellen angereichert werden, die die genetische Modifikation umfassen, indem die genetisch modifizierten Zellen von der verbleibenden Population abgetrennt werden. Vor der Anreicherung können die "genetisch modifizierten" Zellen lediglich etwa 1% oder mehr (z. B. 2% oder mehr, 3% oder mehr, 4% oder mehr, 5% oder mehr, 6% oder mehr, 7% oder mehr, 8% oder mehr, 9% oder mehr, 10% oder mehr, 15% oder mehr, oder 20% oder mehr) der zellulären Population ausmachen. Die Trennung von "genetisch modifizierten" Zellen kann durch beliebige handhabbare Trennverfahren erreicht werden, die im Hinblick auf den verwendeten selektierbaren Marker geeignet sind. Wenn beispielsweise ein Fluoreszenz-Marker inseriert wurde, können die Zellen durch Fluoreszenz-aktivierte Zell-Sortierung abgetrennt werden, und wenn ein Zelloberflächenmarker inseriert wurde, können die Zellen aus der heterogenen Population durch Affinitäts-Separationsverfahren

abgetrennt werden, z. B. durch magnetische Abtrennung, Affinitäts-Chromatographie, "Panning", mit einem Affinitäts-Reagenz, das an eine feste Matrix gebunden ist, oder durch ein anderes handhabbares Verfahren. Verfahren, die eine akkurate Abtrennung bereitstellen, umfassen Vorrichtungen für die Fluoreszenz-aktivierte Zell-Sortierung, die unterschiedliche Entwicklungsstufen aufweisen können, wie beispielsweise Mehrfarbenkanäle, Detektionskanäle für Kleinwinkel- und Stumpfwinkel-Lichtstreuung, Impedanz-Kanäle, usw. Die Zellen können gegen tote Zellen selektiert werden, indem man Farbstoffe verwendet, die mit toten Zellen assoziiert sind (z. B. Propidiumiodid). Es kann ein beliebiges Verfahren verwendet werden, das für die Lebensfähigkeit der genetisch modifizierten Zellen nicht übermäßig schädlich ist. Auf diese Weise werden Zell-Zusammensetzungen erhalten, die in Bezug auf Zellen, welche die modifizierte DNA enthalten, stark angereichert sind. Mit "stark angereichert" ist gemeint, dass die genetisch modifizierten Zellen 70% oder mehr, 75% oder mehr, 80% oder mehr, 85% oder mehr, 90% oder mehr der Zell-Zusammensetzungen ausmachen, z. B. etwa 95% oder mehr, oder 98% oder mehr der Zell-Zusammensetzung. Mit anderen Worten kann die Zusammensetzung eine im Wesentlichen reine Zusammensetzung von genetisch modifizierten Zellen sein.

[0305] Genetisch modifizierte Zellen, die anhand der vorliegend beschriebenen Verfahren erzeugt wurden, können sofort verwendet werden. Alternativ dazu können die Zellen bei Temperaturen von flüssigem Stickstoff eingefroren und für längere Zeiträume gelagert werden, wobei diese nach Auftauen in der Lage sind, erneut verwendet zu werden. In solchen Fällen werden die Zellen üblicherweise in 10% Dimethylsulfoxid (DMSO), 50% Serum, 40% gepuffertes Medium eingefroren, oder in einer anderen derartigen Lösung, die normalerweise im Stand der Technik verwendet wird, um Zellen bei derartigen Gefriertemperaturen zu konservieren, und sie können in einer Weise aufgetaut werden, die üblicherweise im Stand der Technik für das Auftauen von gefroren kultivierten Zellen bekannt ist.

[0306] Die genetisch modifizierten Zellen können in vitro unter verschiedenen Bedingungen kultiviert werden. Die Zellen können in Kultur expandiert werden, d. h. unter Bedingungen angezüchtet werden, die ihre Proliferation unterstützen. Das Kulturmedium kann flüssig oder semi-fest sein, es kann z. B. Agar, Methylzellulose, usw., enthalten. Die Zellpopulation kann in einem geeigneten Nährmedium suspendiert werden, wie beispielsweise in Iscove-modifiziertem DMEM oder in RPMI 1640, das normalerweise mit fötalem Kälberserum (etwa 5–10%), L-Glutamin, einem Thiol, insbesondere 2-Mercaptoethanol, und mit Antibiotika, z. B. Penicillin und Streptomycin, supplementiert ist. Die Kultur kann Wachstumsfaktoren enthalten, auf die die regulatorischen T-Zellen reagieren. Wachstumsfaktoren sind wie vorliegend definiert Moleküle, die in der Lage sind, das Überleben, das Wachstum und/oder die Differenzierung von Zellen entweder in Kultur oder im intakten Gewebe durch spezifische Wirkungen auf einen Transmembran-Rezeptor zu fördern. Wachstumsfaktoren umfassen Polypeptide und Faktoren, bei denen es sich nicht um Polypeptide handelt.

[0307] Zellen, die in dieser Weise genetisch modifiziert wurden, können für Zwecke wie zum Beispiel für die Gentherapie in ein Subjekt transplantiert werden, z. B. um eine Krankheit zu behandeln oder als antivirales, antipathogenes oder ein Anti-Krebs-Therapeutikum, für die Herstellung von genetisch modifizierten Organismen in der Landwirtschaft oder für die biologische Forschung. Das Subjekt kann ein Neugeborenes, ein Jungendlicher oder ein Erwachsener sein. Von besonderem Interesse sind Säugetier-Subjekte. Säugetier-Spezies, die mit dem vorliegenden Verfahren behandelt werden können, umfassen Hunde und Katzen; Pferde; Rinder; Schafe; usw. und Primaten, insbesondere Menschen. Tiermodelle, insbesondere kleine Säugetiere (z. B. Maus, Ratte, Meerschweinchen, Hamster, Hasenartige (z. B. Kaninchen), usw.) können für experimentelle Untersuchungen verwendet werden.

[0308] Die Zellen können dem Subjekt allein oder zusammen mit einem geeigneten Substrat oder einer Matrix bereitgestellt werden, z. B. um deren Wachstum und/oder deren Organisation in dem Gewebe, in das sie transplantiert werden, zu unterstützen. Normalerweise werden mindestens 10^3 Zellen verabreicht werden, beispielsweise 5×10^3 Zellen, 1×10^4 Zellen, 5×10^4 Zellen, 1×10^5 Zellen, 1×10^6 Zellen oder mehr. Die Zellen können auf einem beliebigen der nachfolgenden Wege an das Subjekt verabreicht werden: parenteral, subkutan, intravenös, intrakranial, intraspinal, intraokular oder in die Spinalflüssigkeit. Die Zellen können durch Injektion, Katheter oder ähnliches eingebracht werden. Beispiele für Verfahren für die lokale Bereitstellung, d. h., die Bereitstellung an der Stelle der Verletzung, umfassen z. B. durch ein Ommaya-Reservoir, z. B. für die intrathekale Bereitstellung (siehe z. B. US-Patent-Nrn. 5,222,982 und 5,385,582, die vorliegend durch Bezugnahme eingeschlossen sind); durch Bonus-Injektion, z. B. durch eine Spritze, z. B. in ein Gelenk; durch kontinuierliche Infusion, z. B. durch Kanülierung, z. B. mit Konvektion (siehe z. B. US-Patentanmeldung Nr. 20070254842, die vorliegend durch Bezugnahme eingeschlossen ist; oder durch Implantieren einer Vorrichtung, auf der die Zellen reversibel befestigt wurden (siehe z. B. US-Patentanmeldung-Nm. 20080081064 und 20090196903, die vorliegend durch Bezugnahme eingeschlossen sind). Die Zellen können auch in einen Embryo (z. B. in eine Blastozyste) eingebracht werden, um ein transgenes Tier zu erzeugen (z. B. eine transgene Maus).

[0309] Die Anzahl der Verabreichungen an das Subjekt bei der Behandlung kann variieren. Das Einbringen der genetisch modifizierten Zellen in das Subjekt kann ein einmaliges Ereignis sein; in manchen Situationen jedoch kann eine solche Behandlung eine Verbesserung für einen beschränkten Zeitraum zeigen und eine fortlaufende Anzahl von wiederholten Behandlungen erfordern. In anderen Situationen können mehrere Verabreichungen der genetisch modifizierten Zellen erforderlich sein, bevor eine Wirkung beobachtet wird. Die exakten Protokolle hängen von der Erkrankung oder dem Zustand ab, von der Phase der Erkrankung und von den Eigenschaften des einzelnen Subjekts, das behandelt wird.

[0310] In anderen Aspekten der Erfindung werden die auf DNA gerichtete RNA und/oder das ortsspezifisch modifizierende Polypeptid und/oder das Donor-Polynukleotid dazu verwendet, zelluläre DNA in vivo zu modifizieren, wiederum für Zwecke wie zum Beispiel für die Gentherapie, z. B. um eine Erkrankung zu behandeln oder als antivirales, antipathogenes oder Anti-Krebs-Therapeutikum, für die Erzeugung genetisch modifizierter Organismen in der Landwirtschaft oder für die biologische Forschung. In diesen in vivo-Anwendungen werden eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid direkt an das Individuum verabreicht. Eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid können durch eine Vielzahl von hinreichend bekannten Verfahren für die Verabreichung von Peptiden, Small Molecules und Nukleinsäuren an ein Subjekt verabreicht werden. Eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid können in eine Vielzahl von Formulierungen eingebracht werden. Insbesondere können eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid der vorliegenden Erfindung in pharmazeutische Zusammensetzungen eingebracht werden, indem sie mit geeigneten pharmazeutisch akzeptablen Trägern oder Verdünnungsmitteln kombiniert werden.

[0311] Pharmazeutische Zubereitungen sind Zusammensetzungen, die eine oder mehrere auf DNA gerichtete RNAs und/oder ortsspezifisch modifizierende Polypeptide und/oder Donor-Polynukleotide umfassen, die in einem pharmazeutisch akzeptablen Vehikel vorliegen. "Pharmazeutisch akzeptable Vehikel" können Vehikel sein, die durch eine regulatorische Behörde der Bundesregierung oder einer Landesregierung zugelassen wurden, oder die in der US-Pharmakopöe oder in einer anderen allgemein anerkannten Pharmakopöe zur Verwendung in Säugetieren, wie beispielsweise in Menschen, aufgeführt sind. Der Begriff "Vehikel" bezeichnet ein Verdünnungsmittel, ein Adjuvans, einen Hilfsstoff oder einen Träger, mit dem eine Verbindung der Erfindung zur Verabreichung an ein Säugetier formuliert ist. Solche pharmazeutischen Vehikel können Lipide, z. B. Liposomen, z. B. Liposomen-Dendrimere; Flüssigkeiten, wie beispielsweise Wasser und Öle, einschließlich solcher, die von Erdöl, Tieren, Gemüse stammen oder synthetischen Ursprungs sind, wie beispielsweise Erdnussöl, Sojabohnenöl, Mineralöl, Sesamöl oder ähnliches; Gummiarabikum, Gelatine, Stärkepaste, Talk, Keratin, kolloidale Kieselsäure, Harnstoff und ähnliches umfassen. Darüber hinaus können Hilfsstoffe, Stabilisierungsmittel, Verdickungsmittel, Befeuchtungsmittel und Farbstoffe verwendet werden. Pharmazeutische Zusammensetzungen können zu Zubereitungen in fester, halbfester, flüssiger oder gasförmiger Form formuliert werden, wie beispielsweise zu Tabletten, Kapseln, Pulvern, Granulaten, Salben, Lösungen, Zäpfchen, Injektionen, Inhalationsmitteln, Gelen, Mikrosphären und Aerosolen. Somit kann die Verabreichung einer auf DNA gerichteten RNA und/oder eines ortsspezifisch modifizierenden Polypeptids und/oder eines Donor-Polynukleotids auf verschiedene Arten erreicht werden, einschließlich oraler, bukkaler, rektaler, parenteraler, intraperitonealer, intradermaler, transdermaler, intratrachealer, intraokularer, usw. Verabreichung. Der Wirkstoff kann nach Verabreichung systemisch vorliegen oder er kann durch Verwendung von regionaler Verabreichung, intramuraler Verabreichung, oder durch Verwendung eines Implantats, das die aktive Dosis an der Stelle der Implantation zurückhält, lokalisiert vorliegen. Der Wirkstoff kann für die sofortige Aktivität formuliert sein, oder er kann für die anhaltende Freisetzung formuliert sein.

[0312] Bei einigen Zuständen, insbesondere Zuständen des zentralen Nervensystems, kann es erforderlich sein, Mittel so zu formulieren, dass sie die Blut-Hirn-Schranke (blood-brain-barrier, BBB) überqueren. Eine Strategie für die Verabreichung des Wirkstoffs durch die Blut-Hirn-Schranke (BBB) umfasst die Zerstörung der BBB, entweder durch osmotische Mittel, wie beispielsweise Mannitol oder Leukotriene, oder auf biochemischem Weg durch Verwendung von vasoaktiven Substanzen, wie beispielsweise Bradykinin. Die Möglichkeit zur Verwendung einer BBB-Öffnung zur gezielten Einschleusung spezifischer Wirkstoffe in Hirntumore ist ebenfalls eine Option. Ein BBB-zerstörendes Mittel kann gemeinsam mit den therapeutischen Zusammensetzungen der Erfindung co-verabreicht werden, wenn die Zusammensetzungen durch intravaskuläre Injektion verabreicht werden. Andere Strategien zur Durchquerung der BBB können die Verwendung von endogenen Transportsystemen umfassen, einschließlich Caveolin-1-vermittelte Transzytose, Träger-vermittelte Transporter, wie beispielsweise Glukose- und Aminosäure-Transporter, Rezeptor-vermittelte Transzytose für Insulin oder Transferrin und aktive Efflux-Transporter, wie beispielsweise p-Glykoprotein. Aktive Transport-Reste können ferner an die therapeutischen Zusammensetzungen zur Verwendung im Rahmen der vorliegenden Erfin-

dung konjugiert werden, um den Transport über die Endothelialwand des Blutgefäßes zu gewährleisten. Alternativ dazu kann die Wirkstoff-Bereitstellung von therapeutischen Mitteln hinter die BBB durch lokale Bereitstellung erfolgen, beispielsweise durch intrathekale Bereitstellung, z. B. durch ein Ommaya-Reservoir (siehe beispielsweise US-Patent Nrn. 5,222,982 und 5385582, die vorliegend durch Bezugnahme eingeschlossen sind); durch Bolus-Injektion, z. B. durch eine Spritze, z. B. intravitreal oder intrakranial; durch kontinuierliche Infusion, z. B. durch Kanülierung, z. B. mit Konvektion (siehe z. B. US-Patentanmeldung Nr. 20070254842, die vorliegend durch Bezugnahme eingeschlossen ist), oder durch Implantation einer Vorrichtung, auf der das Mittel reversibel befestigt wurde (siehe beispielsweise US-Patent Nrn. 20080081064 und 20090196903, die vorliegend durch Bezugnahme eingeschlossen sind).

[0313] Üblicherweise wird eine wirksame Menge einer auf DNA gerichteten RNA und/oder eines ortsspezifisch modifizierenden Polypeptids und/oder eines Donor-Polynukleotids bereitgestellt. Wie oben im Zusammenhang mit ex vivo-Verfahren erörtert ist eine wirksame Menge oder eine wirksame Dosis einer auf DNA gerichteten RNA und/oder eines ortsspezifisch modifizierenden Polypeptids und/oder eines Donor-Polynukleotids in vivo die Menge, die eine zweifache Erhöhung oder eine stärkere Erhöhung der Menge an Rekombination zwischen zwei homologen Sequenzen relativ zu einer Negativ-Kontrolle induziert, z. B. einer Zelle, die mit einem leeren Vektor oder einem irrelevanten Polypeptid in Kontakt gebracht wird. Die Menge an Rekombination kann durch ein beliebiges handhabbares Verfahren bestimmt werden, z. B. wie es oben beschrieben und im Stand der Technik bekannt ist. Die Berechnung der wirksamen Menge oder der wirksamen Dosis einer zu verabreichenden auf DNA gerichteten RNA und/oder eines ortsspezifisch modifizierenden Polypeptids und/oder eines Donor-Polynukleotids liegt im Bereich des Fachwissen des Fachmanns und wird durch den Fachmann routinemäßig bestimmt werden. Die endgültig zu verabreichende Menge wird von der Art der Verabreichung und von der Art der zu behandelnden Erkrankung oder des zu behandelnden Zustands abhängen.

[0314] Die wirksame Menge, die an einen bestimmten Patienten verabreicht wird, wird von einer Vielzahl von Faktoren abhängen, von denen sich einige von Patient zu Patient unterscheiden werden. Ein fähiger Arzt wird in der Lage sein, eine wirksame Menge eines therapeutischen Mittels zur Verabreichung an einen Patienten zu bestimmen, um das Fortschreiten des Erkrankungszustands nach Bedarf anzuhalten oder umzukehren. Durch Verwendung von LD50-Daten aus Tieren und anderen Informationen, die für das Mittel erhältlich sind, kann ein Arzt die maximale sichere Dosis für ein Individuum in Abhängigkeit der Art der Verabreichung bestimmen. Beispielsweise kann eine intravenös verabreichte Dosis höher sein als eine intrathekal verabreichte Dosis, weil die Menge der Flüssigkeit, in der die therapeutische Zusammensetzung verabreicht wird, größer ist. In ähnlicher Weise können Zusammensetzungen, die schnell vom Körper ausgeschieden werden, in höheren Dosen verabreicht werden, oder in sich wiederholenden Dosen, um eine therapeutische Konzentration beizubehalten. Durch Verwendung des üblichen Fachwissens wird der fähige Arzt in der Lage sein, die Dosierung eines bestimmten Therapeutikums im Verlaufe routinemäßiger klinischer Versuche zu optimieren.

[0315] Zur Einbringung in ein Medikament kann eine auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid und/oder ein Donor-Polynukleotid von einer geeigneten kommerziellen Bezugsquelle erhalten werden. Allgemein wird vorgeschlagen, dass die gesamte pharmazeutisch wirksame Menge einer auf DNA gerichteten RNA und/oder eines ortsspezifisch modifizierenden Polypeptids und/oder eines Donor-Polynukleotids, die parenteral verabreicht wird, pro Dosis in einem Bereich liegt, der durch eine Dosis-Response-Kurve gemessen werden kann.

[0316] Therapien, die auf einer auf DNA gerichteten RNA und/oder einem ortsspezifisch modifizierenden Polypeptid und/oder Donor-Polynukleotiden basieren, d. h. Zubereitungen einer auf DNA gerichteten RNA und/oder eines ortsspezifisch modifizierenden Polypeptids und/oder eines Donor-Polynukleotids, die für die therapeutische Verabreichung verwendet werden, müssen steril sein. Sterilität wird auf einfache Weise mittels Filtration durch eine sterile Filtrationsmembran (z. B. 0,2 µm Membran) erreicht. Therapeutische Zusammensetzungen werden im Allgemeinen in einen Behälter eingebracht, der einen sterilen Zugangsport hat, z. B. einen Beutel mit intravenöser Lösung oder ein Fläschchen mit einem Stopfen, der mit einer hypodermischen Injektionsnadel durchstechbar ist. Die Therapien, die auf einer auf DNA gerichteten RNA und/oder einem ortsspezifisch modifizierenden Polypeptid und/oder einem Donor-Polynukleotid basieren, können in Einheitsbehältern oder in Mehr-Dosen-Behältern gelagert werden, beispielsweise in versiegelten Ampullen oder Fläschchen, als wässrige Lösung oder als lyophilisierte Formulierung zur Rekonstituierung. Als ein Beispiel für eine lyophilisierte Formulierung werden 10 ml Fläschchen mit 5 ml steril filtrierter 1%iger (w/v) wässriger Lösung der Verbindung befüllt, und die resultierende Mischung wird lyophilisiert. Die Infusionslösung wird hergestellt, indem man die lyophilisierte Verbindung unter Verwendung von bakteriostatischem Wasser für die Injektion rekonstituiert.

[0317] Pharmazeutische Zusammensetzungen können, abhängig von der gewünschten Formulierung, pharmazeutisch akzeptable, nicht toxische Träger oder Verdünnungsmittel umfassen, die als Vehikel definiert sind, welche üblicherweise verwendet werden, um pharmazeutische Zusammensetzungen für die Verabreichung an Tiere oder Menschen zu formulieren. Das Verdünnungsmittel wird so ausgewählt, das es die biologische Aktivität der Kombination nicht beeinflusst. Beispiele für solche Verdünnungsmittel sind destilliertes Wasser, gepuffertes Wasser, physiologische Salzlösung, PBS, Ringer-Lösung, Dextrose-Lösung und Hanks-Lösung. Darüber hinaus kann die pharmazeutische Zusammensetzung oder Formulierung weitere Träger, Adjuvantien oder nicht-toxische, nicht-therapeutische, nicht-immunogene Stabilisierungsmittel, Hilfsmittel und ähnliches enthalten. Die Zusammensetzungen können ferner zusätzliche Substanzen enthalten, um sich physiologischen Bedingungen anzunähern, wie beispielsweise pH-einstellende und puffernde Mittel, Toxizitäts-einstellende Mittel, Befeuchtungsmittel und Detergenzien.

[0318] Die Zusammensetzung kann ferner eines von mehreren Stabilisierungsmitteln enthalten, wie beispielsweise ein antioxidatives Mittel. Wenn die pharmazeutische Zusammensetzung ein Polypeptid enthält, kann das Polypeptid mit verschiedenen hinreichend bekannten Verbindungen komplexiert werden, um die in vivo-Stabilität des Polypeptids zu erhöhen oder auf andere Weise seine pharmakologischen Eigenschaften zu verbessern (z. B. Erhöhung der Halbwertszeit des Polypeptids, Verringerung seiner Toxizität, Erhöhung seiner Löslichkeit oder seiner Aufnahme). Beispiele für solche Modifikationen oder komplexierende Mittel umfassen Sulfat, Gluconat, Citrat und Phosphat. Die Nukleinsäuren oder Polypeptide einer Zusammensetzung können ferner mit Molekülen komplexiert sein, die ihre in vivo-Eigenschaften verbessern. Solche Moleküle umfassen beispielsweise Kohlenhydrate, Polyamine, Aminosäuren, andere Peptide, Ionen (z. B. Natrium, Kalium, Kalzium, Magnesium, Mangan) und Lipide.

[0319] Weitere Hinweise zu Formulierungen, die für verschiedene Verabreichungsarten geeignet sind, können in Remington's Pharmaceutical Sciences, Mace Publishing Company, Philadelphia, Pa., 17. Auflage (1985) gefunden werden. Für einen kurzen Überblick über Verfahren zur Bereitstellung von Wirkstoffen, siehe Langer, Science 249: 1527–1533 (1990).

[0320] Die pharmazeutischen Zusammensetzungen können für prophylaktische und/oder therapeutische Behandlungen verabreicht werden. Toxizität und therapeutische Wirksamkeit des Wirkstoffs können nach pharmazeutischen Standardverfahren in Zellkulturen und/oder in Testtieren bestimmt werden, beispielsweise durch Bestimmung der LD50 (die Dosis, die für 50% der Population letal ist) und die ED50 (die Dosis, die bei 50% der Population therapeutisch wirksam ist). Das Dosis-Verhältnis zwischen toxischer Wirkung und therapeutischer Wirkung ist der therapeutische Index, und dieser kann als Verhältnis LD50/ED50 ausgedrückt werden. Therapien, die große therapeutische Indizes aufweisen, sind bevorzugt.

[0321] Die Daten, die aus der Zellkultur und/oder aus Untersuchungen mit Tieren erhalten werden, können verwendet werden, um einen Dosierungsbereich für den Menschen zu formulieren. Die Dosierung des Wirkstoffs liegt üblicherweise innerhalb eines Bereichs von zirkulierenden Konzentrationen, die den ED50 mit geringer Toxizität einschließen. Die Dosierung kann innerhalb dieses Bereichs in Abhängigkeit der verwendeten Dosierungsform und der verwendeten Verabreichungsart variieren.

[0322] Die Bestandteile, die verwendet werden, um die pharmazeutischen Zusammensetzungen zu formulieren, sind vorzugsweise von hoher Reinheit und im Wesentlichen frei von potentiell schädlichen Kontaminationen (z. B. mindestens National Food(NF)-Qualität, im Allgemeinen mindestens analytische Qualität und typischerweise mindestens pharmazeutische Qualität). Darüber hinaus sind Zusammensetzungen, die für die Verwendung in vivo gedacht sind, üblicherweise steril. Wenn eine bestimmte Verbindung vor ihrer Verwendung synthetisiert werden muss, ist das resultierende Produkt üblicherweise im Wesentlichen frei von jeglichem potentiell toxischem Mittel, insbesondere von jeglichen Endotoxinen, die während der Synthese oder des Aufreinigungsverfahrens vorliegen können. Zusammensetzungen für die parenterale Verabreichung sind ebenfalls steril, im Wesentlichen isotonisch und unter GMP-Bedingungen hergestellt.

[0323] Die wirksame Menge einer therapeutischen Zusammensetzung, die an einen bestimmten Patienten verabreicht werden soll, wird von einer Vielzahl von Faktoren abhängen, von denen sich einige von Patient zu Patient unterscheiden werden. Ein fähiger Arzt wird in der Lage sein, eine wirksame Menge eines therapeutischen Mittels für die Verabreichung an einen Patienten zu bestimmen, um das Fortschreiten des Erkrankungszustands nach Bedarf anzuhalten oder umzukehren. Durch Verwendung von LD50-Daten aus Tieren und anderen Informationen, die für das Mittel erhältlich sind, kann ein Arzt die maximale sichere Dosis für ein Individuum in Abhängigkeit von der Art der Verabreichung bestimmen. Beispielsweise kann eine intravenös verabreichte Dosis höher sein als eine intrathekal verabreichte Dosis, weil die Menge der Flüssigkeit, in der die

therapeutische Zusammensetzung verabreicht wird, größer ist. In ähnlicher Weise können Zusammensetzungen, die schnell vom Körper ausgeschieden werden, in höheren Dosen verabreicht werden, oder in sich wiederholenden Dosen, um eine therapeutische Konzentration beizubehalten. Durch Verwendung des üblichen Fachwissens wird der fähige Arzt in der Lage sein, die Dosierung eines bestimmten Therapeutikums im Verlaufe routinemäßiger klinischer Versuche zu optimieren.

GENETISCH MODIFIZIERTE WIRTSZELLEN

[0324] Die vorliegende Offenbarung stellt genetisch modifizierte Wirtszellen bereit, einschließlich isolierter genetisch modifizierter Wirtszellen, wobei eine vorliegende genetisch modifizierte Wirtszelle Folgendes umfasst (genetisch modifiziert wurde mit): 1.) eine exogene auf DNA gerichtete RNA; 2.) eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; 3.) ein exogenes ortsspezifisch modifizierendes Polypeptid (z. B. ein natürlich vorkommendes Cas9; eine modifizierte, d. h. mutierte Variante von Cas9; ein chimäres Cas9; usw.); 4.) eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert; oder 5.) eine beliebige Kombination der obigen. Eine vorliegende genetisch modifizierte Zelle wird erzeugt, indem man eine Wirtszelle genetisch modifiziert, beispielsweise mit: 1) einer exogenen auf DNA gerichteten RNA; 2.) einer exogenen Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; 3.) einem exogenen ortsspezifisch modifizierenden Polypeptid; 4.) einer exogenen Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert; oder 5.) eine beliebige Kombination der obigen).

[0325] Alle Zellen, die als Ziel-Zelle geeignet sind, sind auch als genetisch modifizierte Wirtszelle geeignet. Beispielsweise kann eine genetisch modifizierte Wirtszelle von Interesse eine Zelle aus einem beliebigen Organismus sein (z. B. eine bakterielle Zelle, eine Archaeen-Zelle, eine Zelle eines einzelligen eukaryontischen Organismus, eine Pflanzenzelle, eine Algenzelle, z. B. *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens* C. Agardh, und ähnliche, eine Pilz-Zelle (z. B. eine Hefezelle), eine tierische Zelle, eine Invertebraten-Zelle (z. B. Fruchtfliege, Nesseltier, Stachelhäuter, Fadenwurm, usw.), eine Zelle von einem Wirbeltier (z. B. Fisch, Amphibie, Reptil, Vogel, Säugetier), eine Zelle von einem Säugetier (z. B. ein Schwein, eine Kuh, eine Ziege, ein Schaf, ein Nagetier, eine Ratte, eine Maus, ein nicht-humaner Primat, ein Mensch, usw.), usw.

[0326] In einigen Ausführungsformen wurde eine genetisch modifizierte Wirtszelle genetisch mit einer exogenen Nukleinsäure modifiziert, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, z. B. mutiertes Cas9 oder eine Variante davon; ein chimäres Cas9, usw.). Die DNA der genetisch modifizierten Zelle kann zum Ziel der Modifikation gemacht werden, indem man eine auf DNA gerichtete RNA (oder eine DNA, die für eine auf DNA gerichtete RNA kodiert, welche die genomische Stelle/Sequenz, die modifiziert werden soll, bestimmt) und wahlweise eine Donor-Nukleinsäure in die Zelle einbringt. In einigen Ausführungsformen ist die Nukleotidsequenz, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise an einen induzierbaren Promotor gekoppelt (z. B. Hitzeschock-Promotor, Tetracyclin-regulierter Promotor, Steroid-regulierter Promotor, Metall-regulierter Promotor, Östrogenrezeptor-regulierter Promotor, usw.). In einigen Ausführungsformen ist die Nukleotidsequenz, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise an einen räumlich beschränkten und/oder zeitlich beschränkten Promotor gekoppelt (z. B. einen Gewebe-spezifischen Promotor, einen Zelltyp-spezifischen Promotor, usw.). In einigen Ausführungsformen ist die Nukleotidsequenz, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert, in funktionsfähiger Weise an einen konstitutiven Promotor gekoppelt.

[0327] In einigen Ausführungsformen liegt eine vorliegende genetisch modifizierte Wirtszelle *in vitro* vor. In einigen Ausführungsformen liegt eine vorliegende genetisch modifizierte Wirtszelle *in vivo* vor. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine prokaryontische Zelle oder ist von einer prokaryontischen Zelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine bakterielle Zelle oder ist von einer bakteriellen Zelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine Archaeen-Zelle oder ist von einer Archaeen-Bakterienzelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine eukaryontische Wirtszelle oder ist von einer eukaryontischen Wirtszelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine Pflanzenzelle oder ist von einer Pflanzenzelle abgeleitet. In einigen Ausführungsformen ist die vorliegende genetisch modifizierte Wirtszelle eine tierische Zelle oder ist von einer tierischen Zelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine Invertebraten-Zelle oder ist von einer Invertebraten-Zelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine Vertebraten-Zelle oder ist von

einer Vertebraten-Zelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine Säugetierzelle oder ist von einer Säugetierzelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine Nagetier-Zelle oder ist von einer Nagetier-Zelle abgeleitet. In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine menschliche Zelle oder ist von einer menschlichen Zelle abgeleitet.

[0328] Die vorliegende Offenbarung stellt ferner Nachkommen einer vorliegenden genetisch modifizierten Zelle bereit, wobei die Nachkommenschaft die gleiche exogene Nukleinsäure oder das gleiche exogene Polypeptid umfassen kann, wie die vorliegende genetisch modifizierte Zelle, von der diese abgeleitet ist. Die vorliegende Offenbarung stellt ferner eine Zusammensetzung bereit, die eine vorliegende genetisch modifizierte Wirtszelle umfasst. Genetisch modifizierte Stammzellen und genetisch modifizierte Vorläuferzellen

[0329] In einigen Ausführungsformen ist eine vorliegende genetisch modifizierte Wirtszelle eine genetisch modifizierte Stammzelle oder Vorläuferzelle. Geeignete Wirtszellen umfassen z. B. Stammzellen (adulte Stammzellen, embryonische Stammzellen, iPS-Zellen, usw.) und Vorläuferzellen (z. B. kardiäre Vorläuferzellen, neuronale Vorläuferzellen, usw.). Geeignete Wirtszellen umfassen Stammzellen und Vorläuferzellen von Säugetieren, einschließlich z. B. Nagetier-Stammzellen, Nagetier-Vorläuferzellen, humane Stammzellen, humane Vorläufer-Zellen, usw. Geeignete Wirtszellen umfassen in vitro-Wirtszellen, z. B. isolierte Wirtszelle.

[0330] In einigen Ausführungsformen umfasst eine vorliegende genetisch modifizierte Wirtszelle eine exogene auf DNA gerichtete RNA-Nukleinsäure. In einigen Ausführungsformen umfasst eine vorliegende genetisch modifizierte Wirtszelle eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, die für eine auf DNA gerichtete RNA kodiert. In einigen Ausführungsformen umfasst eine vorliegende genetisch modifizierte Wirtszelle ein exogenes ortsspezifisch modifizierendes Polypeptid (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder eine Variante von Cas9; ein chimäres Cas9, usw.). In einigen Ausführungsformen umfasst eine vorliegende genetisch modifizierte Wirtszelle eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert. In einigen Ausführungsformen umfasst eine vorliegende genetisch modifizierte Wirtszelle eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für 1) eine auf DNA gerichtete RNA und 2) ein ortsspezifisch modifizierendes Polypeptid kodiert.

[0331] In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Cas9/Csn1-Aminosäuresequenz oder zu den entsprechenden Teilen in einer der in SEQ ID NOS: 1–256 und 795–1346 dargestellten Aminosäuresequenzen umfasst.

ZUSAMMENSETZUNGEN

[0332] Die vorliegende Erfindung stellt eine Zusammensetzung bereit, die eine vorliegende auf DNA gerichtete RNA und/oder ein ortsspezifisch modifizierendes Polypeptid umfasst. In einigen Fällen ist das ortsspezifisch modifizierende Polypeptid ein vorliegendes chimäres Polypeptid. Eine vorliegende Zusammensetzung ist für die Durchführung eines Verfahrens der vorliegenden Offenbarung nützlich, z. B. eines Verfahrens zur ortsspezifischen Modifikation einer Ziel-DNA; eines Verfahrens zur ortsspezifischen Modifikation eines Polypeptids, das mit einer Ziel-DNA assoziiert ist, usw.

Zusammensetzungen, die eine auf DNA gerichtete RNA umfassen

[0333] Die vorliegende Erfindung stellt eine Zusammensetzung bereit, die eine vorliegende auf DNA gerichtete RNA umfasst. Die Zusammensetzung kann neben der auf DNA gerichteten RNA ein oder mehrere der folgenden Dinge umfassen: ein Salz, z. B. NaCl, MgCl₂, KCl, MgSO₄, usw.; ein pufferndes Mittel, z. B. einen Tris-Puffer, N-(2-Hydroxyethyl)piperazin-N'-(2-ethansulfonsäure) (HEPES), 2-(N-Morpholino)ethansulfonsäure (MES), MES-Natriumsalz, 3-(N-Morpholino)propansulfonsäure (MOPS), N-Tris[hydroxymethyl]methyl-3-Aminopropansulfonsäure (TAPS), usw.; ein solubilisierendes Mittel; ein Detergenz, z. B. ein nicht-ionisches Detergenz, wie beispielsweise Tween-20, usw.; einen Nuklease-Inhibitor; und ähnliches. Beispielsweise umfasst eine vorliegende Verbindung in einigen Fällen eine vorliegende auf DNA gerichtete RNA und einen Puffer zur Stabilisierung von Nukleinsäuren.

[0334] In einigen Ausführungsformen ist eine auf DNA gerichtete RNA, die in einer vorliegenden Zusammensetzung vorliegt, rein, z. B. mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens

etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder mehr als 99% rein, wobei "% Reinheit" bedeutet, dass die auf DNA gerichtete RNA in der angegebenen prozentualen Menge frei von anderen Makromolekülen oder Kontaminationen ist, die während der Erzeugung der auf DNA gerichteten RNA vorliegen können.

Zusammensetzungen, die ein vorliegendes chimäres Polypeptid umfassen

[0335] Die vorliegende Erfindung stellt eine Zusammensetzung bereit, die ein vorliegendes chimäres Polypeptid umfasst. Die Zusammensetzungen können neben der auf DNA gerichteten RNA ein oder mehrere der folgenden Dinge umfassen: ein Salz, z. B. NaCl, MgCl₂, KCl, MgSO₄, usw.; ein pufferndes Mittel, z. B. einen Tris-Puffer, HEPES, MES, MES-Natriumsalz, MOPS, TAPS, usw.; ein solubilisierendes Mittel; ein Detergenz, z. B. ein nicht-ionisches Detergenz, wie beispielsweise Tween-20, usw.; einen Protease-Inhibitor; ein reduzierendes Mittel (z. B. Dithiotreitol); und ähnliches.

[0336] In einigen Ausführungsformen ist ein vorliegendes chimäres Polypeptid, das in einer vorliegenden Zusammensetzung vorliegt, rein, z. B. mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder mehr als 99% rein, wobei "% Reinheit" bedeutet, dass das ortsspezifisch modifizierende Polypeptid in der angegebenen prozentualen Menge frei von anderen Proteinen, anderen Makromolekülen oder Kontaminationen ist, die während der Erzeugung des chimären Polypeptids vorliegen können.

Zusammensetzungen, die auf eine DNA gerichtete RNA
und ein ortsspezifisch modifizierendes Polypeptid umfassen

[0337] Die vorliegende Erfindung stellt eine Zusammensetzung bereit, die Folgendes umfasst: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert; und (ii) ein ortsspezifisch modifizierendes Polypeptid, oder ein Polynukleotid, das für selbiges kodiert. In einigen Fällen ist das ortsspezifisch modifizierende Polypeptid ein vorliegendes chimäres ortsspezifisch modifizierendes Polypeptid. In anderen Fällen ist das ortsspezifisch modifizierende Polypeptid ein natürlich vorkommendes ortsspezifisch modifizierendes Polypeptid. In einigen Fällen weist das ortsspezifisch modifizierende Polypeptid eine enzymatische Aktivität auf, die eine Ziel-DNA modifiziert. In anderen Fällen weist das ortsspezifisch modifizierende Polypeptid eine enzymatische Aktivität auf, die ein Polypeptid modifiziert, welches mit einer Ziel-DNA assoziiert ist. In noch anderen Fällen moduliert das ortsspezifisch modifizierende Polypeptid die Transkription der Ziel-DNA.

[0338] Die vorliegende Erfindung stellt eine Zusammensetzung bereit, die Folgendes umfasst: (i) eine auf DNA gerichtete RNA, wie sie oben beschrieben ist, oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0339] In einigen Fällen umfasst eine vorliegende Zusammensetzung folgendes: eine Zusammensetzung, die Folgendes umfasst: (i) eine vorliegende auf DNA gerichtete RNA, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0340] In anderen Ausführungsformen umfasst eine vorliegende Zusammensetzung folgendes: (i) ein Polynukleotid, das für eine vorliegende auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) ein Polynukleotid, das für das ortsspezifisch modifizierende Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf

DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

[0341] In einigen Ausführungsformen umfasst eine vorliegende Zusammensetzung beide RNA-Moleküle eines auf DNA gerichteten RNA-Doppelmoleküls. Somit umfasst in solchen Ausführungsformen eine vorliegende Zusammensetzung eine Aktivator-RNA, die ein Duplex-bildendes Segment umfasst, welches komplementär zu dem Duplex-formenden Segment einer Targeter-RNA ist (siehe **Fig. 1A**). Die Duplex-formenden Segmente der Aktivator-RNA und der Targeter-RNA hybridisieren, wobei der dsRNA-Duplex aus dem Protein-bindenden Segment der auf DNA gerichteten RNA gebildet wird. Die Targeter-RNA stellt ferner das auf DNA gerichtete Segment (einzelnsträngig) der auf DNA gerichteten RNA bereit und richtet somit die auf DNA gerichtete RNA auf eine spezifische Sequenz innerhalb der Ziel-DNA. Als ein nicht-einschränkendes Beispiel umfasst das Duplex-bildende Segment der Aktivator-RNA eine Nukleotidsequenz, die mindestens etwa 70%, mindestens etwa 80%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98% oder 100% Identität mit der Sequenz 5'-UAGCAAGUUAUUUU-3' (SEQ ID NO: 562) aufweist. Als weiteres nicht-einschränkendes Beispiel umfasst das Duplex-bildende Segment der Targeter-RNA eine Nukleotidsequenz, die mindestens etwa 70%, mindestens etwa 80%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98% oder 100% Identität mit der Sequenz 5'-GUUUUAGAGCUA-3' (SEQ ID NO: 679) aufweist.

[0342] Die vorliegende Erfindung stellt eine Zusammensetzung bereit, die Folgendes umfasst: (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0343] In einigen Fällen umfasst eine vorliegende Zusammensetzung beispielsweise: (i) eine auf DNA gerichtete RNA, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0344] Als weiteres Beispiel umfasst eine vorliegende Zusammensetzung in einigen Fällen: (i) ein DNA-Polynukleotid, das für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) ein Polynukleotid, das für das ortsspezifisch modifizierende Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0345] Eine vorliegende Zusammensetzung kann zusätzlich zu (i) einer vorliegenden auf DNA gerichteten RNA oder einem DNA-Polynukleotid, das für selbiges kodiert; und (ii) einem ortsspezifisch modifizierenden Polypeptid, oder einem Polynukleotid, das für selbiges kodiert, eines oder mehrere der folgenden Dinge umfassen: ein Salz, z. B. NaCl, MgCl₂, KCl, MgSO₄, usw.; eine pufferende Substanz, z. B. einen Tris-Puffer, HEPES, MES, MES-Natriumsalz, MOPS, TAPS, usw.; ein solubilisierendes Mittel; ein Detergenz, z. B. ein nicht-ionisches Detergenz, wie beispielsweise Tween-20, usw.; einen Protease-Inhibitor, ein reduzierendes Mittel (z. B. Dithiothreitol); und ähnliches.

[0346] In einigen Fällen sind die einzelnen Bestandteile der Zusammensetzung rein, z. B. ist jedes der Bestandteile mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder mindestens 99% rein. In einigen Fällen sind die einzelnen Bestandteile einer vorliegenden Zusammensetzung rein, bevor sie der Zusammensetzung zugesetzt werden.

[0347] Beispielsweise ist in einigen Ausführungsformen ein ortsspezifisch modifizierendes Polypeptid, das in einer vorliegenden Zusammensetzung vorhanden ist, rein, z. B. mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, mindestens etwa 99% oder mehr als 99% rein, wobei "% Reinheit" bedeutet, dass das ortsspezifisch modifizierende Polypeptid in der angegebenen prozentualen Menge frei von anderen Proteinen (z. B. Proteinen, bei denen es sich nicht um das ortsspezifisch modifizierende Polypeptid handelt), anderen Makromolekülen oder Kontaminationen ist, die während der Erzeugung des ortsspezifisch modifizierenden Polypeptids vorhanden sein können.

KITS

[0348] Die vorliegende Erfindung stellt Kits zur Ausführung eines vorliegenden Verfahrens bereit. Ein vorliegender Kit kann eines oder mehrere der folgenden Dinge umfassen: ein ortsspezifisch modifizierendes Polypeptid; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert; eine auf DNA gerichtete RNA; eine Nukleinsäure, die für eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; eine Aktivator-RNA; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine Aktivator-RNA kodiert; eine Targeter-RNA; und eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine Targeter-RNA kodiert. Ein ortsspezifisch modifizierendes Polypeptid; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert; eine auf DNA gerichtete RNA; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; eine Aktivator-RNA; eine Nukleinsäure, welche eine Nukleotidsequenz umfasst, welche für eine Aktivator-RNA kodiert; eine Targeter-RNA, und eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine Targeter-RNA kodiert, werden oben im Detail beschrieben. Ein Kit kann ferner einen Komplex umfassen, der zwei oder mehrere der folgenden Dinge umfasst: ein ortsspezifisch modifizierendes Polypeptid; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert; eine auf DNA gerichtete RNA; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; eine Aktivator-RNA; eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine Aktivator-RNA kodiert; eine Targeter-RNA; und eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine Targeter-RNA kodiert.

[0349] In einigen Ausführungsformen umfasst ein vorliegendes Kit ein ortsspezifisch modifizierendes Polypeptid oder ein Polynukleotid, das für selbiges kodiert. In einigen Ausführungsformen umfasst das ortsspezifisch modifizierende Polypeptid folgendes: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen weist der Aktivitätsteil des ortsspezifisch modifizierenden Polypeptids eine verringerte oder inaktivierte Nuklease-Aktivität auf. In einigen Fällen ist das ortsspezifisch modifizierende Polypeptid ein chimäres ortsspezifisch modifizierendes Polypeptid.

[0350] In einigen Ausführungsformen umfasst ein vorliegender Kit folgendes: ein ortsspezifisch modifizierendes Polypeptid oder ein Polynukleotid, welches für selbiges kodiert, und ein Reagenz zur Rekonstituierung und/oder Verdünnung des ortsspezifisch modifizierenden Polypeptids. In anderen Ausführungsformen umfasst ein vorliegender Kit eine Nukleinsäure (z. B. DNA, RNA), die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert. In einigen Ausführungsformen umfasst ein vorliegender Kit folgendes: eine Nukleinsäure (z. B. DNA, RNA), die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert; und ein Reagenz zur Rekonstituierung und/oder Verdünnung des ortsspezifisch modifizierenden Polypeptids.

[0351] Ein vorliegender Kit, der ein ortsspezifisch modifizierendes Polypeptid umfasst, oder ein Polynukleotid, das für selbiges kodiert, kann ferner ein oder mehrere zusätzliche Reagenzien umfassen, wobei solche zusätzlichen Reagenzien ausgewählt sein können aus: einem Puffer zur Einbringung des ortsspezifisch modifizierenden Polypeptids in eine Zelle; einem Waschpuffer; einem Kontrollreagenz; einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid; einem Reagenz für die in vitro Herstellung des ortsspezifisch modifizierenden Polypeptids ausgehend von DNA, und ähnliches. In einigen Fällen ist das ortsspezifisch modifizierende Polypeptid, das in dem vorliegenden Kit enthalten ist, ein chimäres ortsspezifisch modifizierendes Polypeptid, wie es oben beschrieben ist.

[0352] In einigen Ausführungsformen umfasst ein vorliegender Kit eine auf DNA gerichtete RNA, oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert. In

einigen Ausführungsformen umfasst die auf DNA gerichtete RNA ein drittes Segment (wie oben beschrieben). In einigen Ausführungsformen umfasst ein vorliegender Kit folgendes: (i) eine auf DNA gerichtete RNA, oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) ein ortsspezifisch modifizierendes Polypeptid, oder ein DNA-Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifisch enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird. In einigen Ausführungsformen weist der Aktivitätsteil des ortsspezifisch modifizierenden Polypeptids keine enzymatische Aktivität auf (umfasst eine inaktivierte Nuklease, z. B. durch Mutation). In einigen Fällen umfasst der Kit eine auf DNA gerichtete RNA und ein ortsspezifisch modifizierendes Polypeptid. In anderen Fällen umfasst der Kit folgendes: (i) eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; und (ii) eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert.

[0353] Als weiteres Beispiel kann ein vorliegender Kit folgendes umfassen: (i) eine auf DNA gerichtete RNA, oder ein DNA-Polynukleotid, das für selbige kodiert, umfassend: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, umfassend: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird. In einigen Fällen umfasst der Kit folgendes: (i) eine auf DNA gerichtete RNA; und ein ortsspezifisch modifizierendes Polypeptid. In anderen Fällen umfasst der Kit folgendes: (i) eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; und (ii) eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert.

[0354] Die vorliegende Offenbarung stellt einen Kit bereit, der Folgendes umfasst: (1) einen rekombinanten Expressions-Vektor, der Folgendes umfasst: (i) eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) eine Nukleotidsequenz, die für das ortsspezifisch modifizierende Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird; und (2) ein Reagenz zur Rekonstituierung und/oder Verdünnung des Expressionsvektors.

[0355] Die vorliegende Offenbarung stellt einen Kit bereit, der Folgendes umfasst: (1) einen rekombinanten Expressionsvektor, der Folgendes umfasst: (i) eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (ii) eine Nukleotidsequenz, die für das ortsspezifisch modifizierende Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird; und (2) ein Reagenz zur Rekonstituierung und/oder Verdünnung des rekombinanten Expressionsvektors.

[0356] Die vorliegende Offenbarung stellt einen Kit bereit, der Folgendes umfasst: (1) einen rekombinanten Expressionsvektor, der eine Nukleinsäure umfasst, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert, umfassend: (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und (ii) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und (2) ein Reagenz zur Rekonstituierung und/oder Verdünnung des rekombinanten Expressionsvektors. In einigen Ausführungsformen dieses Kits umfasst der Kit: einen rekombinanten Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierende Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete

RNA bestimmt wird. In anderen Ausführungsformen dieses Kits, umfasst der Kit: (1) einen rekombinanten Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst: (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

[0357] In einigen Ausführungsformen der oben beschriebenen Kits umfasst der Kit eine Aktivator-RNA oder eine Targeter-RNA. In einigen Ausführungsformen der oben beschriebenen Kits umfasst der Kit ein auf DNA gerichtetes RNA-Einzelmolekül. In einigen Ausführungsformen der oben beschriebenen Kits umfasst der Kit zwei oder mehrere auf DNA gerichtete RNA-Doppelmoleküle oder -Einzelmoleküle. In einigen Ausführungsformen der oben beschriebenen Kits kann eine auf DNA gerichtete RNA (die z. B. zwei oder mehrere auf DNA gerichtete RNAs enthält) als Array bereitgestellt werden (z. B. ein Array aus RNA-Molekülen, ein Array aus DNA-Molekülen, die für auf DNA gerichtete RNA(s) kodieren, usw.). Solche Kits können beispielsweise zur Verwendung in Verbindung mit den oben beschriebenen genetisch modifizierten Wirtszellen nützlich sein, welche ein vorliegendes ortsspezifisch modifizierendes Polypeptid umfassen. In einigen Ausführungsformen der oben beschriebenen Kits umfasst der Kit ferner ein Donor-Polynukleotid, um die gewünschte genetische Modifikation zu bewirken. Bestandteile eines vorliegenden Kits können in getrennten Behältern vorliegen; oder sie können kombiniert in einem einzelnen Behälter vorliegen.

[0358] Alle der oben beschriebenen Kits können ferner ein oder mehrere zusätzliche Reagenzien umfassen, wobei solche zusätzlichen Reagenzien ausgewählt sein können aus: einem Verdünnungspuffer; einem Rekonstituierungspuffer; einem Waschpuffer; einem Kontrollreagenz; einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid; einem Reagenz für die in vitro Herstellung des ortsspezifisch modifizierenden Polypeptids ausgehend von DNA, und ähnliches.

[0359] Zusätzlich zu den oben genannten Bestandteilen kann ein vorliegender Kit ferner Hinweise für die Verwendung der Bestandteile des Kits zur Durchführung der vorliegenden Verfahren umfassen. Diese Anweisungen zur Durchführung der vorliegenden Verfahren sind im Allgemeinen auf einem geeigneten Speichermedium aufgezeichnet. Beispielsweise können die Anweisungen auf ein Substrat, wie beispielsweise Papier oder Plastik, usw., aufgedruckt sein. Somit können die Instruktionen in den Kits als Packungsbeilage vorliegen, als Beschriftung des Behälters des Kits oder Bestandteilen davon (d. h. mit der Verpackung oder einer Unterverpackung assoziiert), usw. In anderen Ausführungsformen liegen die Instruktionen als elektronische Speicherdatei auf einem geeigneten Computerlesbarem Speichermedium vor, z. B. auf CD-ROM, Diskette, Flash-Drive, usw. In noch weiteren Ausführungsformen sind die eigentlichen Instruktionen nicht im Kit vorhanden, sondern es werden lediglich Mittel zum Erhalt der Anweisungen von einer entfernten Quelle, z. B. über das Internet, bereitgestellt. Ein Beispiel für diese Ausführungsform ist ein Kit, der eine Web-Adresse enthält, unter der die Anweisungen eingesehen werden können und/oder von der die Anweisungen heruntergeladen werden können. Wie die Anweisungen sind auch die Mittel zum Erhalt der Anweisungen auf einem geeigneten Substrat aufgezeichnet.

NICHT-HUMANE GENETISCH MODIFIZIERTE ORGANISMEN

[0360] In einigen Ausführungsformen wurde eine genetisch modifizierte Wirtszelle genetisch mit einer exogenen Nukleinsäure modifiziert, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.). Wenn eine solche Zelle ein eukaryontischer einzelliger Organismus ist, kann die modifizierte Zelle als genetisch modifizierter Organismus angesehen werden. In einigen Ausführungsformen ist ein vorliegender genetisch modifizierter Organismus ein Cas9-transgener multizellulärer Organismus.

[0361] In einigen Ausführungsformen kann eine vorliegende genetisch modifizierte, nicht-humane Wirtszelle (z. B. eine Zelle, die genetisch mit einer exogenen Nukleinsäure modifiziert wurde, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert, z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.) einen vorliegenden genetisch modifizierten nicht-humanen Organismus erzeugen (z. B. eine Maus, einen Fisch, einen Frosch, eine Fliege, einen Wurm, usw.). Wenn die genetisch modifizierte Wirtszelle beispielsweise eine pluripotente Stammzelle (d. h. PSC) oder eine Keimzelle (z. B. Spermium, Oozyte, usw.) ist, kann ein vollständiger, genetisch modifizierter Organismus von der genetisch modifizierten Wirtszelle abgeleitet werden. In einigen Ausführungsformen ist die genetisch modifizierte Wirtszelle eine pluripotente Stammzelle (z. B. ESC, iPSC,

pluripotente Pflanzen-Stammzelle, usw.) oder eine Keimzelle (z. B. Spermium, Oozyte, usw.), die entweder in vivo oder in vitro vorliegt und zu einem genetisch modifizierten Organismus führen kann. In einigen Ausführungsformen ist die genetisch modifizierte Wirtszelle eine Vertebraten-PSC (z. B. ESC, iPSC, usw.) und wird verwendet, um einen genetisch modifizierten Organismus zu erzeugen (z. B. durch Injektion einer PSC in eine Blastozyste, um ein chimäres Tier/Mosaik-Tier zu erzeugen, welches dann gekreuzt werden kann, um genetisch modifizierte Organismen zu erzeugen, bei denen es sich nicht um einen chimären Organismus/Mosaik-Organismus handelt; Grafting im Fall von Pflanzen). Jedes handhabbare Verfahren/Protokoll zur Herstellung eines genetisch modifizierten Organismus, einschließlich der vorliegend beschriebenen Verfahren, ist zur Erzeugung einer genetisch modifizierten Wirtszelle, die eine exogene Nukleinsäure umfasst, welche eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert (ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.), geeignet. Verfahren zur Herstellung genetisch modifizierter Organismen sind im Stand der Technik bekannt. Siehe beispielsweise Cho et al., *Curr Protoc Cell Biol.* 2009 Mar; Chapter 19: Unit 19.11: Generation of transgenic mice; Gama et al., *Brain Struct Funct.* 2010 Mar; 214(2–3): 91–109. Epub 2009 Nov 25: Animal transgenesis: an overview; Husaini et al., *GM Crops.* 2011 Jun–Dec; 2(3): 150–62. Epub 2011 Jun 1: Approaches for gene targeting and targeted gene expression in plants.

[0362] In einigen Ausführungsformen umfasst ein genetisch modifizierter Organismus eine Ziel-Zelle für die Verfahren der vorliegenden Erfindung und kann somit als Quelle für Ziel-Zellen angesehen werden. Wenn beispielsweise eine genetisch modifizierte Zelle, die eine exogene Nukleinsäure umfasst, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.), verwendet wird, um einen ein genetisch modifizierten Organismus herzustellen, dann umfassen die Zellen des genetisch modifizierten Organismus die exogene Nukleinsäure, welche die Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.). In einigen dieser Ausführungsformen kann die DNA der Zelle oder der Zellen des genetisch modifizierten Organismus für eine Modifikation ausgewählt werden, indem eine auf DNA gerichtete RNA (oder eine DNA, die für eine auf DNA gerichtete RNA kodiert) und wahlweise eine Donor-Nukleinsäure in die Zelle oder in die Zellen eingebracht wird. Beispielsweise kann die Einbringung einer auf DNA gerichteten RNA (oder einer DNA, die für eine auf DNA gerichtete RNA kodiert) in eine Untergruppe von Zellen (z. B. Gehirnzellen, Darmzellen, Nierenzellen, Lungenzellen, Blutzellen, usw.) des genetisch modifizierten Organismus die DNA solcher Zellen für eine Modifikation anvisieren, wobei der genomische Ort derselben von der auf DNA gerichteten Sequenz der eingebrachten auf DNA gerichteten RNA abhängen wird.

[0363] In einigen Ausführungsformen ist ein genetisch modifizierter Organismus eine Quelle für Zielzellen für die erfindungsgemäßen Verfahren. Beispielsweise kann ein genetisch modifizierter Organismus, der Zellen umfasst, welche genetisch mit einer exogenen Nukleinsäure modifiziert werden, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.), eine Quelle für genetisch modifizierte Zellen darstellen, beispielsweise für PSCs (z. B. ESCs, iPSCs, Spermien, Oozyten, usw.), Neuronen, Vorläuferzellen, Kardiomyozyten, usw.

[0364] In einigen Ausführungsformen ist eine genetisch modifizierte Zelle eine PSC, die eine exogene Nukleinsäure umfasst, welche eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.). Somit kann die PSC in der Weise eine Zielzelle sein, dass die DNA der PSC für eine Modifikation ausgewählt werden kann, indem man in die PSC eine auf DNA gerichtete RNA (oder eine DNA, die für eine auf DNA gerichtete RNA kodiert) und wahlweise eine Donor-Nukleinsäure einbringt, wobei der genomische Ort der Modifikation von der auf DNA gerichteten Sequenz der eingebrachten auf DNA gerichteten RNA abhängen wird. Somit können die vorliegend beschriebenen Verfahren in einigen Ausführungsformen dazu verwendet werden, um die DNA von PSCs, die von einem vorliegenden genetisch modifizierten Organismus abgeleitet wurden, zu modifizieren (z. B. zu deletieren und/oder um eine beliebige gewünschte genomische Stelle zu ersetzen). Solche modifizierten PSCs können dann dazu verwendet werden, um Organismen zu erzeugen, die sowohl (i) eine exogene Nukleinsäure aufweisen, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.) als auch (ii) eine DNA-Modifikation, die in die PSC eingebracht wurde, umfassen.

[0365] Eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.) kann unter der Kontrolle eines unbekannten Promotor stehen (d. h. in funktionsfähiger Weise mit einem unbekannten Promotor verknüpft sein) (z. B. wenn die Nukleinsäure sich auf Zufallsbasis in das Genom einer Wirtszelle integriert), oder sie kann unter der Kontrolle eines bekannten Promotor stehen (d. h. in funktionsfähiger Weise mit einem bekannten Promotor verknüpft sein). Geeignete bekannte Promotoren können jeden bekannten Promotor umfassen und schließen konstitutiv aktive Promotoren (z. B. CMV-Promotor), induzierbare Promotoren (z. B. Hitzeschock-Promotor, Tetracyclin-regulierter Promotor, Steroid-regulierter Promotor, Metall-regulierter Promotor, Östrogenrezeptor-regulierter Promotor, usw.), räumlich beschränkte und/oder zeitlich beschränkte Promotoren (z. B. ein Gewebe-spezifischer Promotor, ein Zelltyp-spezifischer Promotor, usw.) usw., ein.

[0366] Ein vorliegender genetisch modifizierter Organismus (z. B. ein Organismus, dessen Zellen eine Nukleotidsequenz umfassen, die für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.), kann jeder beliebige Organismus sein, einschließlich z. B. eine Pflanze; eine Alge; ein Invertebrat (z. B. ein Nesseltier, ein Stachelhäuter, ein Wurm, eine Fliege, usw.); ein Vertebrat (z. B. ein Fisch (z. B. ein Zebrafisch, ein Pufferfisch, ein Goldfisch, usw.), eine Amphibie (z. B. Salamander, Frosch, usw.), ein Reptil, ein Vogel, ein Säugetier, usw.); ein Huftier (z. B. eine Ziege, ein Schwein, ein Schaf, eine Kuh, usw.); ein Nagetier (z. B. eine Maus, eine Ratte, ein Hamster, ein Meerschweinchen); ein Hasenartiger (z. B. ein Kaninchen); usw.

[0367] In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäure 7–166 oder 731–1003 der in **Fig. 3** dargestellten Cas9/Csn1-Aminosäuresequenz aufweist oder zu den entsprechenden Teilen in einer der in SEQ ID NO: 1–246 und 795–1346 gezeigten Aminosäuresequenzen.

Transgen nicht-humane Tiere

[0368] Wie oben beschrieben wird in manchen Ausführungsformen eine vorliegende Nukleinsäure (z. B. eine Nukleotidsequenz, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, z. B. für ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.) oder ein vorliegender rekombinanter Expressionsvektor als Transgen verwendet, um ein transgenes Tier zu erzeugen, das ein ortsspezifisch modifizierendes Polypeptid erzeugt. Somit stellt die vorliegende Erfindung ferner ein transgenes nicht-humanes Tier bereit, wobei das Tier ein Transgen umfasst, das eine vorliegende Nukleinsäure umfasst, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert, z. B. für ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw., wie es oben beschrieben ist. In einigen Ausführungsformen umfasst das Genom des transgenen nicht-humanen Tiers eine vorliegende Nukleotidsequenz, die für ein ortsspezifisch modifizierendes Polypeptid kodiert. In einigen Ausführungsformen ist das transgene, nicht-humane Tier homozygot in Bezug auf die genetische Modifikation. In einigen Ausführungsformen ist das transgene, nicht-humane Tier heterozygot in Bezug auf die genetische Modifikation. In einigen Ausführungsformen ist das transgene, nicht-humane Tier ein Vertebrat, z. B. ein Fisch (z. B. Zebrafisch, Goldfisch, Pufferfisch, Höhlenfisch, usw.), eine Amphibie (Frosch, Salamander, usw.), ein Vogel (z. B. Huhn, Truthahn, usw.), ein Reptil (z. B. Schlange, Eidechse, usw.), ein Säugetier (z. B. ein Huftier, z. B. ein Schwein, eine Kuh, eine Ziege, ein Schaf, usw.), ein Hasenartiger (z. B. ein Kaninchen); ein Nagetier (z. B. eine Ratte, eine Maus); ein nicht-humaner Primat; usw.), usw.

[0369] Eine exogene Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.) kann unter der Kontrolle eines unbekannten Promotor stehen (d. h. in funktionsfähiger Weise mit einem unbekannten Promotor verknüpft sein) (z. B. wenn die Nukleinsäure sich auf Zufallsbasis in das Genom einer Wirtszelle integriert), oder sie kann unter der Kontrolle eines bekannten Promotor stehen (d. h. in funktionsfähiger Weise mit einem bekannten Promotor verknüpft sein). Geeignete bekannte Promotoren können jeden bekannten Promotor umfassen und schließen konstitutiv aktive Promotoren (z. B. CMV-Promotor), induzierbare Promotoren (z. B. Hitzeschock-Promotor, Tetracyclin-regulierter Promotor, Steroid-regulierter Promotor, Metall-regulierter Promotor, Östrogenrezeptor-regulierter Promotor, usw.), räumlich beschränkte und/oder zeitlich beschränkte Promotoren (z. B. ein Gewebe-spezifischer Promotor, ein Zelltyp-spezifischer Promotor, usw.) usw., ein.

[0370] In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz mit mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Cas9/Csn1-Aminosäuresequenz oder zu den entsprechenden Teilen in einer der in SEQ ID NO: 1–246 und 795–1346 gezeigten Aminosäuresequenzen.

Transgene Pflanzen

[0371] Wie vorliegend beschrieben wird in einigen Ausführungsformen eine vorliegende Nukleinsäure (z. B. eine Nukleotidsequenz, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, z. B. für ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.) oder ein vorliegender rekombinanter Expressionsvektor als Transgen verwendet, um eine transgene Pflanze zu erzeugen, die für ein ortsspezifisch modifizierendes Polypeptid kodiert. Somit stellt die vorliegende Erfindung ferner eine transgene Pflanze bereit, wobei die Pflanze ein Transgen umfasst, welches eine vorliegende Nukleinsäure umfasst, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert, (z. B. für ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder ein Varianten-Cas9; ein chimäres Cas9, usw.), wie es oben beschrieben wird. In einigen Ausführungsformen umfasst das Genom der transgenen Pflanze eine vorliegende Nukleinsäure. In einigen Ausführungsformen ist die transgene Pflanze homozygot in Bezug auf die genetische Modifikation. In einigen Ausführungsformen ist die transgene Pflanze heterozygot in Bezug auf die genetische Modifikation.

[0372] Verfahren zur Einbringung von exogenen Nukleinsäuren in Pflanzenzellen sind im Stand der Technik hinreichend bekannt. Solche Pflanzenzellen werden als "transformiert" angesehen, wie es oben beschrieben ist. Geeignete Verfahren umfassen virale Infektion (z. B. doppelsträngige DNA-Viren), Transfektion, Konjugation, Protoplasten-Fusion, Elektroporation, Particle-Gun-Technologie, Kaliumphosphat-Präzipitation, direkte Mikroinjektion, Siliziumkarbid-Whisker-Technologie, Agrobacterium-vermittelte Transformation und ähnliches. Die Wahl des Verfahrens hängt im Allgemeinen von der Art der zu transformierenden Zelle und von den Umständen ab, unter denen die Transformation erfolgt (d. h. in vitro, ex vivo oder in vivo).

[0373] Transformationsverfahren, die auf dem Bodenbakterium *Agrobacterium tumefaciens* basieren, sind insbesondere für die Einbringung eines exogenen Nukleinsäure-Moleküls in eine vaskuläre Pflanze nützlich. Die Wildtypform von *Agrobacterium* umfasst ein Ti (Tumor-induzierendes) Plasmid, welches das Wachstum von tumorigenen Kronengallen auf den Wirtspflanzen steuert. Der Transfer der Tumor-induzierenden T-DNA-Region des T-Plasmids in ein Pflanzengenom erfordert die vom T-Plasmid kodierten Virulenz-Gene sowie auch die T-DNA-Grenzen, bei denen es sich um eine Gruppe von direkten DNA-Repeats handelt, welche die zu transferierende Region abgrenzen. Ein *Agrobacterium*-basierter Vektor ist eine modifizierte Form des T-Plasmids, in der die Tumor-induzierenden Funktionen gegen die Nukleinsäuresequenz von Interesse, die in den Pflanzenwirt eingebracht werden soll, ausgetauscht werden.

[0374] Eine *Agrobacterium*-vermittelte Transformation verwendet im Allgemeinen co-integrierende Vektoren oder binäre Vektorsysteme, in denen die Bestandteile des T-Plasmids zwischen einem Hilfsvektor, der permanent im *Agrobacterium*-Wirt verbleibt und die Virulenz-Gene trägt, und einem Shuttle-Vektor, der das Gen von Interesse eingegrenzt von T-DNA-Sequenzen enthält, verteilt sind. Eine Vielzahl von binären Vektoren ist im Stand der Technik hinreichend bekannt und kommerziell verfügbar, beispielsweise von Clontech (Palo Alto, Kalifornien). Verfahren zur Co-Kultivierung von *Agrobacterium* mit kultivierten Pflanzenzellen oder verletztem Gewebe, wie beispielsweise Blattgewebe, Wurzelexplantaten, Hypokotyledonen, Stängelteilen oder Knollen sind z. B. ebenfalls im Stand der Technik hinreichend bekannt, siehe z. B. Glick und Thompson, (Hrsg.), *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton, Fla.: CRC Press (1993).

[0375] Eine Mikroprojektil-vermittelte Transformation kann ebenfalls verwendet werden, um eine vorliegende transgene Pflanze zu erzeugen. Dieses Verfahren, das zuerst von Klein et al. (*Nature* 327: 70–73 (1987)) beschrieben wurde, beruht auf Mikroprojektilen, wie beispielsweise Gold oder Wolfram, die durch Präzipitation mit Kalziumchlorid, Spermidin oder Polyethylenglykol mit dem gewünschten Nukleinsäure-Molekül beschichtet sind. Die Mikroprojektilpartikel werden mit hoher Geschwindigkeit in ein Angiospermen-Gewebe beschleunigt, wobei eine Vorrichtung wie beispielsweise die BIOLISTIC PD-1000 (Biorad; Hercules Kalifornien) verwendet wird.

[0376] Eine vorliegende Nukleinsäure kann in der Weise in eine Pflanze eingebracht werden, dass die Nukleinsäure in der Lage ist, in die Pflanzenzelle(n) einzudringen, z. B. durch ein in vivo oder ex vivo-Protokoll. Mit "in vivo" ist gemeint, dass die Nukleinsäure an einen lebenden Körper einer Pflanze verabreicht wird, z. B.

Infiltration. Mit "ex vivo" ist gemeint, dass Zellen oder Explantate außerhalb der Pflanze modifiziert werden, und anschließend solche Zellen oder Organe zu einer Pflanze regeneriert werden. Eine Vielzahl von Vektoren, die für eine stabile Transformation von Pflanzenzellen oder für die Herstellung von transgenen Pflanzen geeignet sind, sind beschrieben worden, einschließlich derjenigen, die in Weissbach und Weissbach, (1989) *Methods for Plant Molecular Biology* Academic Press, und Gelvin et al., (1990) *Plant Molecular Biology Manual*, Kluwer Academic Publishers beschrieben wurden. Spezifische Beispiele umfassen solche, die von einem Ti-Plasmid von *Agrobacterium tumefaciens* abgeleitet sind, wie beispielsweise die von Herrera-Estrella et al. (1983) *Nature* 303: 209, Bevan (1984) *Nucl Acid Res.* 12: 8711–8721, Klee (1985) *Bio/Technolo* 3: 637–642 beschrieben wurden. Alternativ dazu können Nicht-Ti-Vektoren verwendet werden, um die DNA in Pflanzen und Zellen zu transferieren, wobei Übertragungsverfahren mit freier DNA verwendet werden. Durch Verwendung dieser Verfahren können transgene Pflanzen, wie beispielsweise Weizen, Reis (Christou (1991) *Bio/Technology* 9: 957–9 und 4462) und Mais (Gordon-Kamm (1990) *Plant Cell* 2: 603–618) erzeugt werden. Ein unreifer Embryo kann bei Monokotylen ebenfalls ein gutes Zielgewebe für Verfahren zur direkten Bereitstellung von DNA durch Verwendung der Particle-Gun (Weeks et al. (1993) *Plant Physiol* 102: 1077–1084; Vasil (1993) *Bio/Technolo* 10: 667–674; Wan und Lemeaux (1994) *Plant Physiol* 104: 37–48) und für den *Agrobacterium*-vermittelten DNA-Transfer (Ishida et al. (1996) *Nature Biotech* 14: 745–750). Beispielhafte Verfahren für die Einbringung von DNA in Chloroplasten sind biolistisches Bombardment, Polyethylenglykol-Transformation von Protoplasten und Mikroinjektion (Danieli et al *Nat. Biotechnol* 16: 345–348, 1998; Staub et al *Nat. Biotechnol* 18: 333–338, 2000; O'Neill et al *Plant J.* 3: 729–738, 1993; Knoblauch et al *Nat. Biotechnol* 17: 906–909; US-Pat. Nrn. 5, 451,513; 5,545,817; 5,545,818 und 5,576,198; Internat. Anmeldenummer WO 95/16783; und in Boynton et al., *Methods in Enzymology* 217: 510–536 (1993), Svab et al., *Proc. Natl. Acad. Sci. USA* 90: 913–917 (1993), und McBride et al., *Proc. Nat. Acad. Sci. USA* 91: 7301–7305 (1994)). Beliebige Vektoren, die für Verfahren des biolistischen Bombardements, Polyethylenglykol-Transformation von Protoplasten und Mikroinjektion geeignet sind, werden als Targeting-Vektor für die Chloroplasten-Transformation geeignet sein. Ein beliebiger doppelsträngiger DNA-Vektor kann als Transformationsvektor verwendet werden, insbesondere wenn das Verfahren der Einbringung nicht *Agrobacterium* benutzt.

[0377] Pflanzen, die genetisch modifiziert werden können, umfassen Getreide, Futterpflanzen, Früchte, Gemüse, Ölsamen-Nutzpflanzen, Palmen, Waldpflanzen und Weinpflanzen. Spezifische Beispiele für Pflanzen, die modifiziert werden können, sind die folgenden: Mais, Banane, Erdnuss, Erbsen, Sonnenblume, Tomate, Canola, Tabak, Weizen, Gerste, Hafer, Kartoffel, Sojabohnen, Baumwolle, Nelken, Sorghum, Lupine und Reis.

[0378] Ebenfalls bereitgestellt durch die vorliegende Erfindung werden transformierte Pflanzenzellen, Gewebe, Pflanzen und Produkte, welche die transformierten Pflanzenzellen umfassen. Ein Merkmal der vorliegenden transformierten Zellen und Gewebe sowie der Produkte, welche diese umfassen, ist das Vorhandensein einer vorliegenden Nukleinsäure integriert im Genom und das Erzeugen eines ortsspezifisch modifizierenden Polypeptids durch die Pflanzenzellen, z. B. eines natürlich vorkommenden Cas9; eines modifizierten, d. h. mutierten Cas9 oder eines Varianten-Cas9; eines chimären Cas9, usw. Rekombinante Pflanzenzellen der vorliegenden Erfindung sind als Populationen aus rekombinanten Zellen nützlich oder als ein Gewebe, ein Samen, eine Gesamtpflanze, ein Stängel, eine Frucht, ein Blatt, eine Wurzel, eine Blüte, ein Stängel, eine Knolle, ein Korn, ein Tierfutter, ein Pflanzenfeld, und ähnliches.

[0379] Eine Nukleinsäure, die eine Nukleotidsequenz, welche für ein ortsspezifisch modifizierendes Polypeptid kodiert (z. B. für ein natürlich vorkommendes Cas9; ein modifiziertes, d. h. mutiertes Cas9 oder eine Varianten-Cas9; ein chimäres Cas9, usw.), kann unter der Kontrolle eines unbekannten Promotor stehen (d. h. in funktionsfähiger Weise mit einem unbekannten Promotor verknüpft sein) (z. B. wenn die Nukleinsäure sich auf Zufallsbasis in das Genom einer Wirtszelle integriert), oder sie kann unter der Kontrolle eines bekannten Promotor stehen (d. h. in funktionsfähiger Weise mit einem bekannten Promotor verknüpft sein). Geeignete bekannte Promotoren können jeden bekannten Promotor umfassen und schließen konstitutiv aktive Promotoren, induzierbare Promotoren, räumlich beschränkte und/oder zeitlich beschränkte Promotoren, usw.

[0380] In einigen Fällen umfasst das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz mit mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** dargestellten Cas9/Csn1-Aminosäuresequenz oder zu den entsprechenden Teilen in einer der in SEQ ID NO: 1–246 und 795–1346 gezeigten Aminosäuresequenzen.

[0381] Ebenfalls durch die vorliegende Erfindung bereitgestellt wird reproduktives Material einer vorliegenden transgenen Pflanze, wobei reproduktives Material Samen, Nachkommen-Pflanzen und klonales Material umfasst.

DEFINITIONEN – TEIL II

[0382] Der Begriff "natürlich vorkommend" oder "unmodifiziert", wie er vorliegend auf eine Nukleinsäure, ein Polypeptid, eine Zelle oder einen Organismus angewendet wird, bezeichnet eine Nukleinsäure, ein Polypeptid, eine Zelle oder einen Organismus der/die/das in der Natur vorgefunden wird. Beispielsweise ist eine Polypeptid- oder eine Polynukleotidsequenz, die in einem Organismus (einschließlich Viren) vorliegt und von einer Quelle in der Natur isoliert werden kann und nicht bewusst durch einen Menschen im Labor modifiziert wurde, natürlich vorkommend.

[0383] "Heterolog" bedeutet wie vorliegend verwendet eine Nukleotid- oder Polypeptidsequenz, die nicht in der nativen Nukleinsäure bzw. dem nativen Protein vorgefunden wird. Beispielsweise kann ein ortsspezifisches Varianten-Cas9-Polypeptid in einem ortsspezifischen Varianten-Cas9-Fusionspolypeptid an ein heterologes Polypeptid fusioniert sein (d. h. an ein Polypeptid, bei dem es sich nicht um Cas9 handelt). Das heterologe Polypeptid kann eine Aktivität aufweisen (z. B. enzymatische Aktivität), die gleichermaßen durch das ortsspezifische Varianten-Cas9-Fusionspolypeptid ausgeübt wird. Eine heterologe Nukleinsäuresequenz kann an ein ortsspezifisches Varianten-Cas9-Polypeptid (z. B. durch genetische Modifikation) geknüpft sein, um eine Nukleotidsequenz zu erzeugen, welche für ein ortsspezifisches Varianten-Cas9-Fusionspolypeptid kodiert.

[0384] Der Begriff "chimäres Polypeptid" bezeichnet ein Polypeptid, das nicht natürlich vorkommt, z. B. eines, das durch künstliche Kombination von zwei ansonsten getrennten Segmenten aus Aminosäuresequenzen durch menschlichen Eingriff hergestellt wird. Somit ist ein chimäres Polypeptid auch das Ergebnis eines menschlichen Eingriffs. Somit ist ein Polypeptid, das eine chimäre Aminosäuresequenz umfasst, ein chimäres Polypeptid.

[0385] Mit "ortsspezifischem Polypeptid" oder "RNA-bindendem, ortsspezifischem Polypeptid" oder "RNA-bindendem, ortsspezifischem Polypeptid" ist ein Polypeptid gemeint, das RNA bindet und auf eine spezifische DNA-Sequenz gerichtet ist. Ein ortsspezifisches Polypeptid ist wie vorliegend beschrieben durch das RNA-Molekül, an das es gebunden ist, auf eine spezifische DNA-Sequenz gerichtet. Das RNA-Molekül umfasst eine Sequenz, die komplementär zu einer Ziel-Sequenz innerhalb der Ziel-DNA ist, wodurch das gebundene Polypeptid auf eine spezifische Stelle innerhalb der Ziel-DNA (die Ziel-Sequenz) gerichtet ist.

[0386] In manchen Ausführungsformen umfasst eine vorliegende Nukleinsäure (z. B. eine auf DNA gerichtet RNA, eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; eine Nukleinsäure, die für ein ortsspezifisches Polypeptid kodiert; usw.) eine Modifikation oder eine Sequenz, die für eine zusätzliche Eigenschaft sorgt (z. B. modifizierte oder regulierte Stabilität; subzelluläres Targeting; Tracking, z. B. eine fluoreszente Markierung; eine Bindungsstelle für ein Protein oder für einen Proteinkomplex, usw.). Nicht-einschränkende Beispiele umfassen: eine 5'-Kappe (z. B. eine 7-Methylguanylat-Kappe (m^7G)); einen 3'-polyadenylierten Anhang (d. h. einen 3'-Poly(A)-Anhang); eine Riboswitch-Sequenz (z. B. um eine regulierte Stabilität und/oder eine regulierte Zugänglichkeit durch Proteine und/oder Protein-Komplexe zu ermöglichen), eine Modifikation oder eine Sequenz, die die RNA zu einem subzellulären Ort führt (z. B. Nukleus, Mitochondrien, Chloroplasten und ähnliches); eine Modifikation oder Sequenz, die für ein Tracking sorgt (z. B. direkte Konjugation an ein fluoreszentes Molekül, Konjugation an einen Rest, der eine fluoreszente Detektion ermöglicht; eine Sequenz, die eine fluoreszente Detektion ermöglicht, usw.), eine Modifikation oder Sequenz, die eine Bindungsstelle für Proteine bereitstellt (z. B. Proteine, die an der DNA wirken, einschließlich transkriptionaler Aktivatoren, transkriptionaler Repressoren, DNA-Methyltransferasen, DNA-Demethylasen, Histon-Acetyltransferasen, Histon-Deacetylasen und ähnliches); und Kombinationen davon.

[0387] In einigen Ausführungsformen umfasst eine auf DNA gerichtet RNA ein zusätzliches Segment, entweder am 5'-Ende oder am 3'-Ende, das für eines der oben beschriebenen Merkmale sorgt. Beispielsweise kann ein geeignetes drittes Segment eine 5'-Kappe (z. B. eine 7-Methylguanylat-Kappe (m^7G)); einen 3'-polyadenylierten Anhang (d. h. einen 3'-Poly(A)-Anhang); eine Riboswitch-Sequenz (z. B. um eine regulierte Stabilität und/oder eine regulierte Zugänglichkeit durch Proteine und/oder durch Protein-Komplexe zu ermöglichen), eine Modifikation oder eine Sequenz, die die RNA zu einem subzellulären Ort führt (z. B. Nukleus, Mitochondrien, Chloroplasten und ähnliches); eine Modifikation oder Sequenz, die für ein Tracking sorgt (z. B. direkte Konjugation an ein fluoreszentes Molekül, Konjugation an einen Rest, der eine fluoreszente Detektion ermöglicht; eine Sequenz, die eine fluoreszente Detektion ermöglicht, usw.), eine Modifikation oder Sequenz, die eine Bindungsstelle für Proteine bereitstellt (z. B. Proteine, die an der DNA wirken, einschließlich transkriptionaler Aktivatoren, transkriptionaler Repressoren, DNA-Methyltransferasen, DNA-Demethylasen, Histon-Acetyltransferasen, Histon-Deacetylasen und ähnliches); und Kombinationen davon.

[0388] Eine vorliegende auf DNA gerichtete RNA und ein vorliegendes ortsspezifisches Polypeptid bilden einen Komplex (d. h. binden über nicht-kovalente Interaktionen). Die auf DNA gerichtete RNA sorgt für die Ziel-spezifität des Komplexes, indem sie eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz der Ziel-DNA ist. Das ortsspezifische Polypeptid des Komplexes sorgt für die ortsspezifische Aktivität. Mit anderen Worten wird das ortsspezifische Polypeptid durch seine Assoziation mit dem Protein-bindenden Segment der auf DNA gerichteten RNA zu einer Ziel-DNA-Sequenz geführt (z. B. einer Ziel-Sequenz in einer chromosomalen Nukleinsäure; einer Ziel-Sequenz in einer extrachromosomalen Nukleinsäure, z. B. einer episomalen Nukleinsäure, einer kleinen zirkulären Sequenz (minicircle), usw.; einer Ziel-Sequenz in einer mitochondrialen Nukleinsäure; einer Ziel-Sequenz in einer Chloroplasten-Nukleinsäure; einer Ziel-Sequenz in einem Plasmid; usw.) geführt.

[0389] In einigen Ausführungsformen umfasst eine vorliegende auf DNA gerichtete RNA zwei getrennte RNA-Moleküle (RNA-Polynukleotide) und wird vorliegend als "auf DNA gerichtetes RNA-Doppelmolekül" oder als "auf DNA gerichtete RNA aus zwei Molekülen" bezeichnet. In anderen Ausführungsformen ist eine vorliegende auf DNA gerichtete RNA ein RNA-Einzelmolekül (RNA-Einzelpolynukleotid) und wird vorliegend als "auf DNA gerichtetes RNA-Einzelmolekül" bezeichnet. Sofern nicht anders angegeben ist der Begriff "auf DNA gerichtete RNA" umfassend und bezeichnet sowohl auf DNA gerichtete RNA-Einzelmoleküle als auch auf DNA gerichtete RNA-Doppelmoleküle.

[0390] Eine vorliegende auf DNA gerichtete RNA aus zwei Molekülen umfasst zwei getrennte RNA-Moleküle (eine "Targeter-RNA" und eine "Aktivator-RNA"). Jedes dieser beiden RNA-Moleküle einer vorliegenden auf DNA gerichteten RNA aus zwei Molekülen umfasst einen Abschnitt von Nukleotiden die komplementär zueinander sind, so dass die komplementären Nukleotide der beiden RNA-Moleküle hybridisieren, wobei der doppelsträngige RNA-Duplex des Protein-bindenden Segments gebildet wird.

[0391] Ein vorliegendes auf DNA gerichtetes RNA-Einzelmolekül umfasst zwei Abschnitte aus Nukleotiden (eine "Targeter-RNA" und eine "Aktivator-RNA"), die komplementär zueinander sind, kovalent durch zwischenliegende Nukleotide ("Linker" oder "Linker-Nukleotide") miteinander verknüpft sind, und die hybridisieren, um den doppelsträngigen RNA-Duplex (dsRNA-Duplex) des Protein-bindenden Segments zu bilden, was zu einer Haarnadelschleifenstruktur führt. Die Targeter-RNA und die Aktivator-RNA können kovalent über das 3'-Ende der Targeter-RNA und das 5'-Ende der Aktivator-RNA verknüpft sein. Alternativ dazu können die Targeter-RNA und die Aktivator-RNA kovalent über das 5'-Ende der Targeter-RNA und das 3'-Ende der Aktivator-RNA verbunden sein.

[0392] Eine beispielhafte auf DNA gerichtete RNA aus zwei Molekülen umfasst ein crRNA-ähnliches ("CRISPR RNA" oder "Targeter-RNA" oder "crRNA" oder "crRNA-Repeat") Molekül und ein entsprechendes tracrRNA-ähnliches ("trans-wirkendes CRISPR RNA" oder "Aktivator-RNA" oder "tracrRNA") Molekül. Ein crRNA-ähnliches Molekül (Targeter-RNA) umfasst sowohl das auf DNA gerichtete Segment (einzelssträngig) der auf DNA gerichteten RNA und einen Abschnitt ("Duplex-bildendes Segment") von Nukleotiden, die eine Hälfte des dsRNA-Duplex des Protein-bindenden Segments der auf DNA gerichteten RNA bildet. Ein entsprechendes tracrRNA-ähnliches Molekül (Aktivator-RNA) umfasst einen Abschnitt von Nukleotiden ("Duplex-bildendes Segment"), der die andere Hälfte des dsRNA-Duplex des Protein-bindenden Segments der auf DNA gerichteten RNA bildet. Mit anderen Worten ist ein Abschnitt von Nukleotiden eines crRNA-ähnlichen Moleküls komplementär zu und hybridisiert mit einem Abschnitt von Nukleotiden eines tracrRNA-ähnlichen Moleküls, wobei der dsRNA-Duplex der Protein-bindenden Domäne der auf DNA gerichteten RNA gebildet wird. Somit kann gesagt werden, dass jedes crRNA-ähnliche Molekül ein entsprechendes tracrRNA-ähnliches Molekül hat. Das crRNA-ähnliche Molekül stellt zusätzlich das einzelssträngige auf DNA gerichtete Segment bereit. Somit hybridisieren ein crRNA-ähnliches und ein tracrRNA-ähnliches Molekül (als entsprechendes Paar), wobei eine auf DNA gerichtete RNA gebildet wird. Die exakte Sequenz eines jeweiligen crRNA- oder tracrRNA-Moleküls ist charakteristisch für die Spezies, in der die RNA-Moleküle gefunden werden.

[0393] Der Begriff "Aktivator-RNA" wird vorliegend so verwendet, dass er ein tracrRNA-ähnliches Molekül eines auf DNA gerichteten RNA-Doppelmoleküls bezeichnet. Der Begriff "Targeter-RNA" wird vorliegend so verwendet, dass er ein crRNA-ähnliches Molekül eines auf DNA gerichteten RNA-Doppelmoleküls bezeichnet. Der Begriff "Duplex-bildendes Segment" wird vorliegend so verwendet, dass er den Abschnitt von Nukleotiden einer Aktivator-RNA oder einer Targeter-RNA bezeichnet, der zur Bildung des dsRNA-Duplex durch Hybridisierung an einen Abschnitt von Nukleotiden eines entsprechenden Aktivator-RNA- oder Targeter-RNA-Moleküls beiträgt. Mit anderen Worten umfasst eine Aktivator-RNA ein Duplex-bildendes Segment, das komplementär zu dem Duplex-bildenden Segment der entsprechenden Targeter-RNA ist. Somit umfasst eine Aktivator-RNA ein Duplex-bildendes Segment, während eine Targeter-RNA sowohl ein Duplex-bildendes Segment als auch

das DNA-Targeting-Segment einer auf DNA gerichteten RNA umfasst. Somit kann ein vorliegendes auf DNA gerichtetes RNA-Doppelmolekül aus einem beliebigen Paar aus korrespondierender Aktivator-RNA und Targeter-RNA bestehen.

[0394] Eine auf DNA gerichtete RNA aus zwei Molekülen kann so konzipiert sein, dass sie die kontrollierte (d. h. konditionale) Bindung einer Targeter-RNA mit einer Aktivator-RNA ermöglicht. Da eine auf DNA gerichtete RNA aus zwei Molekülen nicht funktionsfähig ist, wenn nicht sowohl die Aktivator-RNA als auch die Targeter-RNA in einem funktionalen Komplex mit dCas9 gebunden sind, kann eine auf DNA gerichtete RNA aus zwei Molekülen induzierbar sein (z. B. durch einen Wirkstoff induzierbar), indem die Bindung zwischen der Aktivator-RNA und der Targeter-RNA induzierbar ausgestaltet wird. Als ein nicht einschränkendes Beispiel können RNA-Aptamere verwendet werden, um die Bindung der Aktivator-RNA mit der Targeter-RNA zu regulieren (d. h. zu kontrollieren). Demgemäß kann die Aktivator-RNA und/oder die Targeter-RNA eine RNA-Aptamer-Sequenz umfassen.

[0395] RNA-Aptamere sind im Stand der Technik bekannt, und es handelt sich bei diesen im Allgemeinen um eine synthetische Version eines Riboswitch. Die Begriffe "RNA-Aptamer" und "Riboswitch" werden vorliegend austauschbar verwendet und umfassen sowohl synthetische als auch natürliche Nukleinsäuresequenzen, die für eine induzierbare Regulation der Struktur (und damit für die Verfügbarkeit von spezifischen Sequenzen) des RNA-Moleküls sorgen, von dem sie einen Teil bilden. RNA-Aptamere umfassen üblicherweise eine Sequenz, die sich in eine bestimmte Struktur faltet (z. B. eine Haarnadelschleife), welche spezifisch einen bestimmten Wirkstoff bindet (z. B. ein Small Molecule). Die Bindung des Wirkstoffs verursacht eine strukturelle Änderung in der Faltung der RNA, wobei dies ein Merkmal der Nukleinsäure, von der das Aptamer einen Teil bildet, ändert. Nicht einschränkende Beispiele: (i) eine Aktivator-RNA mit einem Aptamer könnte nicht in der Lage sein, die entsprechende Targeter-RNA zu binden, es sei denn, das Aptamer ist durch den entsprechenden Wirkstoff gebunden; (ii) eine Targeter-RNA mit einem Aptamer könnte nicht in der Lage sein, an die entsprechende Aktivator-RNA zu binden, es sei denn, das Aptamer wird durch den entsprechenden Wirkstoff gebunden; und (iii) eine Targeter-RNA und eine Aktivator-RNA, von denen jede ein unterschiedliches Aptamer umfasst, welches an einen unterschiedlichen Wirkstoff bindet, könnten nicht in der Lage sein, aneinander zu binden, sofern nicht beide Wirkstoffe vorhanden sind. Wie durch diese Beispiele veranschaulicht wird, kann eine auf DNA gerichtete RNA aus zwei Molekülen so konzipiert sein, dass sie induzierbar ist.

[0396] Beispiele für Aptamere und Riboswitch-Moleküle können beispielweise in Nakamura et al., *Genes Cells*. 2012 Mai; 17(5): 344–64; Vavalle et al., *Future Cardiol*. 2012 Mai; 8(3): 371–82; Citartan et al., *Biosens Bioelectron*. 2012 Apr 15; 34(1): 1–11; und Liberman et al., *Wiley Interdiscip Rev RNA*. 2012 Mai–Juni; 3(3): 369–84 gefunden werden, von denen alle vorliegend durch Bezugnahme in ihrer Gesamtheit eingeschlossen sind.

[0397] Nicht einschränkende Beispiele für Nukleotidsequenzen, die von einer auf DNA gerichteten RNA aus zwei Molekülen umfasst sein können, umfassen Targeter-RNAs (SEQ ID NOs: 566–567), die sich mit dem Duplex-bildenden Segment aus einer der Aktivator-RNAs, die in SEQ ID NOs: 671–678 gezeigt wird, paaren können.

[0398] Ein beispielhaftes auf DNA gerichtetes RNA-Einzelmolekül umfasst zwei komplementäre Abschnitte von Nukleotiden, die hybridisieren, wobei ein dsRNA-Duplex gebildet wird. In einigen Ausführungsformen ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder die DNA, die für diesen Abschnitt kodiert) mindestens etwa 60% identisch zu einer der Aktivator-RNA (tracrRNA)-Sequenzen, die in SEQ ID NOs: 431–562 gezeigt werden, über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden. Beispielsweise ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder der DNA, die für diesen Abschnitt kodiert) mindestens etwa 65% identisch, mindestens etwa 70% identisch, mindestens etwa 75% identisch, mindestens etwa 80% identisch, mindestens etwa 85% identisch, mindestens etwa 90% identisch, mindestens etwa 95% identisch, mindestens etwa 98% identisch, mindestens etwa 99% identisch oder 100% identisch zu einer der tracrRNA-Sequenzen, die in SEQ ID NOs: 431–562 gezeigt werden, über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden.

[0399] In einigen Ausführungsformen ist einer der beiden komplementären Abschnitten von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder der DNA, die für diesen Abschnitt kodiert), mindestens etwa 60% identisch zu einer der Targeter-RNA(crRNA)-Sequenzen, die in SEQ ID NOs: 563–679 gezeigt werden, über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden. Beispielsweise ist einer der beiden komplementären Abschnitte von Nukleotiden des auf DNA gerichteten RNA-Einzelmoleküls (oder der DNA,

die für diesen Abschnitt kodiert) mindestens etwa 65% identisch, mindestens etwa 70% identisch, mindestens etwa 75% identisch, mindestens etwa 80% identisch, mindestens etwa 85% identisch, mindestens etwa 90% identisch, mindestens etwa 95% identisch, mindestens etwa 98% identisch, mindestens etwa 99% identisch oder 100% identisch zu einer der crRNA-Sequenzen, die in SEQ ID NOs: 563–679 gezeigt werden, über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden.

[0400] Wir oben dargestellt bezeichnet eine "Wirtszelle" wie vorliegend verwendet eine in vivo oder in vitro vorliegende eukaryontische Zelle, eine prokaryontische Zelle (z. B. eine bakterielle Zelle oder eine Archaeen-Zelle) oder eine Zelle aus einem multizellulären Organismus (z. B. eine Zelllinie), die als eine unizelluläre Einheit kultiviert wird, wobei eukaryontische und prokaryontische Zellen als Rezipienten für Nukleinsäure verwendet werden können oder für diesen Zweck verwendet wurden und Nachkommen der ursprünglichen Zelle umfassen, die mit der Nukleinsäure transformiert wurde. Es sollte verstanden werden, dass die Nachkommen einer einzelnen Zelle aufgrund von natürlicher, zufälliger oder freiwilliger Mutation hinsichtlich der Morphologie oder des genomischen DNA- oder Gesamt-DNA-Komplements nicht notwendigerweise vollständig identisch zu der ursprünglichen parenteralen Zellen sein müssen. Eine "rekombinante Wirtszelle" (auch als "genetisch modifizierte Wirtszelle" bezeichnet) ist eine Wirtszelle, in die eine heterologe Nukleinsäure eingebracht wurde, z. B. ein Expressionsvektor. Beispielsweise wird eine vorliegende bakterielle Wirtszelle zu einer genetisch modifizierten Wirtszelle, indem man in eine geeignete bakterielle Wirtszelle eine exogene Nukleinsäure einbringt (z. B. ein Plasmid oder ein rekombinanter Expressionsvektor), und eine vorliegende eukaryontische Wirtszelle wird zu einer genetisch modifizierten eukaryontischen Wirtszelle (z. B. eine Säugetier-Keimzelle), indem man eine exogene Nukleinsäure in eine geeignete eukaryontische Wirtszelle einbringt.

[0401] Die in "Definitionen – Teil I" genannten Definitionen sind auch für den vorliegenden Abschnitt anwendbar; siehe "Definitionen – Teil I" für zusätzliche Erläuterungen zu Begriffen.

[0402] Bevor die vorliegende Erfindung weiter beschrieben wird, sollte verstanden werden, dass diese Erfindung nicht auf bestimmte beschriebene Ausführungsformen beschränkt ist, da diese selbstverständlich variieren können. Es sollte auch verstanden werden, dass die vorliegend verwendete Terminologie lediglich zum Zwecke der Beschreibung bestimmter Ausführungsformen verwendet wird und nicht einschränkend zu verstehen ist, da der Umfang der vorliegenden Erfindung lediglich durch die beiliegenden Ansprüche eingeschränkt wird.

[0403] Wenn ein Bereich von Werten angegeben wird, sollte verstanden werden, dass jeder zwischen der oberen und der unteren Grenze dieses Bereichs liegende Wert und jeder anderweitig genannte oder zwischenliegende Wert in dem genannten Bereich, sofern der Zusammenhang nicht eindeutig auf anderes verweist, bis auf das Zehntel der Einheit der unteren Grenze von der Erfindung eingeschlossen ist. Die oberen und unteren Grenzen dieser kleineren Bereiche können unabhängig voneinander in die kleinere Bereiche eingeschlossen werden und sind somit ebenfalls von der Erfindung umfasst, sofern nicht eine Grenze im genannten Bereich speziell ausgeschlossen ist. Wenn der genannte Bereich eine oder beide Grenzen einschließt, so sind Bereiche, die entweder den einen oder beide dieser eingeschlossenen Grenzen ausschließen, ebenfalls von der Erfindung eingeschlossen.

[0404] Sofern es nicht anderweitig definiert ist, haben alle technischen und wissenschaftlichen Bezeichnungen, die vorliegend verwendet werden, die gleiche Bedeutung wie sie üblicherweise vom Fachmann verstanden wird, an den sich diese Erfindung richtet. Obwohl bei der Ausführung oder Überprüfung der vorliegenden Erfindung beliebige Verfahren und Materialien, die den vorliegend beschriebenen Verfahren und Materialien ähneln oder äquivalent zu diesen sind, verwendet werden können, werden nunmehr bevorzugte Verfahren und Materialien beschrieben. Alle Publikationen, die vorliegend genannt werden, sind durch Bezugnahme eingeschlossen und offenbaren und beschreiben die Verfahren und/oder Materialien im Zusammenhang mit denen die Publikationen genannt werden.

[0405] Es muss beachtet werden, dass wie vorliegend und in den anliegenden Ansprüchen verwendet, die Singularformen "ein/eine/einer" und "der/die/das" die Pluralformen einschließen, sofern der Zusammenhang nicht eindeutig auf etwas anderes verweist. So umfasst beispielsweise die Bezugnahme auf "ein enzymatisch inaktives Cas9-Polypeptid" eine Vielzahl solcher Polypeptide, und die Bezugnahme auf "die Ziel-Nukleinsäure" umfasst die Bezugnahme auf eine oder mehrere Ziel-Nukleinsäuren und Äquivalente derselben, die dem Fachmann bekannt sind, usw. Es muss ferner beachtet werden, dass die Ansprüche so abgefasst sein können, dass sie ein beliebiges optionales Element ausschließen. Somit soll diese Anmerkung im Zusammenhang mit der Nennung von Anspruchselementen als vorhergehende Basis für die Verwendung einer ausschließenden Terminologie, wie z. B. "alleinig", "nur" und ähnlichem, oder zur Verwendung als "negative" Beschränkung dienen.

[0406] Es wird anerkannt, dass bestimmte Merkmale der Erfindung, die aus Gründen der Klarheit im Zusammenhang mit getrennten Ausführungsformen beschrieben werden, auch in Kombination in einer einzelnen Ausführungsform bereitgestellt werden können. Umgekehrt können verschiedene Merkmale der Erfindung, die aus Gründen der Kürze im Zusammenhang mit einer einzelnen Ausführungsform beschrieben werden, auch getrennt voneinander in einer geeigneten Unterkombination bereitgestellt werden. Alle Kombinationen von Ausführungsformen, die sich auf die Erfindung beziehen, sind spezifisch von der vorliegenden Erfindung eingeschlossen und werden vorliegend offenbart, als wäre jede Kombination individuell und explizit offenbart. Darüber hinaus sind alle Unterkombinationen der verschiedenen Ausführungsformen und der Elemente davon ebenfalls spezifisch von der vorliegenden Erfindung eingeschlossen und werden vorliegend offenbart, als wäre jede dieser Unterkombinationen individuell und explizit offenbart.

[0407] Die vorliegend besprochenen Veröffentlichungen werden lediglich bereitgestellt, um diese vor dem Anmeldetag der vorliegenden Anmeldung zu offenbaren. Nichts darf vorliegend als Zugeständnis verstanden werden, dass die vorliegende Erfindung nicht berechtigt wäre, einer solchen Veröffentlichung aufgrund einer früheren Erfindung vorzuziehen. Darüber hinaus können sich die Datumsangaben der vorliegend erwähnten Veröffentlichungen von den tatsächlichen Veröffentlichungsdaten unterscheiden, was im Einzelfall bestätigt werden muss.

DETAILLIERTE BESCHREIBUNG – TEIL II

[0408] Die vorliegende Erfindung stellt Verfahren zur Modulierung der Transkription einer Ziel-Nukleinsäure in einer Wirtszelle bereit. Die Verfahren umfassen im Allgemeinen das Inkontaktbringen der Ziel-Nukleinsäure mit einem enzymatisch inaktiven Cas9-Polypeptid und einer single-guide RNA. Diese Verfahren sind bei einer Vielzahl von Anwendungen nützlich, die ebenfalls bereitgestellt werden.

[0409] Ein Verfahren zur Modulierung der Transkription gemäß der vorliegenden Offenbarung überwindet einige der Nachteile von Verfahren, die RNAi verwenden. Ein Verfahren zur Modulation der Transkription gemäß der vorliegenden Offenbarung findet Verwendung bei einer großen Vielzahl von Anwendungen, einschließlich bei Forschungsanwendungen, bei der Wirkstoffentwicklung (z. B. Hochdurchsatz-Screening), bei der Target-Validierung, bei industriellen Anwendungen (z. B. die Herstellung von Nutzpflanzen, mikrobielles Engineering, usw.), bei diagnostischen Anwendungen, bei therapeutischen Anwendungen und bei bildgebenden Verfahren.

VERFAHREN ZUR MODULATION DER TRANSKRIPTION

[0410] Die vorliegende Offenbarung stellt ein Verfahren zur selektiven Modulation der Transkription einer Ziel-DNA in einer Wirtszelle bereit. Das Verfahren umfasst im Allgemeinen: (a) das Einbringen in die Wirtszelle von: (i) einer auf DNA gerichteten RNA oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert; und (ii) einem ortsspezifischen Varianten-Cas9-Polypeptid ("Varianten-Cas9-Polypeptid"), oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für das Varianten-Cas9-Polypeptid kodiert, wobei das Varianten-Cas9-Polypeptid eine verringerte Endodesoxyribonuklease-Aktivität aufweist.

[0411] Die auf DNA gerichtete RNA (vorliegend auch als "crRNA" oder "guide RNA" oder gRNA bezeichnet) umfasst: (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Ziel-Sequenz in einer Ziel-DNA ist; (ii) ein zweites Segment, das mit einem ortsspezifischen Polypeptid interagiert; und (iii) einen transkriptionalen Terminator. Das erste Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Ziel-Sequenz in einer Ziel-DNA ist, wird vorliegend als "Targeting-Segment" bezeichnet. Das zweite Segment, das mit einem ortsspezifischen Polypeptid interagiert, wird vorliegend auch als "Protein-bindende Sequenz" oder "dCas9-bindende Haarnadelschleife" oder "dCas9-Handle" bezeichnet. Mit "Segment" ist ein(e) Segment/Bereich/Region eines Moleküls gemeint, z. B. ein zusammenhängender Abschnitt von Nukleotiden in einer RNA. Die Definition von "Segment" ist soweit es nicht anderweitig in einem bestimmten Zusammenhang definiert ist, nicht auf eine bestimmte Anzahl von Gesamtbasenpaaren beschränkt, und kann Regionen von RNA-Molekülen einschließen, die eine beliebige Länge aufweisen und Regionen mit Komplementarität zu anderen Molekülen einschließen können oder nicht. Eine auf DNA gerichtete RNA gemäß der vorliegenden Erfindung kann ein RNA-Einzelmolekül sein (RNA-Einzelpolynukleotid), welches vorliegend als ein "auf DNA gerichtetes RNA-Einzelmolekül", eine "single guide RNA" oder eine "sgRNA" bezeichnet werden kann. Eine auf DNA gerichtete RNA gemäß der vorliegenden Offenbarung kann zwei RNA-Moleküle umfassen. Der Begriff "auf DNA gerichtete RNA" oder "gRNA" ist umfassend und schließt sowohl auf DNA gerichtete RNAs aus zwei Molekülen als auch auf DNA gerichtete RNA-Einzelmoleküle (d. h. sgRNA) ein.

[0412] Das ortsspezifische Varianten-Cas9-Polypeptid umfasst: (i) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und (ii) einen Aktivitätsteil, der eine verringerte Endodesoxyribonuklease-Aktivität aufweist.

[0413] Die auf DNA gerichtete RNA und das Varianten-Cas9-Polypeptid bilden einen Komplex in der Wirtszelle; der Komplex moduliert selektiv die Transkription einer Ziel-DNA in der Wirtszelle.

[0414] In einigen Fällen sorgt ein Verfahren zur Modulation der Transkription gemäß der vorliegenden Erfindung für eine selektive Modulation (z. B. Verringerung oder Erhöhung) einer Ziel-Nukleinsäure in einer Wirtszelle. Beispielsweise verringert die "selektive" Reduktion der Transkription einer Ziel-Nukleinsäure die Transkription der Ziel-Nukleinsäure um mindestens etwa 10%, mindestens etwa 20%, mindestens etwa 30%, mindestens etwa 40%, mindestens etwa 50%, mindestens etwa 60%, mindestens etwa 70%, mindestens etwa 80%, mindestens etwa 90% oder mehr als 90% im Vergleich zum Ausmaß der Transkription der Ziel-Nukleinsäure in Abwesenheit eines Komplexes aus einer auf DNA gerichteten RNA und eines Varianten-Cas9-Polypeptids. Die selektive Verringerung der Transkription einer Ziel-Nukleinsäure verringert die Transkription der Ziel-Nukleinsäure, verringert jedoch nicht wesentlich die Transkription einer Nicht-Ziel-Nukleinsäure; so wird z. B. die Transkription einer Nicht-Ziel-Nukleinsäure im Vergleich zum Ausmaß der Transkription der Nicht-Ziel-Nukleinsäure in Abwesenheit des Komplexes aus einer auf DNA gerichteten RNA und eines Varianten-Cas9-Polypeptids weniger als 10% (wenn überhaupt) verringert.

Erhöhte Transkription

[0415] Eine "selektiv" erhöhte Transkription einer Ziel-DNA kann die Transkription einer Ziel-DNA um mindestens etwa 1,1-fach (z. B. mindestens etwa 1,2-fach, mindestens etwa 1,3-fach, mindestens etwa 1,4-fach, mindestens etwa 1,5-fach, mindestens etwa 1,6-fach, mindestens etwa 1,7-fach, mindestens etwa 1,8-fach, mindestens etwa 1,9-fach, mindestens etwa 2-fach, mindestens etwa 2,5-fach, mindestens etwa 3-fach, mindestens etwa 3,5-fach, mindestens etwa 4-fach, mindestens etwa 4,5-fach, mindestens etwa 5-fach, mindestens etwa 6-fach, mindestens etwa 7-fach, mindestens etwa 8-fach, mindestens etwa 9-fach, mindestens etwa 10-fach, mindestens etwa 12-fach, mindestens etwa 15-fach, oder mindestens etwa 20-fach) im Vergleich zum Ausmaß der Transkription der Ziel-DNA in Abwesenheit eines Komplexes aus einer auf DNA gerichtete RNA und einem Varianten-Cas9-Polypeptid. Die selektive Erhöhung der Transkription einer Ziel-DNA erhöht die Transkription der Ziel-DNA, verringert jedoch nicht wesentlich die Transkription einer Nicht-Ziel-DNA; so wird z. B. die Transkription einer Nicht-Ziel-DNA im Vergleich zum Ausmaß der Transkription der Nicht-Ziel-DNA in Abwesenheit des Komplexes aus einer auf DNA gerichtete RNA und einem Varianten-Cas9-Polypeptid (wenn überhaupt) weniger als etwa 5-fach (z. B. weniger als etwa 4-fach, weniger als etwa 3-fach, weniger als etwa 2-fach, weniger als etwa 1,8-fach, weniger als etwa 1,6-fach, weniger als etwa 1,4-fach, weniger als etwa 1,2-fach oder weniger als 1,1-fach) erhöht.

[0416] Als nicht einschränkendes Merkmal kann eine Erhöhung erreicht werden, indem man dCas9 an eine heterologe Sequenz fusioniert. Geeignete Fusionspartner umfassen (sind jedoch nicht beschränkt auf) ein Polypeptid, das eine Aktivität bereitstellt, die indirekt die Transkription erhöht, indem es direkt auf die Ziel-DNA oder auf ein Polypeptid (z. B. ein Histon oder ein anderes DNA-bindendes Protein), das mit der Ziel-DNA assoziiert ist, wirkt. Geeignete Fusionspartner umfassen (sind jedoch nicht beschränkt auf) ein Polypeptid, das für Methyltransferase-Aktivität, Demethylase-Aktivität, Acetyltransferase-Aktivität, Deacetylase-Aktivität, Kinase-Aktivität, Phosphatase-Aktivität, Ubiquitin-Ligase-Aktivität, deubiquitinierende Aktivität, Adenylierungs-Aktivität, Deadenylierungs-Aktivität, SUMOylierende Aktivität, deSUMOylierende Aktivität, Ribosylierungs-Aktivität, Deribosylierungs-Aktivität, Myristoylierungs-Aktivität oder Demyristoylierungs-Aktivität.

[0417] Zusätzliche geeignete Fusionspartner umfassen (sind jedoch nicht beschränkt auf) ein Polypeptid, das direkt für eine erhöhte Transkription der Ziel-Nukleinsäure sorgt (z. B. ein Transkriptions-Aktivator oder ein Fragment davon, ein Protein oder ein Fragment davon, das einen Transkriptions-Aktivator rekrutiert, ein Transkriptionsregulator, der auf ein Small Molecule/einen Wirkstoff reagiert, usw.).

[0418] Ein nicht-einschränkendes Beispiel für ein vorliegendes Verfahren, das ein dCas9-Fusionsprotein zur Erhöhung der Transkription in einem Prokaryonten verwendet, umfasst eine Modifikation des bakteriellen Ein-Hybrid(B1H)-Systems oder Zwei-Hybrid(B2H)-Systems. Im B1H-System wird eine DNA-bindende Domäne (BD) an eine bakterielle Transkriptions-Aktivator-Domäne (AD, z. B. an die alpha-Untereinheit der RNA-Polymerase von Escherichia coli (RNAP α)) fusioniert. Somit kann eine vorliegende dCas9 an eine heterologe Sequenz fusioniert werden, die eine AD umfasst. Wenn das vorliegende dCas9-Fusionsprotein an der Region stromaufwärts des Promotors ankommt (dorthin geführt durch die auf DNA gerichtete RNA), dann rekrutiert die

AD (z. B. RNAP α) des dCas9-Fusionsproteins das RNAP-Holoenzym, was zur Transkriptionsaktivierung führt. Im B2H-System ist die BD nicht direkt an die AD fusioniert; stattdessen wird deren Interaktion durch eine Protein-Protein-Interaktion (z. B. GAL11P-GAL4-Interaktion) vermittelt. Um ein solches System zur Verwendung in den vorliegenden Verfahren zu modifizieren, kann dCas9 an eine erste Proteinsequenz fusioniert werden, die für eine Protein-Protein-Interaktion sorgt (z. B. das GAL11P und/oder GAL4-Protein der Hefe), und RNA α kann an eine zweite Proteinsequenz fusioniert werden, welche die Protein-Protein-Interaktion komplettiert (z. B. GAL4 wenn GAL11P an dCas9 fusioniert ist, GAL11P, wenn GAL4 an dCas9 fusioniert ist, usw.). Die Bindungsaffinität zwischen GAL11P und GAL4 erhöht die Effizienz der Bindung und die Transkriptionsrate.

[0419] Ein nicht einschränkendes Beispiel für ein vorliegendes Verfahren, das ein dCas9-Fusionsprotein zur Erhöhung der Transkription in einem Eukaryonten verwendet, umfasst die Fusion von dCas9 an eine Aktivierungsdomäne (AD) (z. B. GAL4, Herpesvirus-Aktivierungsprotein VP16 oder VP64, humaner nuklearer Faktor NF- κ B p65-Untereinheit, usw.). Um das System induzierbar zu machen, kann die Expression des dCas9-Fusionsproteins durch einen induzierbaren Promotor kontrolliert werden (z. B. Tet-AN, Tet-AUS, usw.). Die auf DNA gerichtete RNA kann so beschaffen sein, dass sie auf bekannte Transkriptions-Response-Elemente (z. B. Promotoren, Enhancer, usw.), bekannte stromaufwärts gelegene Aktivierungssequenzen (upstream activating sequences, UAS), Sequenzen, von unbekannter oder bekannter Funktion, von denen angenommen wird, dass sie in der Lage sind, die Expression der Ziel-DNA zu kontrollieren, usw., gerichtet ist.

Zusätzliche Fusionspartner

[0420] Nicht einschränkende Beispiele für Fusionspartner, die für eine erhöhte oder verringerte Transkription sorgen, sind in **Fig. 54** gezeigt und umfassen Transkriptions-Aktivator- und Transkriptions-Repressor-Domänen (z. B. die Krüppel-assoziierte Box (KRAB oder SKD); die Mad mSIN3-Interaktionsdomäne (SID); die ERF-Repressor-Domäne (ERD), usw.). In einigen Fällen wird das dCas9-Fusionsprotein von der auf DNA gerichteten RNA zu einem spezifischen Ort (einer Sequenz) in der Ziel-DNA geführt und übt eine Locus-spezifische Regulation aus, wie beispielsweise die Blockierung der Bindung der RNA-Polymerase an einen Promotor (was selektiv die Transkriptions-Aktivator-Funktion hemmt), und/oder die Modifikation des lokalen Chromatinzustands (z. B. wenn eine Fusionssequenz verwendet wird, welche die Ziel-DNA oder ein mit der Ziel-DNA modifiziertes Polypeptid modifiziert). In einigen Fällen sind die Veränderungen transient (z. B. Repression oder Aktivierung der Transkription). In einigen Fällen sind die Veränderungen vererbbar (z. B. wenn epigenetische Modifikationen an der Ziel-DNA oder an Proteinen, die mit der Ziel-DNA assoziiert sind, vorgenommen werden, z. B. nukleosomale Histone).

[0421] In einigen Ausführungsformen kann die heterologe Sequenz an den C-Terminus des dCas9-Polypeptids fusioniert sein. In einigen Ausführungsformen kann die heterologe Sequenz an den N-Terminus des dCas9-Polypeptids fusioniert sein. In einigen Ausführungsformen kann die heterologe Sequenz an einen internen Teil (d. h. einen Teil, bei dem es sich nicht um den N- oder C-Terminus handelt) des dCas9-Polypeptids fusioniert sein.

[0422] Die biologischen Wirkungen eines Verfahrens, welches ein wie vorliegend beschriebenes dCas9-Fusionsprotein verwendet, kann durch jedes handhabbare Verfahren nachgewiesen werden (z. B. Gen-Expressions-Assays; auf Chromatin basierende Assays, z. B. Chromatin-Immunpräzipitation (ChIP), Chromatin in vivo Assay (CiA), usw.; und ähnliches).

[0423] In einigen Fällen umfasst ein vorliegendes Verfahren die Verwendung von zwei oder mehreren verschiedenen auf DNA gerichteten RNAs. Beispielsweise können zwei verschiedene auf DNA gerichtete RNAs in einer einzelnen Wirtszelle verwendet werden, wobei die beiden verschiedenen auf DNA gerichteten RNAs auf zwei unterschiedliche Ziel-Sequenzen in derselben Ziel-Nukleinsäure gerichtet sind.

[0424] Somit kann beispielsweise ein vorliegendes Verfahren zur transkriptionalen Modulation ferner die Einbringung einer zweiten auf DNA gerichteten RNA oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für die zweite auf DNA gerichtete RNA kodiert, in die Wirtszelle umfassen, wobei die zweite auf DNA gerichtete RNA Folgendes umfasst: (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer zweiten Ziel-Sequenz in der Ziel-DNA ist; (ii) ein zweites Segment, das mit dem ortsspezifischen Polypeptid interagiert, und (iii) einen transkriptionalen Terminator. In einigen Fällen sorgt die Verwendung von zwei unterschiedlichen auf DNA gerichteten RNAs, die auf zwei unterschiedliche Ziel-Sequenzen in der gleichen Ziel-Nukleinsäure gerichtet sind, für eine erhöhte Modulation (z. B. Erhöhung oder Verringerung) der Transkription der Ziel-Nukleinsäure.

[0425] Als weiteres Beispiel können zwei unterschiedliche auf DNA gerichtete RNAs in einer einzelnen Wirtszelle verwendet werden, wobei die beiden unterschiedlichen auf DNA gerichteten RNAs auf zwei unterschiedliche Ziel-Nukleinsäuren gerichtet sind. Somit kann ein vorliegendes Verfahren zur transkriptionalen Modulation beispielsweise ferner die Einbringung einer zweiten auf DNA gerichteten RNA oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für die zweite auf DNA gerichtete RNA kodiert, in die Wirtszelle umfassen, wobei die zweite auf DNA gerichtete RNA Folgendes umfasst: (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Ziel-Sequenz in mindestens einer zweiten Ziel-DNA ist; (ii) ein zweites Segment, das mit dem ortsspezifischen Polypeptid interagiert, und (iii) einen transkriptionalen Terminator.

[0426] In einigen Ausführungsformen umfasst eine vorliegende Nukleinsäure (z. B. eine auf DNA gerichtete RNA; z. B. ein auf DNA gerichtetes RNA-Einzelmolekül, einer Aktivator-RNA, eine Targeter-RNA, usw.; ein Donor-Polynukleotid; eine Nukleinsäure, die für ein ortsspezifisch modifizierendes Polypeptid kodiert; usw.) eine Modifikation oder eine Sequenz, die für ein zusätzliches gewünschtes Merkmal sorgt (z. B. modifizierte oder regulierte Stabilität; subzelluläres Targeting; Tracking, z. B. eine fluoreszente Markierung; eine Bindungsstelle für ein Protein oder für einen Proteinkomplex; usw.). Nicht einschränkende Beispiele umfassen eine 5'-Kappe (z. B. eine 7-Methylguanylat-Kappe (m⁷G)); einen 3'-polyadenylierten Anhang (d. h. einen 3'-Poly(A)-Anhang); eine Riboswitch-Sequenz oder eine Aptamer-Sequenz (z. B. um eine regulierte Stabilität und/oder eine regulierte Zugänglichkeit durch Proteine und/oder Protein-Komplexe zu ermöglichen); eine Terminator-Sequenz; eine Sequenz, die einen dsRNA-Duplex bildet (z. B. eine Haarnadelschleife); eine Modifikation oder eine Sequenz, welche die RNA gezielt zu einer subzellulären Stelle führt (z. B. Nukleus, Mitochondrien, Chloroplasten, und ähnliches); eine Modifikation oder eine Sequenz, die für ein Tracking sorgt (z. B. direkte Konjugation an ein fluoreszentes Molekül, Konjugation an einen Rest, der eine fluoreszente Detektion ermöglicht, eine Sequenz, die eine fluoreszente Detektion ermöglicht, usw.); eine Modifikation oder eine Sequenz, die eine Bindungsstelle für Proteine bereitstellt (z. B. Proteine, die auf die DNA wirken, einschließlich Transkriptions-Aktivatoren, Transkriptions-Repressoren, DNA-Methyltransferasen, DNA-Demethylasen, Histon-Acetyltransferasen, Histon-Deacetylasen, und ähnliches) und Kombinationen derselben. Auf DNA gerichtetes Segment

[0427] Das auf DNA gerichtete Segment (oder "die auf DNA gerichtete Sequenz") einer auf DNA gerichteten RNA ("crRNA") umfasst eine Nukleotidsequenz, die komplementär zu einer spezifischen Sequenz innerhalb der Ziel-DNA (dem komplementären Strang der Ziel-DNA) ist.

[0428] Mit anderen Worten interagiert das auf DNA gerichtete Segment einer vorliegenden auf DNA gerichteten RNA mit einer Ziel-DNA in einer Sequenz-spezifischen Weise über Hybridisierung (d. h. Basenpaarung). Somit kann die Nukleotidsequenz des auf DNA gerichteten Segments variieren und bestimmt die Stelle innerhalb der Ziel-DNA, an der die auf DNA gerichtete RNA und die Ziel-DNA interagieren werden. Das auf DNA gerichtete Segment einer vorliegenden auf DNA gerichteten RNA kann so modifiziert werden (z. B. durch genetische Modifikation), dass diese an jede gewünschte Sequenz innerhalb einer Ziel-DNA bindet.

[0429] Das auf DNA gerichtete Segment kann eine Länge von etwa 12 Nukleotiden bis etwa 100 Nukleotiden aufweisen. Beispielsweise kann das auf DNA gerichtete Segment eine Länge von etwa 12 Nukleotiden (nt) bis etwa 80 nt, von etwa 12 nt bis etwa 50 nt, von etwa 12 nt bis etwa 40 nt, von etwa 12 nt bis etwa 30 nt, von etwa 12 nt bis etwa 25 nt, von etwa 12 nt bis etwa 20 nt, oder von etwa 12 nt bis etwa 19 nt aufweisen. Beispielsweise kann das auf DNA gerichtete Segment eine Länge von etwa 19 nt bis etwa 20 nt, von etwa 19 nt bis etwa 25 nt, von etwa 19 nt bis etwa 30 nt, von etwa 19 nt bis etwa 35 nt, von etwa 19 nt bis etwa 40 nt, von etwa 19 nt bis etwa 45 nt, von etwa 19 nt bis etwa 50 nt, von etwa 19 nt bis etwa 60 nt, von etwa 19 nt bis etwa 70 nt, von etwa 19 nt bis etwa 80 nt, von etwa 19 nt bis etwa 90 nt, von etwa 19 nt bis etwa 100 nt, von etwa 20 nt bis etwa 25 nt, von etwa 20 nt bis etwa 30 nt, von etwa 20 nt bis etwa 35 nt, von etwa 20 nt bis etwa 40 nt, von etwa 20 nt bis etwa 45 nt, von etwa 20 nt bis etwa 50 nt, von etwa 20 nt bis etwa 60 nt, von etwa 20 nt bis etwa 70 nt, von etwa 20 nt bis etwa 80 nt, von etwa 20 nt bis etwa 90 nt, von etwa 20 nt bis etwa 100 nt.

[0430] Die Nukleotidsequenz (die auf DNA gerichtete Sequenz) des auf DNA gerichteten Segments, die komplementär zu einer Nukleotidsequenz (Ziel-Sequenz) der Ziel-DNA ist, kann eine Länge von mindestens etwa 12 nt haben. Beispielsweise kann die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, das komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, eine Länge von mindestens etwa 12 nt haben, von mindestens etwa 15 nt, von mindestens etwa 18 nt, von mindestens etwa 19 nt, von mindestens etwa 20 nt, von mindestens etwa 25 nt, von mindestens etwa 30 nt, von mindestens etwa 35 nt oder von mindestens etwa 40 nt haben. Beispielsweise kann die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, eine Länge von etwa 12 Nukleotiden (nt) bis etwa 80 nt, von etwa 12 nt bis etwa 50 nt, von etwa 12 nt bis etwa 45 nt, von etwa 12 nt bis etwa 40 nt, von etwa 12 nt bis

etwa 35 nt, von etwa 12 nt bis etwa 30 nt, von etwa 12 nt bis etwa 25 nt, von etwa 12 nt bis etwa 20 nt, von etwa 12 nt bis etwa 19 nt, von etwa 19 nt bis etwa 20 nt, von etwa 19 nt bis etwa 25 nt, von etwa 19 nt bis etwa 30 nt, von etwa 19 nt bis etwa 35 nt, von etwa 19 nt bis etwa 40 nt, von etwa 19 nt bis etwa 45 nt, von etwa 19 nt bis etwa 50 nt, oder von etwa 19 nt bis etwa 60 nt, von etwa 20 nt bis etwa 25 nt, von etwa 20 nt bis etwa 30 nt, von etwa 20 nt bis etwa 35 nt, von etwa 20 nt bis etwa 40 nt, von etwa 20 nt bis etwa 45 nt, von etwa 20 nt bis etwa 50 nt, oder von etwa 20 nt bis etwa 60 nt aufweisen. Die Nukleotidsequenz (die auf DNA gerichtete Sequenz) des auf DNA gerichteten Segments, die komplementär zu einer Nukleinsäuresequenz (Ziel-Sequenz) der Ziel-DNA ist, kann eine Länge von mindestens etwa 12 nt aufweisen.

[0431] In einigen Fällen ist die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, etwa 20 Nukleotide lang. In einigen Fällen ist die auf DNA gerichtete Sequenz des auf DNA gerichteten Segments, die komplementär zu einer Ziel-Sequenz der Ziel-DNA ist, 19 Nukleotide lang.

[0432] Die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA kann mindestens 60% betragen (z. B. mindestens 65%, mindestens 70%, mindestens 75%, mindestens 80%, mindestens 85%, mindestens 90%, mindestens 95%, mindestens 97%, mindestens 98%, mindestens 99% oder 100%). In einigen Fällen kann die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA 100% über die 7 zusammenhängenden Nukleotide am 5'-Ende der Ziel-Sequenz des komplementären Strangs der Ziel-DNA betragen. In einigen Fällen kann die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA mindestens 60% über etwa 20 zusammenhängende Nukleotide betragen. In einigen Fällen kann die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA 100% über die 14 zusammenhängenden Nukleotide am 5'-Ende der Ziel-Sequenz des komplementären Strangs der Ziel-DNA betragen und lediglich 0% über die restliche Sequenz. In einem solchen Fall kann die auf DNA gerichtete Sequenz als 14 Nukleotide lang angesehen werden. In einigen Fällen kann die prozentuale Komplementarität zwischen der auf DNA gerichteten Sequenz des auf DNA gerichteten Segments und der Ziel-Sequenz der Ziel-DNA 100% über die 7 zusammenhängenden Nukleotide am 5'-Ende der Ziel-Sequenz des komplementären Strangs der Ziel-DNA betragen und lediglich 0% über die restliche Sequenz. In einem solchen Fall kann die auf DNA gerichtete Sequenz als 7 Nukleotide lang betrachtet werden.

Protein-bindendes Segment

[0433] Das Protein-bindende Segment (d. h. die "Protein-bindende Sequenz") einer auf DNA gerichteten RNA interagiert mit einer ortsspezifischen Polypeptid-Variante. Wenn die Variante des ortsspezifischen Cas9-Polypeptids gemeinsam mit der auf DNA gerichteten RNA an eine Ziel-DNA bindet, wird die Transkription der Ziel-DNA verringert.

[0434] Das Protein-bindende Segment der auf DNA gerichteten RNA umfasst zwei komplementäre Abschnitte von Nukleotiden, die miteinander hybridisieren, um einen doppelsträngigen DNA-Duplex zu bilden (dsRNA-Duplex).

[0435] Das Protein-bindende Segment einer auf DNA gerichteten RNA der vorliegenden Offenbarung umfasst zwei Abschnitte von Nukleotiden (eine Targeter-DNA und eine Aktivator-RNA), die komplementär zueinander sind, kovalent miteinander durch zwischenliegende Nukleotide verknüpft sind (z. B. im Falle eines auf DNA gerichteten RNA-Einzelmoleküls) ("Linker" oder "Linker-Nukleotide") und hybridisieren, wobei der doppelsträngige RNA-Duplex (dsRNA-Duplex oder "dCas9-bindende Haarnadelschleife") des Protein-bindenden Segments gebildet wird, was zu einer Haarnadelschleifen-Struktur führt. Diese Haarnadelschleifen-Struktur ist schematisch in **Fig. 39A** gezeigt. Die Targeter-RNA und die Aktivator-RNA können kovalent über das 3'-Ende der Targeter-RNA und das 5'-Ende der Aktivator-RNA verknüpft sein. Alternative dazu können die Targeter-RNA und die Aktivator-RNA kovalent über das 5'-Ende der Targeter-RNA und das 3'-Ende der Aktivator-RNA kovalent verknüpft sein.

[0436] Das Protein-bindende Segment kann eine Länge von etwa 10 Nukleotiden bis etwa 100 Nukleotiden aufweisen, z. B. von etwa 10 Nukleotiden (nt) bis etwa 20 nt, von etwa 20 nt bis etwa 30 nt, von etwa 30 nt bis etwa 40 nt, von etwa 40 nt bis etwa 50 nt, von etwa 50 nt bis etwa 60 nt, von etwa 60 nt bis etwa 70 nt, von etwa 70 nt bis etwa 80 nt, von etwa 80 nt bis etwa 90 nt oder von etwa 90 bis etwa 100 nt. Beispielsweise kann das

Protein-bindende Segment eine Länge von etwa 15 Nukleotiden (nt) bis etwa 80 nt, von etwa 15 nt bis etwa 50 nt, von etwa 15 nt bis etwa 40 nt, von etwa 15 nt bis etwa 30 nt oder von etwa 15 bis etwa 25 nt aufweisen.

[0437] Der dsRNA-Duplex des Protein-bindenden Segments kann eine Länge von etwa 6 Basenpaaren (bp) bis etwa 50 bp aufweisen. Beispielsweise kann der dsRNA-Duplex des Protein-bindenden Segments eine Länge von etwa 6 bp bis etwa 40 bp, von etwa 6 bp bis etwa 30 bp, von etwa 6 bp bis etwa 25 bp, von etwa 6 bp bis etwa 20 bp, von etwa 6 bp bis etwa 15 bp, von etwa 8 bp bis etwa 40 bp, von etwa 8 bp bis etwa 30 bp, von etwa 8 bp bis etwa 25 bp, von etwa 8 bp bis etwa 20 bp oder von etwa 8 bp bis etwa 15 bp aufweisen. Beispielsweise kann der dsRNA-Duplex des Protein-bindenden Segments eine Länge von 8 bp bis etwa 10 bp, von etwa 10 bp bis etwa 15 bp, von etwa 15 bp bis etwa 18 bp, von etwa 18 bp bis etwa 20 bp, von etwa 20 bp bis etwa 25 bp, von etwa 25 bp bis etwa 30 bp, von etwa 30 bp bis etwa 35 bp, von etwa 35 bp bis etwa 40 bp oder von etwa 40 bp bis etwa 50 bp aufweisen. In einigen Ausführungsformen weist der dsRNA-Duplex des Protein-bindenden Segments eine Länge von 36 Basenpaaren auf. Die prozentuale Komplementarität zwischen den Nukleotidsequenzen, die hybridisieren, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden, kann mindestens etwa 60% betragen. Beispielsweise kann die prozentuale Komplementarität zwischen den Nukleotidsequenzen, die hybridisieren, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden mindestens etwa 65%, mindestens etwa 70%, mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 98%, oder mindestens etwa 99% betragen. In einigen Fällen beträgt die prozentuale Komplementarität zwischen den Nukleotidsequenzen, die hybridisieren, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden, 100%.

[0438] Der Linker kann eine Länge von etwa 3 Nukleotiden bis etwa 100 Nukleotiden aufweisen. Beispielsweise kann der Linker eine Länge von etwa 3 Nukleotiden (nt) bis etwa 90 nt, von etwa 3 Nukleotiden (nt) bis etwa 80 nt, von etwa 3 Nukleotiden (nt) bis etwa 70 nt, von etwa 3 Nukleotiden (nt) bis etwa 60 nt, von etwa 3 Nukleotiden (nt) bis etwa 50 nt, von etwa 3 Nukleotiden (nt) bis etwa 40 nt, von etwa 3 Nukleotiden (nt) bis etwa 30 nt, von etwa 3 Nukleotiden (nt) bis etwa 20 nt oder von etwa 3 Nukleotiden (nt) bis etwa 10 nt aufweisen. Beispielsweise kann der Linker eine Länge von etwa 3 nt bis etwa 5 nt, von etwa 5 nt bis etwa 10 nt, von etwa 10 nt bis etwa 15 nt, von etwa 15 nt bis etwa 20 nt, von etwa 20 nt bis etwa 25 nt, von etwa 25 nt bis etwa 30 nt, von etwa 30 nt bis etwa 35 nt, von etwa 35 nt bis etwa 40 nt, von etwa 40 nt bis etwa 50 nt, von etwa 50 nt bis etwa 60 nt, von etwa 60 nt bis etwa 70 nt, von etwa 70 nt bis etwa 80 nt, von etwa 80 nt bis etwa 90 nt oder von etwa 90 nt bis etwa 100 nt aufweisen. In einigen Ausführungsformen besteht der Linker einer auf DNA gerichteten RNA aus 4 nt.

[0439] Nicht einschränkende Beispiele für Nukleotidsequenzen, die in ein geeignetes Protein-bindendes Segment (d. h. dCas9-Handle) eingebracht werden können, sind in den SEQ ID NOs: 563–582 gezeigt (siehe die **Fig. 8** und **Fig. 9** für Beispiele).

[0440] In einigen Fällen umfasst ein geeignetes Protein-bindendes Segment eine Nukleotidsequenz, die in 1, 2, 3, 4 oder 5 Nukleotiden von einer der oben aufgeführten Sequenzen abweicht.

Stabilitäts-Kontrollsequenz (z. B. transkriptionales Terminator-Sequenz)

[0441] Eine Stabilitäts-Kontrollsequenz beeinflusst die Stabilität einer RNA (z. B. einer auf DNA gerichteten RNA, einer Targeter-RNA, einer Aktivator-RNA, usw.). Ein Beispiel für eine geeignete Stabilitäts-Kontrollsequenz ist ein transkriptionales Terminator-Segment (d. h. eine Transkriptions-Terminations-Sequenz). Ein transkriptionales Terminator-Segment einer vorliegenden auf DNA gerichteten RNA kann eine Gesamtlänge von etwa 10 Nukleotiden bis etwa 100 Nukleotiden aufweisen, z. B. von etwa 10 Nukleotiden (nt) bis etwa 20 nt, von etwa 20 nt bis etwa 30 nt, von etwa 30 nt bis etwa 40 nt, von etwa 40 nt bis etwa 50 nt, von etwa 50 nt bis etwa 60 nt, von etwa 60 nt bis etwa 70 nt, von etwa 70 nt bis etwa 80 nt, von etwa 80 nt bis etwa 90 nt oder von etwa 90 nt bis etwa 100 nt. Beispielsweise kann das transkriptionale Terminator-Segment eine Länge von etwa 15 Nukleotiden (nt) bis etwa 80 nt, von etwa 15 nt bis etwa 50 nt, von etwa 15 nt bis etwa 40 nt, von etwa 15 nt bis etwa 30 nt oder von etwa 15 nt bis etwa 25 nt.

[0442] In einigen Fällen ist die Transkriptions-Terminations-Sequenz eine solche, die in einer eukaryontischen Zelle funktionsfähig ist. In einigen Fällen ist die Transkriptions-Terminations-Sequenz eine solche, die in einer prokaryontischen Zelle funktionsfähig ist.

[0443] Nicht einschränkende Beispiele für Nukleotidsequenzen, die in eine Stabilitäts-Kontrollsequenz eingebracht werden können (z. B. in ein transkriptionales Terminations-Segment oder in ein beliebiges Segment der auf DNA gerichteten RNA, das für eine erhöhte Stabilität sorgt) umfassen die in SEQ ID NOs: 683–696 darge-

stellten Sequenzen, beispielsweise, 5'-UAAUCCCCACAGCCGCCAGUUCCGCUGGCGGCAUUUU-5' (SEQ ID NO: 795) (eine Rho-unabhängige trp-Terminationsstelle).

Zusätzliche Sequenzen

[0444] In einigen Ausführungsformen umfasst eine auf DNA gerichtete RNA mindestens ein zusätzliches Segment, entweder am 5'-Ende oder am 3'-Ende. Beispielsweise kann ein geeignetes zusätzliches Segment eine 5'-Kappe (z. B. eine 7-Methylguanylat-Kappe (m7G)); einen 3'-polyadenylierten Anhang (z. B. einen 3'-Poly(A)-Anhang); eine Riboswitch-Sequenz (z. B. um eine regulierte Stabilität und/oder eine regulierte Zugänglichkeit durch Proteine und/oder Protein-Komplexe zu ermöglichen); eine Sequenz, die einen dsRNA-Duplex bildet (z. B. eine Haarnadelschleife); eine Sequenz, welche die RNA gezielt zu einer subzellulären Stelle führt (z. B. Nukleus, Mitochondrien, Chloroplasten, und ähnliche); eine Modifikation oder eine Sequenz, die für ein Tracking sorgt (z. B. direkte Konjugation an ein fluoreszentes Molekül, Konjugation an einen Rest, der eine fluoreszente Detektion ermöglicht, eine Sequenz, die eine fluoreszente Detektion ermöglicht, usw.); eine Modifikation oder eine Sequenz, die eine Bindungsstelle für Proteine bereitstellt (z. B. Proteine, die auf die DNA wirken, einschließlich Transkriptions-Aktivatoren, Transkriptions-Repressoren, DNA-Methyltransferasen, DNA-Demethylasen, Histon-Acetyltransferasen, Histon-Deacetylase, und ähnliches) und Kombinationen derselben.

Mehrere auf DNA gerichtete RNAs gleichzeitig

[0445] In einigen Ausführungsformen werden mehrere auf DNA gerichtete RNAs gleichzeitig in derselben Zelle verwendet, um gleichzeitig die Transkription an verschiedenen Orten auf der gleichen Ziel-DNA oder auf unterschiedlichen Ziel-DNAs zu modulieren. In einigen Ausführungsformen sind zwei oder mehrere auf DNA gerichtete RNAs auf dasselbe Gen oder auf dasselbe Transkript oder auf denselben Locus gerichtet. In einigen Ausführungsformen sind zwei oder mehrere auf DNA gerichtete RNAs auf unterschiedliche, nicht verwandte Loci gerichtet. In einigen Ausführungsformen sind zwei oder mehrere der auf DNA gerichtete RNAs auf unterschiedliche, jedoch miteinander verwandte Loci gerichtet.

[0446] Weil die auf DNA gerichteten RNAs klein und robust sind, können sie gleichzeitig auf dem gleichen Expressionsvektor vorliegen, und sie können sogar unter der gleichen Transkriptionskontrolle stehen, sofern dies gewünscht ist. In einigen Ausführungsformen werden zwei oder mehrere (z. B. 3 oder mehr, 4 oder mehr, 5 oder mehr, 10 oder mehr, 15 oder mehr, 20 oder mehr, 25 oder mehr, 30 oder mehr, 35 oder mehr, 40 oder mehr, 45 oder mehr oder 50 oder mehr) auf DNA gerichtete RNAs gleichzeitig in einer Ziel-Zelle exprimiert (vom gleichen oder von unterschiedlichen Vektoren). Die exprimierten auf DNA gerichteten RNAs können in unterschiedlicher Weise von dCas9-Proteinen aus unterschiedlichen Bakterien, wie beispielsweise aus *S. pyogenes*, *S. thermophilus*, *L. innocua* und *N. meningitidis* erkannt werden.

[0447] Um verschiedene auf DNA gerichtete RNAs zu exprimieren, kann ein künstliches RNA-Prozessierungssystem, welches durch die Csy4-Endoribonuklease vermittelt wird, verwendet werden. Mehrere auf DNA gerichtete RNAs können in einem Tandem-Array auf einem Vorläufer-Transkript (z. B. exprimiert von einem U6-Promotor) zusammengeführt werden und durch eine Csy4-spezifische RNA-Sequenz getrennt werden. Das co-exprimierte Csy4-Protein spaltet das Vorläufer-Transkript in mehrere auf DNA gerichtete RNAs. Vorteile einer Verwendung eines RNA-Prozessierungssystems umfassen: erstens, es besteht keine Notwendigkeit für die Verwendung von mehreren Promotoren; zweitens, da alle auf DNA gerichteten RNAs von einem Vorläufer-Transkript prozessiert werden, sind deren Konzentrationen im Hinblick auf eine ähnliche dCas9-Bindung normalisiert.

[0448] Csy4 ist ein kleines Endoribonuklease(RNase)-Protein, das vom Bakterium *Pseudomonas aeruginosa* abgeleitet ist. Csy4 erkennt spezifisch eine minimale RNA-Haarnadelschleife aus 17 bp und zeigt eine schnelle (< 1 Min) und hocheffiziente (> 99,9%) RNA-Spaltung. Im Gegensatz zu den meisten anderen RNasen bleibt das gespaltene RNA-Fragment stabil und funktionell aktiv. Die auf Csy4 basierende RNA-Spaltung kann in einem künstlichen RNA-Prozessierungssystem verwendet werden. In diesem System werden die RNA-Haarnadelschleifen aus 17 bp zwischen mehreren RNA-Fragmenten inseriert, die als ein Vorläufer-Transkript ausgehend von einem einzelnen Promotor transkribiert werden. Die co-Expression von Csy4 ist bei der Erzeugung einzelner RNA-Fragmente wirksam.

Ortsspezifisches Polypeptid

[0449] Wie oben erwähnt bilden eine vorliegende auf DNA gerichtete RNA und eine Variante des ortsspezifischen Cas9-Polypeptids einen Komplex. Die auf DNA gerichtete RNA sorgt für die Ziel-Spezifität des Komplexes, indem sie eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz der Ziel-DNA ist.

[0450] Die Variante des ortsspezifischen Cas9-Polypeptids weist eine verringerte Endodesoxyribonuklease-Aktivität auf. Beispielsweise weist eine Variante des ortsspezifischen Cas9-Polypeptids, die zur Verwendung in einem Verfahren zur Transkriptionsmodulation der vorliegenden Offenbarung geeignet ist, weniger als etwa 20%, weniger als etwa 15%, weniger als etwa 10%, weniger als etwa 5%, weniger als etwa 1% oder weniger als etwa 0,1% der Endodesoxyribonuklease-Aktivität eines Wildtyp-Cas9-Polypeptids, z. B. eines Wildtyp-Cas9-Polypeptids auf, das eine in **Fig. 3** gezeigte Aminosäuresequenz (SEQ ID NO: 8) umfasst. In einigen Ausführungsformen weist die Variante des ortsspezifischen Cas9-Polypeptids im Wesentlichen keine nachweisbare Endodesoxyribonuklease-Aktivität auf. In einigen Ausführungsformen, in denen ein ortsspezifisches Polypeptid eine verringerte katalytische Aktivität aufweist (z. B. wenn ein Cas9-Protein eine D10-, G12-, G17-, E762-, H840-, N854-, N863-, H982-, H983-, A984-, D986-, und/oder eine A987-Mutation aufweist, z. B. D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A und/oder D986A), kann das Polypeptid weiterhin in ortsspezifischer Weise an die Ziel-DNA binden (weil es weiterhin durch eine auf DNA gerichtete RNA zu einer Ziel-DNA-Sequenz geführt wird), solange sie die Fähigkeit beibehält, mit der auf DNA gerichteten RNA zu interagieren.

[0451] In einigen Fällen umfasst eine geeignete Variante des ortsspezifischen Cas9-Polypeptids eine Aminosäuresequenz, die mindestens etwa 75%, mindestens etwa 80%, mindestens etwa 85%, mindestens etwa 90%, mindestens etwa 95%, mindestens etwa 99% oder 100% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Cas9/Csn1-Aminosäuresequenz (SEQ ID NO: 8) aufweist oder zu den entsprechenden Teilen in einer der Aminosäuresequenzen von SEQ ID NOs: 1–256 und 795–1346.

[0452] In einigen Fällen kann die Variante des ortsspezifischen Cas9-Polypeptids den komplementären Strang der Ziel-DNA spalten, weist jedoch eine verringerte Fähigkeit zur Spaltung des nicht-komplementären Strangs der Ziel-DNA auf. Beispielsweise kann die Variante des ortsspezifischen Cas9-Polypeptids eine Mutation (Aminosäure-Substitution) aufweisen, welche die Funktion der RuvC-Domäne (z. B. "Domäne 1" von **Fig. 3**) verringert. Als nicht einschränkendes Beispiel ist in einigen Fällen die Variante des ortsspezifischen Cas9-Polypeptids eine D10A (Aspartat zu Alanin) Mutation der in **Fig. 3** gezeigten Aminosäuresequenz (oder die entsprechende Mutation einer der Aminosäuresequenzen, die in SEQ ID NOs: 1–256 und 795–1346 gezeigt sind).

[0453] In einigen Fällen kann die Variante des ortsspezifischen Cas9-Polypeptids den nicht-komplementären Strang der Ziel-DNA spalten, weist jedoch eine verringerte Fähigkeit auf, den komplementären Strang der Ziel-DNA zu spalten. Beispielsweise kann die Variante des ortsspezifischen Cas9-Polypeptids eine Mutation (Aminosäure-Substitution) aufweisen, die die Funktion der HNH-Domäne (RuvC/HNH/RuvC-Domänen-Motive, "Domäne 2" von **Fig. 3**) verringert. Als nicht einschränkendes Beispiel kann die Variante des ortsspezifischen Cas9-Polypeptids in einigen Fällen eine H840A (Histidin zu Alanin in Aminosäure-Position 840 von SEQ ID NO: 8) sein oder die entsprechende Mutation einer der Aminosäuresequenzen, die in SEQ ID NOs: 1–256 und 795–1346 gezeigt sind.

[0454] In einigen Fällen weist die Variante des ortsspezifischen Cas9-Polypeptids eine verringerte Fähigkeit zur Spaltung sowohl des komplementären als auch des nicht-komplementären Stranges der Ziel-DNA auf. Als nicht einschränkendes Beispiel umfasst in einigen Fällen die Variante des ortsspezifischen Cas9-Polypeptids sowohl eine D10A- als auch eine H840A-Mutation der in **Fig. 3** dargestellten Aminosäuresequenz (oder die entsprechende Mutation in einer der Aminosäuresequenzen, die in SEQ ID NOs: 1–256 und 795–1346 gezeigt sind).

[0455] Andere Reste können mutiert werden, um die gleiche Wirkung zu erreichen (d. h. Inaktivierung des einen oder anderen Nuklease-Teils). Als nicht einschränkende Beispiele können die Reste D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, und/oder A987 (oder die entsprechenden Mutationen der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Proteine) verändert (d. h. substituiert) werden (siehe **Fig. 3**, **Fig. 5**, **Fig. 11A**, und Tabelle 1 für weitere Informationen bezüglich der Konservierung von Aminosäure-Resten in Cas9). Darüber hinaus sind auch andere Mutationen als Alanin-Substitutionen geeignet.

[0456] In einigen Fällen ist die Variante des ortsspezifischen Cas9-Polypeptids ein Fusionspolypeptid (ein "Varianten-Cas9-Fusionspolypeptid"), d. h. ein Fusionspolypeptid, das Folgendes umfasst: (i) eine Variante

des ortsspezifischen Cas9-Polypeptids; und (ii) ein kovalent verbundenes heterologes Polypeptid (vorliegend auch als "Fusionspartner" bezeichnet).

[0457] Die heterologen Polypeptide können eine Aktivität zeigen (z. B. eine enzymatische Aktivität), die auch von dem Varianten-Cas9-Fusionspolypeptid gezeigt wird (z. B. Methyltransferase-Aktivität, Acetyltransferase-Aktivität, Kinase-Aktivität, ubiquitinierende Aktivität, usw.). Eine heterologe Nukleinsäuresequenz kann an eine andere Nukleinsäuresequenz geknüpft werden (z. B. durch genetische Modifikation), um eine chimäre Nukleotidsequenz zu erzeugen, die für ein chimäres Polypeptid kodiert. In einigen Ausführungsformen wird eine Varianten-Cas9-Fusionspolypeptid erzeugt, indem eine Cas9-Polypeptid-Variante mit einer heterologen Sequenz fusioniert wird, die für eine subzelluläre Lokalisierung sorgt (d. h. die heterologe Sequenz ist eine subzelluläre Lokalisierungs-Sequenz, z. B. ein nukleares Ortsbestimmungssignal (NLS) für das Targeting zum Zellkern; ein mitochondriales Ortsbestimmungssignal für das Targeting zu den Mitochondrien, ein Chloroplasten-Ortsbestimmungssignal für das Targeting zu einem Chloroplasten; eine ER-Retentionssignal, und ähnliches). In einigen Ausführungsformen kann die heterologe Sequenz einen Anhang (Tag) bereitstellen (d. h. die heterologe Sequenz ist eine nachweisbare Markierung), um das Tracking und/oder die Aufreinigung zu erleichtern (z. B. ein fluoreszentes Protein, z. B. grünes fluoreszentes Protein (GFP), YFP, RFP, CFP, mCherry, tdTomato und ähnliches; einen Histidin-Tag, z. B. ein 6xHis-Tag; einen Hämagglutinin(HA)-Tag; einen FLAG-Tag; einen Myc-Tag; und ähnliches). In einigen Ausführungsformen kann die heterologe Sequenz für eine erhöhte oder verringerte Stabilität sorgen (d. h. die heterologe Sequenz ist ein Stabilitäts-Kontrollpeptid, z. B. ein Degron, welches in einigen Fällen kontrollierbar ist (z. B. eine Temperatur-sensitive oder Wirkstoff-kontrollierbare Degron-Sequenz, siehe unten)). In einigen Ausführungsformen kann die heterologe Sequenz für eine erhöhte oder verringerte Transkription der Ziel-DNA sorgen (d. h. die heterologe Sequenz ist ein Transkriptions-Modulations-Sequenz, z. B. ein Transkriptionsfaktor/Aktivator oder ein Fragment davon, ein Protein oder Fragment davon, das einen Transkriptionsfaktor/Aktivator rekrutiert, ein Transkriptionsrepressor oder ein Fragment davon, ein Protein oder ein Fragment davon, das einen Transkriptionsrepressor rekrutiert, ein Transkriptionsregulator, der auf ein Small Molecule/einen Wirkstoff reagiert, usw.). In einigen Ausführungsformen kann die heterologe Sequenz eine Bindungsdomäne bereitstellen (d. h. die heterologe Sequenz ist eine Protein-bindende Sequenz, z. B. um für die Möglichkeit zu sorgen, dass ein chimäres dCas9-Polypeptid an ein anderes Protein von Interesse bindet, z. B. an ein DNA- oder Histon-modifizierendes Protein, einen Transkriptionsfaktor oder einen Transkriptionsrepressor, ein rekrutierendes Protein, usw.).

[0458] Geeignete Fusionspartner, die für eine erhöhte oder verringerte Stabilität sorgen, umfassen (sind jedoch nicht beschränkt auf) Degron-Sequenzen. Fachleute verstehen, dass Degron-Sequenzen Aminosäuresequenzen sind, welche die Stabilität des Proteins, von dem sie einen Teil bilden, kontrollieren. Beispielsweise wird die Stabilität eines Proteins, das eine Degron-Sequenz umfasst, zumindest teilweise durch die Degron-Sequenz kontrolliert. In einigen Fällen ist ein geeignetes Degron konstitutiv, sodass das Degron seinen Einfluss auf die Proteinstabilität unabhängig von einer experimentellen Kontrolle ausübt (d. h. das Degron ist nicht Wirkstoff-induzierbar, Temperatur-induzierbar, usw.). In einigen Fällen verleiht das Degron dem Varianten-Cas9-Polypeptid eine kontrollierbare Stabilität, so dass das Varianten-Cas9-Polypeptid "eingeschaltet" (d. h. stabil) oder "abgeschaltet" (d. h. instabil, degradiert) werden kann, abhängig von den gewünschten Umständen. Wenn beispielsweise das Degron ein Temperatur-sensitives Degron ist, kann das Varianten-Cas9-Polypeptid unterhalb einer Schwellentemperatur (z. B. 42°C, 41°C, 40°C, 39°C, 38°C, 37°C, 36°C, 35°C, 34°C, 33°C, 32°C, 31°C, 30°C, usw.) funktionsfähig ("an", stabil) sein, jedoch oberhalb der Schwellentemperatur nicht funktionsfähig ("aus", degradiert) sein. Als weiteres Beispiel kann das Vorhandensein oder das Fehlen eines Wirkstoffs das Protein vom Status "aus" (d. h. instabil) zum Status "an" (d. h. stabil) umschalten und umgekehrt, wenn das Degron ein Wirkstoff-induzierbares Degron ist. Ein beispielhaftes Wirkstoff-induzierbares Degron ist vom Protein FKBP12 angeleitet. Die Stabilität des Degrons wird durch das Vorhandensein oder das Fehlen eines Small Molecules kontrolliert, welches an das Degron bindet.

[0459] Beispiele für geeignete Degrons umfassen (sind jedoch nicht beschränkt auf) solche Degrons, die durch Shield-1, DHFR, Auxine und/oder die Temperatur kontrolliert werden. Nicht einschränkende Beispiele für geeignete Degrons sind im Stand der Technik bekannt (z. B. Dohmen et al., *Science*, 1994. 263(5151) : S. 1273–1276: Heat-inducible degon: a method for constructing temperature-sensitive mutants; Schoeber et al., *Am J Physiol Renal Physiol*. 2009 Jan; 296(1): F204–11: Conditional fast expression and function of multimeric TRPV5 channels using Shield-1; Chu et al., *Bioorg Med Chem Lett*. 2008 Nov 15; 18(22): 5941–4: Recent progress with FKBP-derived destabilizing domains; Kanemaki, *Pflugers Arch*. 2012 Dec 28: Frontiers of protein expression control with conditional degrons; Yang et al., *Mol Cell*. 2012 Nov 30; 48(4): 487–8: Titivated for destruction: the methyl degon; Barbour et al., *Biosci Rep*. 2013 Jan 18; 33(1): Characterization of the bipartite degon that regulates ubiquitin-independent degradation of thymidylate synthase; und Greussing et al., *J Vis Exp*. 2012 Nov 10; (69): Monitoring of ubiquitin-proteasome activity in living cells using a Degron

(dgn)-destabilized green fluorescent protein(GFP)-based reporter protein; von denen alle vorliegend in ihrer Gesamtheit durch Bezugnahme eingeschlossen sind).

[0460] Beispielhafte Degron-Sequenzen sind hinreichend charakterisiert und getestet worden, sowohl in Zellen als auch in Tieren. Somit erzeugt die Fusionierung von dCas9 an eine Degron-Sequenz ein "schaltbares" und "induzierbares" dCas9-Polypeptid. Alle vorliegend beschriebenen Fusionspartner können in jeder gewünschten Kombination verwendet werden. Als ein nicht-einschränkendes Beispiel zur Veranschaulichung dieses Punkts kann ein dCas9-Fusionsprotein eine YFP-Sequenz für die Detektion umfassen, eine Degron-Sequenz für die Stabilität und eine Transkriptions-Aktivator-Sequenz für die Erhöhung der Transkription der Ziel-DNA. Darüber hinaus ist die Anzahl der Fusionspartner, die in einem dCas9-Fusionsprotein verwendet werden können, unbegrenzt. In einigen Fällen umfasst ein dCas9-Fusionsprotein ein oder mehrere (z. B. zwei oder mehrere, drei oder mehrere, vier oder mehrere, fünf oder mehrere) heterologe Sequenzen.

[0461] Geeignete Fusionspartner umfassen (sind jedoch nicht beschränkt auf) ein Polypeptid, das für eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität, von denen jede auf die direkte Modifikation der DNA gerichtet sein kann (z. B. Methylierung der DNA) oder auf die Modifikation eines mit der DNA assoziierten Polypeptid (z. B. eines Histons oder DNA-bindenden Proteins). Weitere geeignete Fusionspartner umfassen, sind jedoch nicht beschränkt auf, Begrenzungselemente (z. B. CTCF), Proteinen und Fragmente derselben, die für ein peripheres Recruitment sorgen (z. B. Lamin A, Lamin B, usw.), und Protein-Docking-Elemente (z. B. FKBP/FRB, Pi11/Aby1, usw.).

[0462] Beispiele für verschiedene zusätzliche Fusionspartner (oder Fragmente davon) für ein vorliegendes ortsspezifisches Varianten-Cas9-Polypeptid umfassen (sind jedoch nicht beschränkt auf) die in **Fig. 54** aufgeführten.

[0463] In einigen Ausführungsformen kann ein vorliegendes ortsspezifisch modifizierendes Polypeptid Kodon-optimiert sein. Diese Art der Optimierung ist im Stand der Technik bekannt und umfasst die Mutation von ursprünglich fremder DNA, so dass die Kodon-Präferenzen des beabsichtigten Wirtsorganismus oder der beabsichtigten Wirtszelle nachgeahmt wird, während dasselbe Protein kodiert wird. Somit werden die Kodons verändert, wobei jedoch das kodierte Protein unverändert bleibt. Wenn beispielsweise die beabsichtigte Zielzelle eine humane Zelle wäre, würde ein humanes, Kodon-optimiertes dCas9 (oder eine dCas9-Variante) ein geeignetes ortsspezifisch modifizierendes Polypeptid darstellen. Als weiteres nicht einschränkendes Beispiel wäre, wenn die beabsichtigte Wirtszelle eine Mauszelle ist, ein in Bezug auf die Maus Kodon-optimiertes Cas9 (oder eine Variante, z. B. eine enzymatisch inaktive Variante) ein geeignetes Cas9-Polypeptid. Während eine Kodon-Optimierung nicht erforderlich ist, ist diese akzeptabel und kann in einigen Fällen bevorzugt sein. Wirtszellen

[0464] Ein Verfahren der vorliegenden Offenbarung zur Modulierung der Transkription kann verwendet werden, um die transkriptionale Modulation in mitotischen oder post-mitotischen Zellen in vivo und/oder ex vivo und/oder in vitro zu induzieren. Da die auf DNA gerichtete RNA durch Hybridisierung an die Ziel-DNA für Spezifität sorgt, kann es sich bei der mitotischen und/oder post-mitotischen Zelle um eine von vielen verschiedenen Wirtszellen handeln, wobei geeignete Wirtszellen folgendes umfassen (jedoch nicht darauf beschränkt sind) : eine bakterielle Zelle, eine Archaeen-Zelle, ein einzelliger eukaryontischer Organismus, eine Pflanzenzelle, eine Algenzelle, z. B. *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens* C. agardh, und ähnliche; eine Pilz-Zelle; eine tierische Zelle; eine Zelle von einem Invertebraten (z. B. Insekt, Nesseltier, Stachelhäuter, Fadenwurm, usw.); ein eukaryontischer Parasit (z. B. ein Malaria-Parasit, z. B. *Plasmodium falciparum*; ein Helminth; usw.); eine Zelle von einem Wirbeltier (z. B. Fisch, Amphibie, Reptil, Vogel, Säugetier); eine Zelle von einem Säugetier, z. B. eine Zelle von einem Nagetier, eine Zelle von einem Menschen, eine Zelle von einem nicht-menschlichen Primaten, usw.). Geeignete Wirtszellen umfassen natürlich vorkommende Zellen; genetisch modifizierte Zellen (z. B. Zellen, die genetisch im Labor modifiziert wurden, z. B. durch "menschlichen Eingriff"); und Zellen, die in irgendeiner Weise in vitro manipuliert wurden. In manchen Fällen ist die Wirtszelle isoliert.

[0465] Jede Art von Zelle kann von Interesse sein (z. B. eine Stammzelle, z. B. eine embryonische (ES)-Zelle, eine induzierte pluripotente Stammzelle (iPS), eine Keimzelle; eine somatische Zelle, z. B. ein Fibroblast, eine hämatopoetische Zelle, ein Neuron, eine Muskelzelle, eine Knochenzelle, eine Hepatozyte, eine Bauchspeicheldrüse-Zelle; eine in vitro oder in vivo vorliegende embryonische Zelle eines Embryos in einem beliebigen

Stadium, z. B. das 1-Zell-, 2-Zell-, 4-Zell-, 8-Zell-, usw., Stadium eines Zebrafisch-Embryos; usw.). Die Zellen können von etablierten Zelllinien stammen, oder es kann sich um primäre Zellen handeln, wobei „primäre Zellen“, „primäre Zelllinien“ und „primäre Kulturen“ vorliegend austauschbar verwendet werden und sich auf Zellen und Zellkulturen beziehen, die von einem Subjekt abgeleitet wurden, und denen es erlaubt wurde, in vitro für eine begrenzte Anzahl von Passagen zu wachsen, d. h. Aufspaltungen der Kultur. Beispielsweise umfassen primäre Kulturen solche Kulturen, die kein Mal, einmal, zweimal, viermal, fünfmal, zehnmal oder fünfzehnmal passagiert wurden, jedoch nicht häufig genug, um die Krisenphase zu durchlaufen. Primäre Zelllinien werden für weniger als zehn Passagen in vitro gehalten. Die Ziel-Zellen sind in vielen Ausführungsformen einzellige Organismen oder werden in Kultur angezüchtet.

[0466] Wenn die Zellen primäre Zelle sind, können solche Zellen von einem Individuum durch ein beliebiges handhabbares Verfahren erhalten werden. Beispielsweise können Leukozyten in praktischer Weise durch Aphärese, Leukozyten-Aphärese, Dichtegradienten-Zentrifugation, usw., erhalten werden, während Zellen von Geweben, wie beispielsweise Haut, Muskel, Knochenmark, Milz, Leber, Bauchspeicheldrüse, Lunge, Darm, Magen, usw. in praktischer Weise durch Biopsie erhalten werden. Eine geeignete Lösung kann für eine Dispersion oder eine Suspension der geernteten Zellen verwendet werden. Eine solche Lösung wird üblicherweise eine ausgewogene Salzlösung sein, z. B. normale Salzlösung, Phosphatgepufferte Salzlösung (PBS), ausgewogene Salzlösung nach Hanks, usw., die in praktischer Weise mit fötalem Kälberserum oder anderen natürlich vorkommenden Faktoren supplementiert sein kann, in Verbindung mit einem akzeptablen Puffer bei geringer Konzentration, z. B. von 5–25 mM. Praktische Puffer umfassen HEPES, Phosphatpuffer, Lactatpuffer, usw. Die Zellen können direkt verwendet werden, oder sie können in eingefrorenem Zustand für lange Zeiträume gelagert werden. Nach dem Auftauen können sie erneut verwendet werden. In solchen Fällen werden die Zellen üblicherweise in 10% Dimethylsulfoxid (DMSO), 50% Serum, 40% gepuffertes Medium oder in einer ähnlichen Lösung, die üblicherweise im Stand der Technik zur Aufbewahrung von Zellen bei Frosttemperaturen verwendet wird, eingefroren, und sie werden auf eine Weise aufgetaut, die im Stand der Technik für das Auftauen von gefrorenen kultivierten Zellen bekannt ist. Einbringen von Nukleinsäure in eine Wirtszelle

[0467] Eine auf DNA gerichtete RNA oder eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für selbige kodiert, kann in eine Wirtszelle durch eine Vielzahl von hinreichend bekannten Verfahren eingebracht werden. Wenn ein vorliegendes Verfahren die Einbringung einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, in eine Wirtszelle erfordert, so kann eine solche Nukleinsäure durch eine Vielzahl von hinreichend bekannten Verfahren in eine Wirtszelle eingebracht werden.

[0468] Verfahren zur Einbringung einer Nukleinsäure in eine Wirtszelle sind im Stand der Technik bekannt, und jedes beliebige Verfahren kann verwendet werden, um eine Nukleinsäure (z. B. ein Expressionskonstrukt) in eine Stammzelle oder in eine Vorläuferzelle einzubringen. Geeignete Verfahren umfassen, z. B. virale Infektion oder Bakteriophagen-Infektion, Transfektion, Konjugation, Protoplasten-Fusion, Lipofektion, Elektroporation, Kaliumphosphat-Präzipitation, Polyethylenimin(PEI)-vermittelte Transfektion, DEAE-Dextran-vermittelte Transfektion, Liposomen-vermittelte Transfektion, "Particle Gun"-Technologie, Kalziumphosphat-Präzipitation, direkte Mikroinjektion, Nanopartikel-vermittelte Nukleinsäure-Bereitstellung (siehe z. B., Panyam et al., Adv Drug Deliv Rev. 2012 Sep 13. pii: 50169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023), und ähnliches.

NUKLEINSÄUREN

[0469] Die vorliegende Offenbarung stellt eine isolierte Nukleinsäure bereit, die eine Nukleotidsequenz umfasst, welche für eine vorliegende auf DNA gerichtete RNA kodiert. In einigen Fällen umfasst eine vorliegende Nukleinsäure ferner eine Nukleotidsequenz, die für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert.

[0470] In einigen Ausführungsformen umfasst ein vorliegendes Verfahren das Einbringen von einer oder mehreren Nukleinsäuren, die Nukleotidsequenzen umfassen, welche für eine auf DNA gerichtete RNA und/oder ein ortsspezifisches Varianten-Cas9-Polypeptid kodieren, in eine Wirtszelle (oder eine Population von Wirtszellen). In einigen Ausführungsformen liegt eine Zelle, die eine Ziel-DNA umfasst, in vitro vor. In einigen Ausführungsformen liegt eine Zelle, die eine Ziel-DNA umfasst, in vivo vor. Geeignete Nukleinsäuren, die Nukleotidsequenzen umfassen, welche für eine auf DNA gerichtete RNA und/oder ein ortsspezifisches Polypeptid kodieren, umfassen Expressionsvektoren, wobei ein Expressionsvektor, der eine Nukleotidsequenz umfasst, die für eine auf DNA gerichtete RNA und/oder ein ortsspezifisches Polypeptid kodiert, ein "rekombinanter Expressionsvektor" ist.

[0471] In einigen Ausführungsformen ist der rekombinante Expressionsvektor ein virales Konstrukt, z. B. ein rekombinantes Adeno-assoziiertes Virus-Konstrukt (siehe beispielsweise US-Patent Nr. 7,078,387), ein rekombinantes adenovirales Konstrukt, ein rekombinantes lentivirales Konstrukt, ein rekombinantes retrovirales Konstrukt, usw.

[0472] Geeignete Expressionsvektoren umfassen (sind jedoch nicht beschränkt auf) virale Vektoren (z. B. virale Vektoren, die auf Vaccinia-Virus; Poliovirus; Adenovirus (siehe beispielsweise Lee et al., Invest Ophthalmol Vis Sci 35: 2543 2549, 1994; Borrás et al., Gene Ther 6: 515 524, 1999; Li and Davidson, PNAS 92: 7700 7704, 1995; Sakamoto et al., Hum Gene Ther 5:1088 1097, 1999; WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 und WO 95/00655); Adeno-assoziiertem Virus (siehe z. B., Ali et al., Hum Gene Ther 9: 81 86, 1998, Flannery et al., PNAS 94: 6916 6921, 1997; Bennett et al., Invest Ophthalmol Vis Sci 38: 2857 2863, 1997; Jomary et al., Gene Ther 4: 683 690, 1997, Rolling et al., Hum Gene Ther 10: 641 648, 1999; Ali et al., Hum Mol Genet 5: 591 594, 1996; Srivastava in WO 93/09239, Samulski et al., J. Vir. (1989) 63: 3822–3828; Mendelson et al., Virol. (1988) 166: 154–165; und Flotte et al., PNAS (1993) 90: 10613–10617); SV40; Herpes Simplex-Virus; humanem Immundefizienzvirus (siehe z. B. Miyoshi et al., PNAS 94: 10319 23, 1997; Takahashi et al., J Virol 73: 7812 7816, 1999); einem retroviralen Vektor (z. B. murinem Leukämievirus, Milz-Nekrose-Virus und Vektoren, die von Retroviren abgeleitet sind, wie beispielsweise Rous-Sarkoma-Virus, Harvey-Sarkoma-Virus, Vogel-Leukose-Virus, Lentivirus, humanem Immundefizienz-Virus, myeloproliferativem Sarkoma-Virus und Brusttumor-Virus) basieren), und ähnliches.

[0473] Zahlreiche geeignete Expressionsvektoren sind dem Fachmann bekannt und viele sind kommerziell verfügbar. Die folgenden Vektoren werden beispielhaft bereitgestellt: für eukaryontische Wirtszellen: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG und pSVLSV40 (Pharmacia). Beliebige andere Vektoren können jedoch ebenfalls verwendet werden, solange diese kompatibel mit der Wirtszelle sind.

[0474] Abhängig von dem Wirt-Vektor-System, das verwendet wird, können eine beliebige Vielzahl geeigneter Transkriptions- und Translations-Kontrollelemente, einschließlich konstitutiver und induzierbarer Promotoren, Transkriptions-Verstärkungselemente, Transkriptions-Terminatoren, usw., im Expressionsvektor verwendet werden (siehe z. B. Bitter et al. (1987) Methods in Enzymology, 153: 516–544).

[0475] In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, in funktionsfähiger Weise mit einem Kontrollelement verknüpft, z. B. einem transkriptionalen Kontrollelement, wie beispielsweise einem Promotor. Das transkriptionale Kontrollelement kann entweder in einer eukaryontischen Zelle, z. B. in einer Säugetierzelle, oder in einer prokaryontischen Zelle, z. B. in einer bakteriellen Zelle oder in einer Archaeen-Zelle funktionsfähig sein. In einigen Ausführungsformen ist eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA und/oder ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert in funktionsfähiger Weise an mehrere Kontrollelemente gekoppelt, die eine Expression der Nukleotidsequenz, welche für die auf DNA gerichtete RNA und/oder ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, sowohl in prokaryontischen als auch eukaryontischen Zellen erlaubt.

[0476] Ein Promotor kann ein konstitutiv aktiver Promotor (d. h. ein Promotor, der konstitutiv in einem aktiven Zustand/"an" ist), es kann ein induzierbarer Promotor sein (d. h. ein Promotor, dessen Zustand aktiv/"an" oder inaktiv/"aus" durch externe Reize kontrolliert wird, z. B. durch das Vorhandensein einer bestimmten Temperatur, Verbindung oder eines Proteins), es kann ein räumlich beschränkter Promotor sein (d. h. ein transkriptionales Kontroll-Element, ein Enhancer, usw.) (z. B. Gewebe-spezifischer Promotor, Zelltyp-spezifischer Promotor, usw.), und es kann ein zeitlich beschränkter Promotor sein (d. h. der Promotor ist während spezifischer Phasen der embryonalen Entwicklung oder während spezifischer Phasen eines biologischen Prozesses, z. B. Haarfollikelzyklus in Mäusen, im Zustand "AN" oder "AUS").

[0477] Geeignete Promotoren können von Viren abgeleitet werden und können daher als virale Promotoren bezeichnet werden, oder sie können von beliebigen Organismen abgeleitet werden, einschließlich prokaryontischer oder eukaryontischer Organismen. Geeignete Promotoren können verwendet werden, um die Expression durch eine beliebige RNA-Polymerase zu steuern (z. B. pol I, pol II, pol III). Beispielhafte Promotoren umfassen (sind jedoch nicht beschränkt auf) den SV40-Early-Promotor, den Long-Terminal-Repeat(LTR)-Promotor aus dem Brusttumor-Virus der Maus, den Adenovirus-Major-Late-Promotor (Ad MLP), einen Herpes Simplex-Virus(HSV)-Promotor, einen Cytomegalovirus(CMV)-Promotor, wie z. B. die CMV-Immediate-Early-Promotor-Region (CMVIE), ein Rous Sarkoma-Virus(RSV)-Promotor, einen humanen U6-Small-Nuclear-Promotor (U6) (Miyagishi et al., Nature Biotechnology 20, 497–500 (2002)), einen verstärkten U6-Promotor (z. B. Xia et al., Nucleic Acids Res. 2003 Sep 1; 31(17)), einen humanen H1-Promotor (H1), und ähnliche.

[0478] Beispiele für induzierbare Promotoren umfassen (sind jedoch nicht beschränkt auf) den T7-RNA-Polymerase-Promotor, den T3-RNA-Polymerase-Promotor, den Isopropyl-Beta-D-thiogalactopyranosid(IPTG)-regulierten Promotor, den Laktose-induzierten Promotor, den Hitzeschock-Promotor, den Tetracyclin-regulierten Promotor (z. B. Tet-AN, Tet-AUS, usw.), den Steroid-regulierten Promotor, den Metall-regulierten Promotor, den Östrogenrezeptor-regulierten Promotor, usw. Induzierbare Promotoren können daher durch Moleküle reguliert werden, die folgendes umfassen (jedoch nicht darauf beschränkt sind): Doxycyclin; RNA-Polymerase, z. B. T7-RNA-Polymerase; einen Östrogenrezeptor; eine Östrogenrezeptor-Fusion; usw.

[0479] In einigen Ausführungsformen ist der Promotor ein räumlich beschränkter Promotor (d. h. ein Zelltyp-spezifischer Promotor, ein Gewebe-spezifischer Promotor, usw.), sodass der Promotor in einem multizellulären Organismus in einer Untergruppe von spezifischen Zellen aktiv (d. h. "An") ist. Räumlich beschränkte Promotoren können auch als Enhancer, Transkriptions-Kontrollelemente, Kontrollsequenzen, usw., bezeichnet werden. Jeder beliebige räumlich beschränkte Promotor kann verwendet werden, und die Auswahl eines geeigneten Promotors (z. B. ein Gehirn-spezifischer Promotor, ein Promotor, der die Expression in einer Untergruppe von Neuronen steuert, ein Promotor, der die Expression in der Keimbahn steuert, ein Promotor, der die Expression in den Lungen steuert, ein Promotor, der die Expression in den Muskeln steuert, ein Promotor, der die Expression in Inselzellen der Bauspeicheldrüse steuert, usw.) wird vom Organismus abhängen. Beispielsweise sind verschiedene räumlich beschränkte Promotoren für Pflanzen, Fliegen, Würmer, Säugetiere, Mäuse, usw., bekannt. Somit kann ein räumlich beschränkter Promotor verwendet werden, um die Expression einer Nukleinsäure, die für ein vorliegendes ortsspezifisch modifizierendes Polypeptid kodiert, in einer großen Vielzahl von verschiedenen Geweben und Zelltypen in Abhängigkeit vom Organismus zu regulieren. Einige räumlich beschränkte Promotoren sind auch in zeitlicher Hinsicht beschränkt, sodass der Promotor während bestimmter Phasen der embryonalen Entwicklung oder während bestimmter Phasen eines biologischen Prozesses ist (z. B. Haarfollikelzyklus in Mäusen) im Zustand "An" oder im Zustand "Aus" ist.

[0480] Beispiele für räumlich beschränkte Promotoren umfassen (sind jedoch nicht beschränkt auf) Neuronen-spezifische Promotoren, Adipozyten-spezifische Promotoren, Kardiomyozyten-spezifische Promotoren, für glatte Muskelzellen spezifische Promotoren, Photorezeptor-spezifische Promotoren, usw. Räumlich beschränkte Neuronen-spezifische Promotoren umfassen (sind jedoch nicht beschränkt auf) einen Neuronen-spezifischen-Enolase(NSE)-Promotor (siehe z. B., EMBL HSENO2, X51956); einen Promotor der aromatischen Aminosäure-Decarboxylase (AADC); einen Neurofilament-Promotor (siehe z. B. GenBank HUMNFL, L04147); einen Synapsin-Promotor (z. B. GenBank HUMSYNIB, M55301); einen thy-1-Promotor (siehe z. B. Chef et al. (1987) Cell 51: 7–19; und Llewellyn, et al. (2010) Nat. Med. 16(10): 1161–1166); einen Serotonin-Rezeptor-Promotor (siehe z. B. GenBank S62283); einen Tyrosin-Hydroxylase-Promotor (TH) (siehe z. B. Oh et al. (2009) Gene Ther 16: 437; Sasaoka et al. (1992) Mol. Brain Res. 16: 274; Boundy et al. (1998) J. Neurosci. 18: 9989; und Kaneda et al. (1991) Neuron 6: 583–594); einen GnRH-Promotor (siehe z. B. Radovick et al. (1991) Proc. Natl. Acad. Sci. USA 88: 3402–3406); einen L7-Promotor (siehe z. B. Oberdick et al. (1990) Science 248: 223–226); einen DNMT-Promotor (siehe z. B. Bartge et al. (1988) Proc. Natl. Acad. Sci. USA 85: 3648–3652); einen Enkephalin-Promoter (siehe z. B. Comb et al. (1988) EMBO J. 17: 3793–3805); einen Myelin-basisches-Protein(MBP)-Promoter; einen Promoter der Ca²⁺-Calmodulin-abhängigen Proteinkinase II-alpha (CamKIIα) (siehe z. B. Mayford et al. (1996) Proc. Natl. Acad. Sci. USA 93: 13250; und Casanova et al. (2001) Genesis 31: 37); einen CMV-Enhancer/Blutplättchen-abgeleiteter Wachstumsfaktor-β-Promoter (siehe z. B. Liu et al. (2004) Gene Therapy 11: 52–60); und ähnliche.

[0481] Räumlich beschränkte, Adipozyten-spezifische Promotoren umfassen (sind jedoch nicht beschränkt auf) einen aP2-Gen-Promotor/Enhancer, z. B. eine Region von –5,4 kb to +21 bp eines humanen aP2-Gens (siehe z. B. Tozzo et al. (1997) Endocrinol. 138: 1604; Ross et al. (1990) Proc. Natl. Acad. Sci. USA 87: 9590; und Pavjani et al. (2005) Nat. Med. 11: 797); einen Glukose-Transporter-4(GLUT4)-Promoter (siehe z. B. Knight et al. (2003) Proc. Natl. Acad. Sci. USA 100: 14725); einen Fettsäure-Translokase(FAT/CD36)-Promoter (siehe z. B. Kuriki et al. (2002) Biol. Pharm. Bull. 25: 1476; und Sato et al. (2002) J. Biol. Chem. 277: 15703); einen Stearoyl-CoA-Desaturase-1(SCD1)-Promoter (Tabor et al. (1999) J. Biol. Chem. 274: 20603); einen Leptin-Promotor (siehe z. B. Mason et al. (1998) Endocrinol. 139: 1013; und Chen et al. (1999) Biochem. Biophys. Res. Comm. 262: 187); einen Adiponectin-Promoter (siehe z. B. Kita et al. (2005) Biochem. Biophys. Res. Comm. 331: 484; und Chakrabarti (2010) Endocrinol. 151: 2408); einen Adipsin-Promoter (siehe z. B. Platt et al. (1989) Proc. Natl. Acad. Sci. USA 86: 7490); einen Resistin-Promoter (siehe z. B. Seo et al. (2003) Molec. Endocrinol. 17: 1522); und ähnliche.

[0482] Räumlich beschränkte, Kardiomyozyten-spezifische Promotoren umfassen (sind jedoch nicht beschränkt auf) Kontrollsequenzen, die von den folgenden Genen abgeleitet sind: Myosin-leichte Kette-2, α-Myosin-schwere Kette, AE3, kardiäres Troponin-C, kardiäres Actin, und ähnliche. Franz et al. (1997) Cardiovasc.

Res. 35: 560–566; Robbins et al. (1995) Ann. N. Y. Acad. Sci. 752: 492–505; Linn et al. (1995) Circ. Res. 76: 584–591; Parmacek et al. (1994) Mol. Cell. Biol. 14: 1870–1885; Hunter et al. (1993) Hypertension 22: 608–617; und Sartorelli et al. (1992) Proc. Natl. Acad. Sci. USA 89: 4047–4051.

[0483] Räumlich beschränkte, für glatte Muskeln spezifische Promotoren umfassen (sind jedoch nicht beschränkt auf) einen SM22 α -Promotor (siehe z. B. Akyürek et al. (2000) Mol. Med. 6: 983; und US-Patent Nr. 7, 169,874); einen Smoothelin-Promoter (siehe z. B. WO 2001/018048); einen Promotor von α -Actin des glatten Muskels; und ähnliche. Beispielsweise konnte gezeigt werden, dass eine 0,4 kb Region des SM22 α -Promotors, die zwei CARG-Elemente enthält, die Zell-spezifische Expression im vaskulären glatten Muskel vermittelt (siehe z. B. Kim, et al. (1997) Mol. Cell. Biol. 17, 2266–2278; Li, et al., (1996) J. Cell Biol. 132, 849–859; und Moessler, et al. (1996) Development 122, 2415–2425).

[0484] Räumlich beschränkte, Photorezeptor-spezifische Promotoren umfassen (sind jedoch nicht beschränkt auf) einen Rhodopsin-Promotor; einen Rhodopsin-Kinase-Promotor (Young et al. (2003) Ophthalmol. Vis. Sci. 44: 4076); einen Beta-Phosphodiesterase-Gen-Promotor (Nicoud et al. (2007) J. Gene Med. 9: 1015); einen Retinitis pigmentosa-Gen-Promoter (Nicoud et al. (2007) siehe oben); einen Interphotorezeptor-Retinoid-Bindungsprotein (IRBP) Gen-Enhancer (Nicoud et al. (2007) siehe oben); einen IRBP-Gen-Promoter (Yokoyama et al. (1992) Exp Eye Res. 55: 225); und ähnliche.

BIBLIOTHEKEN

[0485] Die vorliegende Offenbarung stellt eine Bibliothek von auf DNA gerichteten RNAs bereit. Die vorliegende Offenbarung stellt eine Bibliothek aus Nukleinsäuren bereit, die Nukleotide umfassen, welche für auf DNA gerichtete RNAs kodieren. Eine vorliegende Bibliothek aus Nukleinsäuren, die Nukleotide umfassen, welche für auf DNA gerichtete RNAs kodieren, können eine Bibliothek aus rekombinanten Expressionsvektoren umfassen, die Nukleotide umfassen, welche für auf DNA gerichtete RNAs kodieren.

[0486] Eine vorliegende Bibliothek kann von etwa 10 einzelnen Mitgliedern bis etwa 10¹² einzelnen Mitgliedern umfassen; z. B. kann eine vorliegende Bibliothek von etwa 10 einzelnen Mitgliedern bis etwa 10² einzelnen Mitgliedern, von etwa 10² einzelnen Mitgliedern bis etwa 10³ einzelnen Mitgliedern, von etwa 10³ einzelnen Mitgliedern bis etwa 10⁵ einzelnen Mitgliedern, von etwa 10⁵ einzelnen Mitgliedern bis etwa 10⁷ einzelnen Mitgliedern, von etwa 10⁷ einzelnen Mitgliedern bis etwa 10⁹ einzelnen Mitgliedern, oder von etwa 10⁹ einzelnen Mitgliedern bis etwa 10¹² einzelnen Mitgliedern umfassen.

[0487] Ein "einzelnes Mitglied" einer vorliegenden Bibliothek unterscheidet sich von anderen Mitgliedern der Bibliothek in der Nukleotidsequenz des auf DNA gerichteten Segments der auf DNA gerichteten RNA. Somit kann z. B. jedes einzelne Mitglied einer vorliegenden Bibliothek die gleiche oder im Wesentlichen die gleiche Nukleotidsequenz des Protein-bindenden Segments umfassen wie alle anderen Mitglieder der Bibliothek; und es kann die gleiche oder im Wesentlichen die gleiche Nukleotidsequenz des transkriptionalen Terminations-Segments umfassen wie alle anderen Mitglieder der Bibliothek; es unterscheidet sich jedoch von anderen Mitgliedern der Bibliothek in der Nukleotidsequenz des auf DNA gerichteten Segments der auf DNA gerichteten RNA. Auf diese Weise kann die Bibliothek Mitglieder umfassen, die an unterschiedliche Ziel-Nukleinsäuren binden.

ANWENDUNGEN

[0488] Ein Verfahren zur Modulierung der Transkription gemäß der vorliegenden Offenbarung kann in einer Vielzahl von Anwendungen verwendet werden, die ebenfalls bereitgestellt werden. Diese Anwendungen umfassen Forschungsanwendungen; diagnostische Anwendungen; industrielle Anwendungen und Behandlungsanwendungen.

[0489] Forschungsanwendungen umfassen z. B. die Bestimmung der Wirkung der Verringerung oder der Erhöhung der Transkription einer Ziel-Nukleinsäure z. B. auf die Entwicklung, den Metabolismus, die Expression eines stromabwärts gelegenen Gens, und ähnliches.

[0490] Es kann eine Hochdurchsatz-Genom-Analyse unter Verwendung eines vorliegenden Verfahren zur Transkriptionsmodulation durchgeführt werden, in der lediglich das auf DNA gerichtete Segment der auf DNA gerichteten RNA variiert werden muss, während das Protein-bindende Segment und das Transkriptions-Terminations-Segment (in einigen Fällen) konstant gehalten werden können. Eine Bibliothek (z. B. eine vorliegende Bibliothek), die eine Vielzahl von Nukleinsäuren umfasst, welche in der Genom-Analyse verwendet wer-

den, würde folgendes umfassen: einen Promotor, der in funktionsfähiger Weise mit einer Nukleotidsequenz verknüpft ist, die für eine auf DNA gerichtete RNA kodiert, wobei jede Nukleinsäure ein unterschiedliches auf DNA gerichtetes Segment umfassen würde, ein übliches Protein-bindendes Segment und ein übliches Transkriptions-Terminations-Segment. Ein Chip könnte mehr als 5×10^4 einzigartige auf DNA gerichtete RNAs umfassen. Anwendungen würden die Phänotypisierung im großen Ausmaß, eine Gene-to-Function-Mapping und eine metagenomische Analyse umfassen.

[0491] Die vorliegend offenbarten Verfahren finden Verwendung auf dem Gebiet des metabolischen Engineerings. Da die Transkriptionsmengen wirksam und zuverlässig durch Erstellen von geeigneten auf DNA gerichteten RNAs kontrolliert werden können, wie es vorliegend offenbart wird, kann die Aktivität der metabolischen Wege (z. B. der biosynthetischen Wege) präzise kontrolliert und eingestellt werden, indem die Mengen der spezifischen Enzyme innerhalb eines metabolischen Wegs von Interesse kontrolliert werden (z. B. durch erhöhte oder verringerte Transkription). Metabolische Wege von Interesse umfassen solche, die für chemische Produktion (Feinchemikalien, Treibstoff, Antibiotika, Toxine, Agonisten, Antagonisten, usw.) und/oder für die Wirkstoffproduktion verwendet werden.

[0492] Biosynthesewege von Interesse umfassen, sind jedoch nicht beschränkt auf, (1) den Mevalonat-Weg (z. B. HMG-CoA-Reduktase-Weg) (wandelt Acetyl-CoA zu Dimethylallyl-Pyrophosphat (DMAPP) und Isopentenyl-Pyrophosphat (IPP) um, die für die Biosynthese einer großen Vielzahl von Biomolekülen verwendet werden, einschließlich Terpenoiden/Isoprenoiden), (2) den Nicht-Mevalonat-Weg (d. h. den "2-C-Methyl-D-Erythritol-4-Phosphat/1-desoxy-D-xylulose 5-phosphat-Weg" oder "MEP/DOXP-Weg" oder "DXP-Weg") (produziert ebenfalls DMAPP und IPP, allerdings durch Umwandlung von Pyruvat und Glyceraldehyd-3-Phosphat in DMAPP und IPP über einen zum Mevalonat-Weg alternativen Weg), (3) der Polyketid-Synthese-Weg (erzeugt eine Vielzahl von Polyketiden durch eine Vielzahl von Polyketid-Synthase-Enzyme). Polyketide umfassen natürlich vorkommende kleine Moleküle, die für die Chemotherapie verwendet werden (z. B. Tetracyclin und Makrolide), und industriell wichtige Polyketide umfassen Rapamycin (Immunsuppressivum), Erythromycin (Antibiotikum), Lovastatin (Anti-Cholesterin-Wirkstoff) und Epothilon B (Anti-Krebsmittel), (4) Fettsäuresyntheseweg, (5) der DAHP(3-Desoxy-D-arabino-heptulosonat-7-phosphat)-Syntheseweg, (6) Wege, die potentielle Biotreibstoffe erzeugen (wie beispielsweise kurzkettige Alkohole und Alkane, Fettsäuremethylester und Fettalkohole, Isoprenoide, usw.) usw.

Netzwerke und Kaskaden

[0493] Die vorliegend offenbarten Verfahren können verwendet werden, um integrierte Kontroll-Netzwerke zu erstellen (d. h. eine Kaskade oder Kaskaden). Beispielsweise kann eine vorliegende auf DNA gerichtete RNA/ein ortsspezifisches Varianten-Cas9-Polypeptid verwendet werden, um die Expression einer anderen auf DNA gerichteten RNA oder eines anderen vorliegenden ortsspezifischen Varianten-Cas9-Polypeptids zu kontrollieren (d. h. zu modulieren, z. B. erhöhen, verringern). Beispielsweise kann eine erste auf DNA gerichtete RNA so erstellt werden, dass sie auf die Modulation der Transkription eines zweiten chimären dCas9-Polypeptids gerichtet ist, das eine Funktion aufweist, die sich vom ersten ortsspezifischen Varianten-Cas9-Polypeptid unterscheidet (z. B. Methyltransferase-Aktivität, Demethylase-Aktivität, Acetyltransferase-Aktivität, Deacetylase-Aktivität, usw.). Da verschiedene dCas9-Proteine (die z. B. von unterschiedlichen Spezies abgeleitet sind) einen unterschiedlichen Cas9-Handle (d. h. Protein-bindendes Segment) erfordern können, kann das zweite chimäre dCas9-Polypeptid darüber hinaus von einer anderen Spezies abgeleitet werden als das erste oben genannte dCas9-Polypeptid. Somit kann in einigen Fällen das zweite chimäre dCas9-Polypeptid so ausgewählt werden, dass es nicht mit der ersten auf DNA gerichteten RNA interagiert. In anderen Fällen kann das zweite chimäre dCas9-Polypeptid so ausgewählt werden, dass es mit der ersten auf DNA gerichteten RNA interagiert. In einigen dieser Fälle können die Aktivitäten der zwei (oder mehreren) dCas9-Proteine miteinander konkurrieren (z. B. wenn die Polypeptide gegensätzliche Aktivitäten aufweisen), oder sie können sich synergistisch unterstützen (z. B. wenn die Polypeptide eine ähnliche oder synergistische Aktivität aufweisen). Wie oben erwähnt können in ähnlicher Weise beliebige der Komplexe (d. h. auf DNA gerichtete RNA/dCas9-Polypeptid) in dem Netzwerk so erstellt werden, dass sie andere auf DNA gerichtete RNAs oder dCas9-Polypeptide kontrollieren. Weil eine vorliegende auf DNA gerichtete RNA und ein vorliegendes ortsspezifisches Varianten-dCas9-Polypeptid zu einer beliebigen gewünschten DNA-Sequenz geleitet werden können, können die vorliegend beschriebenen Verfahren verwendet werden, um die Expression eines beliebigen gewünschten Ziels zu kontrollieren und zu regulieren. Die integrierten Netzwerke (d. h. Kaskaden von Interaktionen), die erstellt werden können, reichen von sehr einfach bis sehr komplex, wobei keinerlei Einschränkung besteht.

[0494] In einem Netzwerk, in dem zwei oder mehrere Bestandteile (z. B. auf DNA gerichtete RNAs, Aktivator-RNAs, Targeter-RNAs oder dCas9-Polypeptide) jeweils unter der regulatorischen Kontrolle eines anderen

Komplexes aus auf DNA gerichteter RNA/dCas9-Polypeptid stehen, kann das Niveau der Expression einer Komponente des Netzwerks das Niveau der Expression einer anderen Komponente des Netzwerks beeinflussen (z. B. Erhöhung oder Verringerung der Expression). Durch diesen Mechanismus kann die Expression einer Komponente die Expression einer anderen Komponente im gleichen Netzwerk beeinflussen, und das Netzwerk kann eine Mischung von Komponenten enthalten, welche die Expression von anderen Komponenten erhöhen, sowie auch Komponenten, welche die Expression von anderen Komponenten verringern. Wie ein Fachmann leicht verstehen wird, sind die obigen Beispiele, bei denen das Expressionsniveau einer Komponente das Expressionsniveau von einer oder mehreren anderen Komponenten beeinflussen kann, lediglich zum Zwecke der Veranschaulichung aufgeführt und sind nicht einschränkend. Eine zusätzliche Komplexitätsebene kann wahlweise in ein Netzwerk eingebracht werden, wenn eine oder mehrere Komponenten so modifiziert werden (wie oben beschrieben), dass sie manipulierbar sind (d. h. unter experimenteller Kontrolle, z. B. Temperaturkontrolle; Wirkstoffkontrolle, d. h. Wirkstoff-induzierbare Kontrolle; Lichtkontrolle; usw.).

[0495] Als ein nicht einschränkendes Beispiel kann eine erste auf DNA gerichtete RNA an einen Promotor einer zweiten auf DNA gerichteten RNA binden, welche die Expression eines therapeutischen/metabolischen Ziel-Gens kontrolliert. In einem solchen Fall aktiviert die konditionale Expression der ersten auf DNA gerichteten RNA indirekt das therapeutische/metabolische Gen. RNA-Kaskaden dieser Art sind nützlich, beispielsweise für die einfache Umwandlung eines Repressors in einen Aktivator, und sie können verwendet werden, um die Logik oder Dynamik der Expression eines Ziel-Gens zu kontrollieren.

[0496] Ein vorliegendes Verfahren zur Modulation der Transkription kann ferner für die Wirkstoffentdeckung und die Ziel-Validierung verwendet werden.

KITS

[0497] Die vorliegende Offenbarung stellt einen Kit zur Durchführung eines vorliegenden Verfahrens bereit. Ein vorliegendes Kit umfasst: (a) eine auf DNA gerichtete RNA der vorliegenden Offenbarung oder eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst: (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Ziel-Sequenz in der Ziel-DNA ist; (ii) ein zweites Segment, das mit einem ortsspezifischen Polypeptid interagiert; und (iii) einen transkriptionalen Terminator; und (b) einen Puffer. In einigen Fällen umfasst die Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert, ferner eine Nukleotidsequenz, die für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, das relativ zum Wildtyp-Cas9 eine verringerte Desoxyribonuklease-Aktivität aufweist.

[0498] In einigen Fällen umfasst ein Kit ferner ein ortsspezifisches Varianten-Cas9-Polypeptid, das relativ zum Wildtyp-Cas9 eine verringerte Desoxyribonuklease-Aktivität aufweist.

[0499] In einigen Fällen umfasst ein vorliegender Kit ferner eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, das relativ zum Wildtyp-Cas9 eine verringerte Desoxyribonuklease-Aktivität aufweist.

[0500] Ein vorliegender Kit kann ferner ein oder mehrere zusätzliche Reagenzien umfassen, wobei solche zusätzlichen Reagenzien ausgewählt sein können aus: einem Puffer; einem Waschpuffer; einem Kontrollreagenz; einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid; einem Reagenz für die in vitro-Herstellung des ortsspezifischen Varianten-Cas9-Polypeptids ausgehend von DNA; und ähnliches. In einigen Fällen ist das ortsspezifische Varianten-Cas9-Polypeptid, welches in einem vorliegenden Kit enthalten ist, ein ortsspezifisches Varianten-Cas9-Fusionspolypeptid, wie es oben beschrieben ist.

[0501] Komponenten eines vorliegenden Kits können in getrennten Behältern vorliegen; oder sie können in einen einzigen Behälter kombiniert werden.

[0502] Zusätzlich zu den oben genannten Bestandteilen kann ein vorliegender Kit ferner Hinweise für die Verwendung der Bestandteile des Kits zur Durchführung der vorliegenden Verfahren umfassen. Diese Anweisungen zur Durchführung der vorliegenden Verfahren sind im Allgemeinen auf einem geeigneten Speichermedium aufgezeichnet. Beispielsweise können die Anweisungen auf ein Substrat, wie beispielsweise Papier oder Plastik, usw., aufgedruckt sein. Somit können die Instruktionen in den Kits als Packungsbeilage vorliegen, als Beschriftung des Behälters des Kits oder Bestandteilen davon (d. h. mit der Verpackung oder einer Unterverpackung assoziiert), usw. In anderen Ausführungsformen liegen die Instruktionen als elektronische Speicherdatei auf einem geeigneten Computer-lesbarem Speichermedium vor, z. B. auf CD-ROM, Diskette,

Flash-Drive, usw. In noch weiteren Ausführungsformen sind die eigentlichen Instruktionen nicht im Kit vorhanden, sondern es werden lediglich Mittel zum Erhalt der Anweisungen von einer entfernten Quelle, z. B. über das Internet, bereitgestellt. Ein Beispiel für diese Ausführungsform ist ein Kit, das eine Web-Adresse enthält, unter der die Anweisungen eingesehen werden können und/oder von der die Anweisungen heruntergeladen werden können. Wie die Anweisungen sind auch die Mittel zum Erhalt der Anweisungen auf einem geeigneten Substrat aufgezeichnet.

BEISPIELE

[0503] Die folgenden Beispiele werden so bereitgestellt, dass sie Fachleute mit einer vollständigen Offenbarung und Beschreibung der Herstellung und Verwendung der vorliegenden Erfindung versehen, und sie sind nicht dazu gedacht, den Umfang dessen, was die Erfinder als ihre Erfindung ansehen, einzuschränken, und sie sind auch nicht dazu gedacht, den Eindruck zu erwecken, dass die nachfolgenden Experimente die gesamten oder die einzigen durchgeführten Experimente darstellen. Es wurde sich bemüht, bei den Zahlenwerten auf Genauigkeit zu achten (z. B. Mengen, Temperaturen, usw.), wobei einige experimentelle Fehler und Abweichungen jedoch berücksichtigt werden sollten. Sofern nicht anderweitig angegeben, sind Teile Gewichtsteile, Molekulargewicht ist durchschnittliches Molekulargewicht, Temperatur ist in Grad Celsius und Druck ist atmosphärischer Druck oder annähernd atmosphärischer Druck. Es können Standardabkürzungen verwendet werden, z. B. bp, Basenpaar(e); kb, Kilobase(n); pl, Pikoliter; s oder sec Sekunde(n); min, Minute(n); h oder hr, Stunde(n); aa Aminosäure(n); kb, Kilobase(n); bp, Basenpaar(e); nt, Nukleotid(e); i. m., intramuskulär; i. p., intraperitoneal; s. c., subkutan; und ähnliches

Beispiel 1: Verwendung von Cas9 zur Erzeugung von Modifikationen in Ziel-DNA

MATERIAL UND METHODEN

Bakterielle Stämme und Kulturbedingungen

[0504] *Streptococcus pyogenes*, kultiviert in THY-Medium (Todd Hewitt Broth (THB, Bacto, Becton Dickinson), supplementiert mit 0,2% Hefeextrakt (Oxoid)) oder auf TSA (Trypticase-Soja-Agar, BBL, Becton Dickinson), supplementiert mit 3% Schafblut, wurde in einer Atmosphäre, die mit 5% CO₂ supplementiert war, bei 37°C ohne Schütteln inkubiert. *Escherichia coli*, kultiviert in Luria-Bertani (LB)-Medium und -Agar wurde bei 37°C mit Schütteln inkubiert. Sofern erforderlich, wurden geeignete Antibiotika dem Medium in den folgenden Endkonzentrationen zugesetzt: Ampicillin, 100 µg/ml für *E. coli*; Chloramphenicol, 33 µg/ml für *Escherichia coli*; Kanamycin, 25 µg/ml für *E. coli* und 300 µg/ml für *S. pyogenes*. Das bakterielle Zellwachstum wurde regelmäßig durch Messung der optischen Dichte von Kultur-Aliquots bei 620 nm unter Verwendung eines Mikroplatten-Lesegeräts (SLT Spectra-Lesegerät) gemessen.

Transformation von bakteriellen Zellen

[0505] Die Plasmid-DNA-Transformation in *E. coli*-Zellen wurde gemäß einem Standard-Hitzeschock-Protokoll durchgeführt. Die Transformation von *S. pyogenes* wurde wie zuvor beschrieben mit einigen Modifikationen durchgeführt. Der Transformationsassay, der durchgeführt wurde, um die CRISPR/Cas-Aktivität auf die Plasmid-Erhaltung in vivo zu verfolgen, wurde im Wesentlichen wie zuvor beschrieben ausgeführt. Kurz gesagt wurden elektrokompente Zellen von *S. pyogenes* auf dieselbe Zelldichte angeglichen und mit 500 ng Plasmid-DNA elektroporiert. Jede Transformation wurde zwei- oder dreimal ausplattiert, und das Experiment wurde für die statistische Analyse dreimal unabhängig mit unterschiedlichen Chargen von kompetenten Zellen durchgeführt. Die Transformations-Wirksamkeiten wurden als CFU (Kolonie-bildende Einheiten, colony forming units) pro µg DNA berechnet. Kontroll-Transformationen wurden mit sterilem Wasser und Rückgratvektor pEC85 durchgeführt.

DNA-Manipulation

[0506] DNA-Manipulationen, einschließlich DNA-Präparation, Amplifikation, Verdau, Ligation, Aufreinigung, Agarose-Gel-Elektrophorese wurden nach Standardverfahren mit geringfügigen Modifikationen durchgeführt. Protospacer-Plasmide für die in vitro-Spaltung sowie *S. pyogenes*-Transformationsassays wurden wie zuvor beschrieben aufgebaut (4). Es wurden zusätzliche, auf pUC19 basierende Protospacer-Plasmide für in vitro-Spaltungsassays erzeugt, indem zusammengelagerte Oligonukleotide zwischen verdauten EcoRI- und BamHI-Schnittstellen in pUC19 ligiert wurden. Das Plasmid, das ein GFP-Gen besitzt, ist zuvor beschrieben worden (41). Es wurden Kits (Qiagen) für die DNA-Aufreinigung und die Plasmid-Präparation verwendet. Die Plasmid-

Mutagenese wurde unter Verwendung von QuikChange® II XL-Kits (Stratagene) oder unter Verwendung des QuikChange site-directed Mutagenesis-Kits (Agilent) durchgeführt. VBC-Biotech Services, Sigma-Aldrich and Integrated DNA Technologies stellten die synthetischen Oligonukleotide und RNAs bereit.

Oligonukleotide für in vitro-Transkriptions-Matrizen

Matrizen für in vitro transkribierte CRISPR Typ II-tracrRNA und crRNAs von *S. pyogenes* (für tracrRNA – PCR auf chr. DNA SF370; für crRNA – Zusammenlagerung von zwei Oligonukleotiden)

T7-tracrRNA (75 nt)

OLEC1521 (F 5' tracrRNA): SEQ ID NO: 340

OLEC1522 (R 3' tracrRNA): SEQ ID NO: 341

T7-crRNA (Matrize)

OLEC2176 (F crRNA-sp1): SEQ ID NO: 342

OLEC2178 (R crRNA-sp1): SEQ ID NO: 343

OLEC2177 (F crRNA-sp2): SEQ ID NO: 344

OLEC2179 (R crRNA-sp2): SEQ ID NO: 345

Matrizen für in vitro-transkribierte *N. meningitidis*-tracrRNA und künstlich erzeugte crRNA-sp2 (für tracrRNA – PCR auf chr. DNA 22491; für crRNA – Zusammenlagerung von zwei Oligonukleotiden)

T7-tracrRNA

OLEC2205 (F predicted 5'): SEQ ID NO: 346

OLEC2206 (R predicted 3'): SEQ ID NO: 347

T7-crRNA (Matrize)

OLEC2209 (F sp2(speM) + *N. m.* repeat): SEQ ID NO: 348

OLEC2214 (R sp2(speM) + *N. m.* repeat): SEQ ID NO: 349

Matrizen für in vitro-transkribierte *L. innocua* tracrRNA und künstlich erzeugte crRNA-sp2 (für tracrRNA – PCR auf chr. DNA Clip11262; für crRNA – Zusammenlagerung von zwei Oligonukleotiden)

T7-tracrRNA

OLEC2203 (F predicted 5'): SEQ ID NO: 350

OLEC2204 (R predicted 3'): SEQ ID NO: 351

T7-crRNA (Matrize)

OLEC2207 (F sp2(speM) + *L. in.* repeat): SEQ ID NO: 352

OLEC2212 (R sp2(speM) + *L. in.* repeat): SEQ ID NO: 353

Oligonukleotide für die Konstruktion von Plasmiden mit Protospacern für in vitro- und in vivo-Untersuchungen

Plasmide für speM (Spacer 2 (CRISPR Type II-A, SF370; Protospacer-Prophage 08232.3 von MGAS8232)-Analyse in vitro und in *S. pyogenes* (Matrize: chr. DNA MGAS8232 oder Plasmide, die speM-Fragmente enthalten)

pEC287

OLEC1555 (F speM): SEQ ID NO: 354

OLEC1556 (R speM): SEQ ID NO: 355

pEC488

OLEC2145 (F speM): SEQ ID NO: 356
OLEC2146 (R speM): SEQ ID NO: 357

pEC370

OLEC1593 (F pEC488 Protospacer 2 A22G): SEQ ID NO: 358
OLEC1594 (R pEC488 Protospacer 2 A22G): SEQ ID NO: 359

pEC371

OLEC1595 (F pEC488 Protospacer 2 T10C): SEQ ID NO: 360
OLEC1596 (R pEC488 Protospacer 2 T10C): SEQ ID NO: 361

pEC372

OLEC2185 (F pEC488 Protospacer 2 T7A): SEQ ID NO: 362
OLEC2186 (R pEC488 Protospacer 2 T7A): SEQ ID NO: 363

pEC373

OLEC2187 (F pEC488 Protospacer 2 A6T): SEQ ID NO: 364
OLEC2188 (R pEC488 Protospacer 2 A6T): SEQ ID NO: 365

pEC374

OLEC2235 (F pEC488 Protospacer 2 A5T): SEQ ID NO: 366
OLEC2236 (R pEC488 Protospacer 2 A5T): SEQ ID NO: 367

pEC375

OLEC2233 (F pEC488 Protospacer 2 A4T): SEQ ID NO: 368
OLEC2234 (R pEC488 Protospacer 2 A4T): SEQ ID NO: 369

pEC376

OLEC2189 (F pEC488 Protospacer 2 A3T): SEQ ID NO: 370
OLEC2190 (R pEC488 Protospacer 2 A3T): SEQ ID NO: 371

pEC377

OLEC2191 (F pEC488 Protospacer 2 PAM G1C): SEQ ID NO: 372
OLEC2192 (R pEC488 Protospacer 2 PAM G1C): SEQ ID NO: 373

pEC378

OLEC2237 (F pEC488 Protospacer 2 PAM GG1, 2CC): SEQ ID NO: 374
OLEC2238 (R pEC488 Protospacer 2 PAM GG1, 2CC): SEQ ID NO: 375

Plasmide für SPy_0700 (Spacer 1(CRISPR Type II-A, SF370; Protospacer-Prophage ø370.1 von SF370)-Analyse in vitro und in *S. pyogenes* (Matrize: chr. DNA SF370 oder Plasmide, die SPy_0700 Fragmente enthalten)

pEC489

OLEC2106 (F Spy_0700): SEQ ID NO: 376
OLEC2107 (R Spy_0700): SEQ ID NO: 377

pEC573

OLEC2941 (F PAM TG1, 2GG): SEQ ID NO: 378

OLEC2942 (R PAM TG1, 2GG): SEQ ID NO: 379

Oligonukleotide zur Verifizierung von Plasmid-Konstrukten und Schnittstellen durch Sequenzierungs-Analyse

Co1E1 (pEC85)

oliRN228 (R-Sequenzierung): SEQ ID NO: 380

speM (pEC287)

OLEC1557 (F-Sequenzierung): SEQ ID NO: 381

OLEC1556 (R-Sequenzierung): SEQ ID NO: 382

repDEG-pAMbeta1 (pEC85)

OLEC787 (F-Sequenzierung): SEQ ID NO: 383

Oligonukleotide für in vitro-Spaltungsassays

crRNA

Spacer 1 crRNA (1-42): SEQ ID NO: 384

Spacer 2 crRNA (1-42): SEQ ID NO: 385

Spacer 4 crRNA (1-42): SEQ ID NO: 386

Spacer 2 crRNA (1-36): SEQ ID NO: 387

Spacer 2 crRNA (1-32): SEQ ID NO: 388

Spacer 2 crRNA (11-42): SEQ ID NO: 389

tracrRNA

(4-89): SEQ ID NO: 390

(15-89): SEQ ID NO: 391

(23-89): SEQ ID NO: 392

(15-53): SEQ ID NO: 393

(15-44): SEQ ID NO: 394

(15-36): SEQ ID NO: 395

(23-53): SEQ ID NO: 396

(23-48): SEQ ID NO: 397

(23-44): SEQ ID NO: 398

(1-26): SEQ ID NO: 399

chimäre RNAs

Spacer 1 – chimäre A: SEQ ID NO: 400

Spacer 1 – chimäre B: SEQ ID NO: 401

Spacer 2 – chimäre A: SEQ ID NO: 402

Spacer 2 – chimäre B: SEQ ID NO: 403

Spacer 4 – chimäre A: SEQ ID NO: 404

Spacer 4 – chimäre B: SEQ ID NO: 405

GFP 1: SEQ ID NO: 406

GFP2: SEQ ID NO: 407

GFP3: SEQ ID NO: 408

GFP4: SEQ ID NO: 409

GFP5: SEQ ID NO: 410

DNA-Oligonukleotide als Substrate für Spaltungsassays (Protospacer in Fettdruck, PAM unterstrichen)

Protospacer 1 – komplementär – WT: SEQ ID NO: 411
 Protospacer 1 – nicht-komplementär – WT: SEQ ID NO: 412
 Protospacer 2 – komplementär – WT: SEQ ID NO: 413
 Protospacer 2 – nicht-komplementär – WT: SEQ ID NO: 414
 Protospacer 4 – komplementär – WT: SEQ ID NO: 415
 Protospacer 4 – nicht-komplementär – WT: SEQ ID NO: 416
 Protospacer 2 – komplementär – PAM1: SEQ ID NO: 417
 Protospacer 2 – nicht-komplementär – PAM1: SEQ ID NO: 418
 Protospacer 2 – komplementär – PAM2: SEQ ID NO: 419
 Protospacer 2 – nicht-komplementär – PAM2: SEQ ID NO: 420
 Protospacer 4 – komplementär – PAM1: SEQ ID NO: 421
 Protospacer 4 – nicht-komplementär – PAM1: SEQ ID NO: 422
 Protospacer 4 – komplementär – PAM2: SEQ ID NO: 423
 Protospacer 4 – nicht-komplementär – PAM2: SEQ ID NO: 424

In vitro-Transkription und Aufreinigung von RNA

[0507] RNA wurde in vitro unter Verwendung des T7 Flash-in vitro-Transkriptions-Kits (Epicentre, Illumina Company) und mittels PCR erzeugten DNA-Matrizen, die eine T7-Promotor-Sequenz trugen, transkribiert. Die RNA wurde im Gel aufgereinigt, und die Qualität wurde vor der Verwendung überprüft. Die Primer, die für die Präparation der RNA-Matrizen von *S. pyogenes* SF370, *Listeria innocua* Clip 11262 und *Neisseria meningitidis* A Z2491 verwendet wurden, sind oben beschrieben.

Protein-Reinigung

[0508] Die für Cas9 kodierende Sequenz (Reste 1–1368) wurde ausgehend von genomischer DNA von *S. pyogenes* SF370 amplifiziert und unter Verwendung einer Ligations-unabhängigen Klonierung (ligation independent cloning, LIC) in einen maßgefertigten, auf pET-basierenden Expressionsvektor insertiert. Das resultierende Fusionskonstrukt enthielt ein N-terminales Hexahistidin-Maltose-bindendes Protein (His6-MBP)-Tag, gefolgt von einer Peptidsequenz, die eine Tabak-Etch-Virus (TEV)-Protease-Spaltungsstelle enthielt. Das Protein wurde in *E. coli*-Stamm BL21-Rosetta 2 (DE3) (EMD Biosciences) exprimiert, in 2xTY-Medium bei 18°C für 16 Stunden angezüchtet, gefolgt von einer Induktion mit 0,2 mM IPTG. Das Protein wurde durch eine Kombination von Affinitätschromatographie-, Ionenaustauschchromatographie- und Größenausschlusschromatographie-Schritten gereinigt. Kurz gesagt wurden die Zellen in 20 mM Tris, pH 8,0, 500 mM NaCl, 1 mM TCEP (supplementiert mit Protease-Inhibitor-Cocktail (Roche)) in einem Homogenisator (Avestin) lysiert. Das geklärte Lysat wurde in einem Batch-Ansatz an Ni-NTA-Agarose (Qiagen) gebunden. Die Säule wurde intensiv mit 20 mM Tris pH 8,0, 500 mM NaCl gewaschen, und das gebundene Protein wurde in 20 mM Tris pH 8,0, 250 mM NaCl, 10% Glycerol eluiert. Der His6-MBP-Affinitäts-Tag wurde durch Spaltung mit TEV-Protease entfernt, während das Protein über Nacht gegen 20 mM HEPES pH 7,5, 150 mM KCl, 1 mM TCEP, 10% Glycerol dialysiert wurde. Das gespaltene Cas9-Protein wurde von dem Fusions-Tag durch Aufreinigung auf einer 5 ml SP-Sepharose-HiTrap-Säule (GE Life Sciences) getrennt und mit einem linearen Gradienten von 100 mM – 1 M KCl eluiert. Das Protein wurde weiter durch Größenaustauschchromatographie auf einer Superdex 200 16/60-Säule in 20 mM HEPES pH 7,5, 150 mM KCl und 1 mM TCEP gereinigt. Eluiertes Protein wurde auf etwa 8 mg/ml konzentriert, in flüssigem Stickstoff schockgefroren und bei –80°C gelagert. D10A-, H840A- und D10A/H840A-Punktmutanten von Cas9 wurden unter Verwendung des QuikChange site-directed Mutagenesis-Kits (Agilent) erzeugt und durch DNA-Sequenzierung bestätigt. Die Proteine wurden aufgereinigt, indem man das gleiche Vorgehen anwendete wie beim Wildtyp-Cas9-Protein.

[0509] Cas9-Orthologe von *Streptococcus thermophilus* (LMD-9, YP_820832.1), *L. innocua* (Clip11262, NP_472073.1), *Campylobacter jejuni* (subsp. *jejuni* NCTC 11168, YP_002344900.1) und *N. meningitidis* (Z2491, YP_002342100.1) wurden als Fusionsproteine His6-MBP (*N. meningitidis* und *C. jejuni*), His6-Thioredoxin (*L. innocua*) und His6-GST (*S. thermophilus*) in BL21 Rosetta (DE3) pLysS-Zellen (Novagen) exprimiert und im Wesentlichen wie für das *S. pyogenes* Cas9 beschrieben aufgereinigt, wobei die folgenden Änderungen vorgenommen wurden. Aufgrund der großen Mengen von co-gereinigten Nukleinsäuren wurden alle vier Cas9-Proteine durch einen zusätzlichen Heparin-Sepharose-Schritt vor der Gel-Filtration gereinigt, wobei das gebundene Protein mit einem linearen Gradienten von 100 mM bis 2 M KCl eluiert wurde. Dies entfernte erfolgreich Nukleinsäure-Kontaminationen von den Proteinen von *C. jejuni*, *N. meningitidis* und *L. innocua*, war jedoch bei der Entfernung der co-gereinigten Nukleinsäuren aus der Cas9-Präparation von *S. thermophilus* nicht

erfolgreich. Alle Proteine wurden auf 1–8 mg/ml in 20 mM HEPES pH 7,5, 150 mM KCl und 1 mM TCEP konzentriert, in flüssigem N₂ schockgefroren und bei –80°C gelagert.

Plasmid-DNA-Spaltungsassays

[0510] Synthetische oder in vitro transkribierte tracrRNA und crRNA wurde vor der Reaktion in einem vorge-lagerten Schritt zusammengelagert, indem diese auf 95°C erhitzt und langsam auf Raumtemperatur abgekühlt wurden. Native oder durch Restriktionsverdau linearisierte Plasmid-DNA (300 ng (etwa 8 nM)) wurde für 60 min bei 37°C mit gereinigtem Cas9-Protein (50–500 nM) und tracrRNA:crRNA-Duplex (50–500 nM, 1:1) in einem Cas9-Plasmid-Spaltungspuffer (20 mM HEPES pH 7,5, 150 mM KCl, 0,5 mM DTT, 0,1 mM EDTA) mit oder ohne 10 mM MgCl₂ inkubiert. Die Reaktionen wurden mit 5 × DNA-Ladepuffer abgestoppt, der 250 mM EDTA enthielt, durch 0,8 oder 1%ige Agarose-Gel-Elektrophorese aufgetrennt und mit Ethidiumbromid-Färbung sichtbar gemacht. Für die Cas9-Mutanten-Spaltungsassays wurden die Reaktionen vor der Beladung des Agarose-Gels mit 5 × SDS Ladepuffer gestoppt (30% Glycerol, 1,2% SDS, 250 mM EDTA).

Metall-abhängiger Spaltungsassay

[0511] Protospacer 2-Plasmid-DNA (5 nM) wurde für eine Stunde bei 37°C mit Cas9 (50 nM) inkubiert, das mit 50 nM tracrRNA:crRNA-sp2 in Spaltungspuffer (20 mM HEPES pH 7,5, 150 mM KCl, 0,5 mM DTT, 0,1 mM EDTA), der mit 1, 5 oder 10 mM MgCl₂, 1 oder 10 mM MnCl₂, CaCl₂, ZnCl₂, CoCl₂, NiSO₄ oder CuSO₄ supplementiert war, vorinkubiert worden war. Die Reaktion wurde durch Zugabe von 5 × SDS Ladepuffer (30% Glycerol, 1,2% SDS, 250 mM EDTA) abgestoppt, durch 1%ige Agarose-Gel-Elektrophorese aufgetrennt und durch Ethidiumbromid-Färbung sichtbar gemacht.

[0512] Cas9 (25 nM) wurde 15 min bei 37°C in Spaltungspuffer (20 mM HEPES pH 7,5, 150 mM KCl, 10 mM MgCl₂, 0,5 mM DTT, 0,1 mM EDTA) mit Duplex-tracrRNA:crRNA-sp2 (25 nM, 1:1) prä-inkubiert, oder mit den beiden RNAs (25 nM), die zuvor nicht zusammengelagert worden waren, und die Reaktion wurde durch Zugabe von Protospacer 2-Plasmid-DNA (5 nM) gestartet. Die Reaktionsmischung wurde bei 37°C inkubiert. In definierten Zeitintervallen wurden Proben aus der Reaktion entnommen, und es wurde 5 × SDS Ladepuffer (30% Glycerol, 1,2% SDS, 250 mM EDTA) zugegeben, um die Reaktion zu stoppen, und die Spaltung wurde durch 1%ige Agarose-Gel-Elektrophorese und Ethidiumbromid-Färbung verfolgt. Das gleiche wurde für den Einzel-Turnover-Reaktionsverlauf ohne Präinkubation von Cas9 und RNA durchgeführt, wobei Protospacer 2-Plasmid-DNA (5 nM) in Spaltungspuffer mit Duplex-tracrRNA:crRNA-sp2 (25 nM) oder mit beiden RNAs (25 nM), die zuvor nicht zusammengelagert worden waren, gemischt und die Reaktion wurde durch Zugabe von Cas9 (25 nM) gestartet. Der prozentuale Anteil der Spaltung wurde durch Densitometrie analysiert, und der Durchschnitt von drei unabhängigen Experimenten wurde gegen die Zeit aufgetragen. Die Daten wurden durch nicht-lineare Regressionsanalyse angepasst, und die Spaltungsraten (k_{obs} [min⁻¹]) wurden berechnet.

Mehrfach-Turnover-Assay

[0513] Cas9 (1 nM) wurde für 15 min bei 37°C mit zuvor zusammengelagertem tracrRNA:crRNA-sp2 (1 nM, 1:1) in Spaltungspuffer (20 mM HEPES pH 7,5, 150 mM KCl, 10 mM MgCl₂, 0,5 mM DTT, 0,1 mM EDTA) prä-inkubiert. Die Reaktion wurde durch Zugabe von Protospacer 2-Plasmid-DNA (5 nM) gestartet. Zu definierten Zeitintervallen wurden Proben entnommen, und die Reaktion wurde durch Zugabe von 5 × SDS-Ladepuffer (30% Glycerol, 1,2% SDS, 250 mM EDTA) gestoppt. Die Spaltungsreaktion wurde durch 1%ige Agarose-Gel-Elektrophorese aufgetrennt, mit Ethidiumbromid gefärbt, und die prozentuale Spaltung wurde durch Densitometrie analysiert. Die Ergebnisse von vier unabhängigen Experimenten wurden gegen die Zeit aufgetragen (min).

Oligonukleotid-DNA-Spaltungs-Assay

[0514] DNA-Oligonukleotide (10 pmol) wurden durch Inkubation mit 5 Einheiten T4-Polynukleotid-Kinase (New England Biolabs) und etwa 3–6 pmol (etwa 20–40 mCi) [γ -³²P]-ATP (Promega) in 1 × T4 Polynukleotid-Kinase-Reaktionspuffer bei 37°C für 30 min in einer 50 μ l-Reaktion radioaktiv markiert. Nach Hitze-Inaktivierung (65°C für 20 min) wurden die Reaktionen durch eine Illustra MicroSpin G-25-Säule (GE Healthcare) gereinigt, um Markierungen, die nicht eingebaut wurden, zu entfernen. Durch Zusammenlagerung von markierten Oligonukleotiden mit äquimolaren Mengen von nicht markiertem, komplementärem Oligonukleotid bei 95°C für 3 min, gefolgt von langsamem Abkühlen auf Raumtemperatur, wurden Duplex-Substrate (100 nM) erzeugt. Für die Spaltungs-Assays wurden tracrRNA und crRNA durch Erhitzen auf 95°C für 30 sec, gefolgt von langsamem Abkühlen auf Raumtemperatur zusammengelagert. Cas9 (500 nM Endkonzentration) wurde mit den zusam-

mengelagerten tracrRNA:crRNA-Duplex (500 nM) in Spaltungs-Assay-Puffer (20 mM HEPES pH 7,5, 100 mM KCl, 5 mM MgCl₂, 1 mM DTT, 5% Glycerol) in einem Gesamtvolumen von 9 µl vorinkubiert. Die Reaktionen wurden von Zugabe von 1 µl Ziel-DNA (10 nM) gestartet und für eine Stunde bei 37°C inkubiert. Die Reaktionen wurden durch Zugabe von 20 µl Lade-Farbstoff (5 mM EDTA, 0,025% SDS, 5% Glycerol in Formamid) gequencht und für 5 min auf 95°C erhitzt. Spaltprodukte wurden auf 12%igen, denaturierenden Polyacrylamid-Gelen, die 7 M Harnstoff enthielten, aufgetrennt und durch Phosphor-Imaging sichtbar gemacht (Storm, GE Life Sciences). Spaltungs-Assays, welche die PAM-Erfordernisse untersuchten (**Fig. 13B**) wurden unter Verwendung von DNA-Duplex-Substraten durchgeführt, die zuvor zusammengelagert, auf einem 8%igen nativen Acrylamid-Gel aufgereinigt, und anschließend an beiden 5'-Ende radioaktiv markiert worden waren. Die Reaktionen wurden konzipiert und analysiert wie es oben beschrieben ist.

Elektrophoretische Mobilitäts-Shift-Assays

[0515] Ziel-DNA-Duplexe wurden gebildet, indem man jeden Strang (10 nmol) in deionisierten Wasser mischte, für 3 min auf 95°C erhitzte und langsam auf Raumtemperatur abkühlte. Alle DNAs wurden auf 8%igen nativen Gelen, die 1 × TBE enthielten, gereinigt. DNA-Banden wurden durch UV-Shadowing sichtbar gemacht, ausgeschnitten und eluiert, indem man Gel-Stückchen in DEPC-behandeltem H₂O einweichte. Eluierte DNA wurde mit Ethanol präzipitiert und in DEPC-behandeltem H₂O gelöst. Die DNA-Proben wurden am 5'-Ende mit [γ-³²P]-ATP markiert, wobei eine T4-Polynukleotid-Kinase (New England Biolabs) für 30 min bei 37°C verwendet wurde. PNK wurde bei 65°C für 20 min hitzedenaturiert, und nicht eingebaute radioaktive Markierung wurde unter Verwendung einer Illustra MicroSpin G-25-Säule (GE Healthcare) entfernt. Die Bindungsassays wurden in Puffer durchgeführt, der 20 mM HEPES pH 7,5, 100 mM KCl, 5 mM MgCl₂, 1 mM DTT und 10% Glycerol in einem Gesamtvolumen von 10 µl enthielt. Die D10A/H840A-Doppelmutante von Cas9 wurde mit äquimolaren Mengen von zuvor zusammengelagerten tracrRNA:crRNA-Duplex programmiert und von 100 pM bis 1 µM titriert. Radioaktiv markierte DNA wurde in einer Endkonzentration von 20 pM zugegeben. Die Proben wurden für 1 Stunde bei 37°C inkubiert und bei 4°C auf einem 8%igen nativen Polyacrylamid-Gel, das 1 × TBE und 5 mM MgCl₂ enthielt, aufgetrennt. Die Gele wurden getrocknet, und die DNA wurde durch Phosphor-Imaging sichtbar gemacht.

In silico-Analyse von DNA- und Protein-Sequenzen

[0516] Das Vektor-NTI-Packet (Invitrogen) wurde für die DNA-Sequenz-Analyse (Vector NTI) und die vergleichende Sequenz-Analyse von Proteinen (AlignX) verwendet.

In silico-Modeling der RNA-Struktur und co-Faltung.

[0517] In silico-Vorhersagen wurden unter Verwendung der Vienna-RNA-Package-Algorithmen (42, 43) durchgeführt. RNA-Sekundärstrukturen und co-Faltungsmodelle wurden mit RNAfold bzw. RNAcifold vorhergesagt und mit VARNA (44) sichtbar gemacht.

ERGEBNISSE

[0518] Bakterien und Archaeen haben durch RNA vermittelte, adaptive Verteidigungssysteme entwickelt, die "clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-assoziiert (Cas)" genannt werden und die Organismen vor eindringenden Viren und Plasmiden schützen (1–3). Wir zeigen, dass in einer Untergruppe dieser Systeme die reife crRNA, welche mit einer trans-aktivierenden crRNA (tracrRNA) Basengepaart ist, eine Struktur aus zwei RNAs bildet, welche das CRISPR-assoziierte Protein Cas9 so steuert, dass es Doppelstrang(ds)-Brüche in der Ziel-DNA einbringt. An Stellen, die komplementär zu der crRNA-Guide-Sequenz sind, spaltet die Cas9-HNH-Nuklease-Domäne den komplementären Strang, während die Cas9-RuvC-ähnliche Domäne den nicht-komplementären Strang spaltet. Die duale tracrRNA:crRNA steuert, wenn sie als einzelne RNA-Chimäre erstellt wird, auch eine Sequenz-spezifische Cas9-dsDNA-Spaltung. Diese Studien offenbaren eine Familie von Endonukleasen, die duale RNAs für die ortsspezifische DNA-Spaltung verwenden und verweisen auf die Fähigkeit, das System für ein RNA-programmierbares Genom-Editing zu nutzen.

[0519] CRISPR/Cas-Verteidigungssysteme benutzen kleine RNAs für die Sequenz-spezifische Detektion und das Silencing von fremden Nukleinsäuren. CRISPR/Cas-Systeme sind aus Cas-Genen zusammengesetzt, die in (einem) Operon(s) und (einem) CRISPR-Array(s) organisiert sind, welche aus Sequenzen bestehen, die auf das Genom gerichtet sind (Spacer genannt), welche von identischen Repeats unterbrochen sind (1–3). Eine CRISPR/Cas-vermittelte Immunität tritt in drei Schritten auf. In der adaptiven Phase reagieren Bakterien und Archaeen, die einen oder mehrere CRISPR-Loci umfassen, auf eine Bedrohung durch einen Virus oder

ein Plasmid, indem sie kurze Fragmente fremder Sequenz (Protospacer) am proximalen Ende des CRISPR-Arrays in das Wirtschromosom integrieren (1–3). In den Expressions- und Interferenzphasen führt die Transkription des Repeat-Spacer-Elements in Vorläufer-CRISPR-RNA(prä-crRNA)-Moleküle und eine nachfolgende enzymatische Spaltung die kurzen crRNAs, die sich mit den komplementären Protospacer-Sequenzen von eindringenden Virus- oder Plasmid-Zielen paaren kann (4–11). Die Zielerkennung durch crRNAs steuert das Silencing der fremden Gene durch Cas-Proteine, die gemeinsam mit den crRNAs funktionieren (10, 12–20).

[0520] Es gibt drei Arten von CRISPR/Cas-Systemen (21–23). Die Systeme vom Typ I und Typ III weisen einige übergeordnete Merkmale auf: spezialisierte Cas-Endonukleasen prozessieren die prä-crRNAs und jede crRNA lagert sich nach der Reifung zu einem großen Multi-Cas-Proteinkomplex zusammen, der in der Lage ist, Nukleinsäuren, die komplementär zu der crRNA sind, zu erkennen und zu spalten. Im Gegensatz dazu prozessieren Systeme vom Typ II prä-crRNAs durch einen anderen Mechanismus, bei dem eine trans-aktivierende crRNA (tracrRNA), die komplementär zu den Repeat-Sequenzen in der prä-crRNA ist, die Prozessierung durch die doppelsträngige (ds) RNA-spezifische Ribonuklease-RNase III in Gegenwart des Cas9(früher Csn1 genannt)-Proteins auslöst (**Fig. 15**) (4, 24). Es wird angenommen, dass Cas9 das einzige Protein ist, das für das crRNA-gesteuerte Silencing von Fremd-DNA verantwortlich ist (25–27).

[0521] Wir zeigen, dass Cas9-Proteine in Systemen des Typs II eine Familie von Enzymen darstellen, die eine Basen-gepaarte Struktur erfordern, welche zwischen der aktivierenden tracrRNA und der für das Targeting sorgenden crRNA gebildet wird, um die Ziel-dsDNA zu spalten. Die ortsspezifische Spaltung erfolgt an Stellen, die sowohl durch die Basenpaarungs-Komplementarität zwischen der crRNA und der Ziel-Protospacer-DNA als auch durch ein kurzes Motiv [als Protospacer-benachbartes Motiv (protospacer adjacent motif), (PAM) bezeichnet], welches der komplementären Region in der Ziel-DNA benachbart ist, bestimmt werden. Unsere Studie zeigt ferner, dass die Cas9-Endonuklease-Familie mit einzelnen RNA-Molekülen programmiert werden kann, um spezifische DNA-Stellen zu spalten, wodurch die Entwicklung eines einfachen und vielseitigen durch RNA-gesteuerten Systems zur Erzeugung von dsRNA-Brüchen für das Genom-Targeting und das Genom-Editing bereitgestellt wird.

Cas9 ist eine DNA-Endonuklease, die durch zwei RNAs geführt wird.

[0522] Es wurde angenommen, dass Cas9, bei dem es sich um das kennzeichnende Protein von Systemen des Typs II handelt, sowohl an der crRNA-Reifung als auch an der von crRNA geführten DNA-Interferenz beteiligt ist (**Fig. 15**) (4, 25–27). Cas9 ist an der crRNA-Reifung beteiligt (4), wobei jedoch seine direkte Beteiligung an der Zerstörung der Ziel-DNA nicht erforscht wurde. Um zu testen, ob und auf welche Weise Cas9 in der Lage sein könnte, Ziel-DNA zu spalten, haben wir ein Überexpressionssystem verwendet, um Cas9-Protein zu reinigen, welches aus dem Erreger *Streptococcus pyogenes* abgeleitet wurde (**Fig. 16**, siehe ergänzende Materialien und Methoden), und wir haben die Fähigkeit desselben zur Spaltung einer Plasmid-DNA oder eines Oligonukleotid-Duplex geprüft, der eine Protospacer-Sequenz enthielt, die komplementär zu einer reifen crRNA war, und ein angenommenes PAM. Wir haben herausgefunden, dass reife crRNA alleine nicht in der Lage war, die Cas9-katalysierte Spaltung von Plasmid-DNA zu steuern (**Fig. 10A** und **Fig. 17A**). Die Zugabe von tracrRNA, die sich mit der Repeat-Sequenz der crRNA paaren kann und für die crRNA-Reifung in diesem System essentiell ist, brachte Cas9 dazu, Plasmid-DNA zu spalten (**Fig. 10A** und **Fig. 17A**). Die Spaltungsreaktion erforderte sowohl Magnesium als auch die Gegenwart einer crRNA-Sequenz, die komplementär zu der DNA war; eine crRNA, die zur tracrRNA-Basenpaarung in der Lage war, jedoch eine unverwandte Ziel-DNA-Bindungssequenz enthielt, unterstützte die Cas9-katalysierte Plasmid-Spaltung nicht (**Fig. 10A**; **Fig. 17A**, vergleiche crRNA-sp2 mit crRNA-sp1; und **Fig. 18A**). Wir erhielten ähnliche Ergebnisse mit einem kürzeren linearen dsDNA-Substrat (**Fig. 10B** und **Fig. 17, B und C**). Somit ist die trans-aktivierende tracrRNA eine kleine nicht-kodierende RNA mit zwei wichtigen Funktionen: Auslösen der Prozessierung von prä-crRNA durch das Enzym RNase III (4) und anschließende Aktivierung der crRNA-gesteuerten DNA-Spaltung durch Cas9.

[0523] Die Spaltung sowohl von Plasmid-DNA als auch von kurzer linearer dsDNA durch ein tracrRNA:crRNA-gesteuertes Cas9 ist ortsspezifisch (**Fig. 10, C bis E** und **Fig. 19, Fig. A und B**). Die Spaltung von Plasmid-DNA erzeugte glatte Enden an einer Position drei Basenpaare stromaufwärts der PAM-Sequenz (**Fig. 10, C und E** und **Fig. 19A und C**) (26). In ähnlicher Weise wird der DNA-Strang, der komplementär zu der Ziel-Bindungssequenz in der crRNA (dem komplementären Strang) ist, innerhalb von kurzen dsDNA-Duplexen an einer Stelle drei Basenpaare stromaufwärts der PAM-Sequenz gespalten (**Fig. 10, D und E** und **Fig. 19B und C**). Der nicht-komplementäre DNA-Strang wird an einer oder an mehreren Stellen innerhalb von drei bis acht Basenpaaren stromaufwärts vom PAM gespalten. Eine weitere Untersuchung ergab, dass der nicht-komplementäre Strang zuerst endonukleolytisch gespalten und anschließend durch eine 3'-5'-Exonuklease-Aktivität zurechtgeschnitten wird (**Fig. 18B**). Die Spaltungsraten durch Cas9 unter Bedingungen eines Einzel-Turnovers reichten von

0,3 bis 1 min⁻¹, was zu den Werten von Restriktions-Endonukleasen vergleichbar ist (**Fig. 20A**), wobei die Inkubation von Wildtyp (WT) Cas9-tracrRNA:crRNA-Komplex mit 5-fachem molaren Überschuss an Substrat-DNA den Beweis erbrachte, dass das von einer dualen RNA geführte Cas9 ein Mehrfach-Turnover-Enzym ist (**Fig. 20B**). Im Gegensatz zum CRISPR-Typ I-Kaskade-Komplex (18), spaltet Cas9 sowohl linearisierte Plasmide als auch Plasmide in Supercoiled-Form (**Fig. 10A** und **Fig. 11A**). Somit kann ein eindringendes Plasmid prinzipiell von Cas9-Proteinen, die mit verschiedenen crRNAs programmiert sind, mehrfach gespalten werden.

[0524] Fig. 10(A). Cas9 wurde mit einer crRNA-sp2 (eine crRNA, die eine Spacer 2-Sequenz enthielt) aus 42 Nukleotiden in Gegenwart oder in Abwesenheit einer tracrRNA aus 75 Nukleotiden programmiert. Der Komplex wurde zu zirkulärer oder mit XhoI linearisierter Plasmid-DNA gegeben, die eine Sequenz umfasste, welche komplementär zu Spacer 2 war, sowie ein funktionsfähiges PAM. crRNA-sp1, Spezifitäts-Kontrolle; M, DNA-Marker; kbp, Kilobasenpaare. Siehe **Fig. 17A**. (B) Cas9 wurde mit crRNA-sp2 und tracrRNA (Nukleotide 4–89) programmiert. Der Komplex wurde mit doppelsträngiger oder einzelsträngiger DNA inkubiert, die eine Sequenz umfasste, welche komplementär zu Spacer 2 war, sowie ein funktionsfähiges PAM (4). Die komplementären oder nicht-komplementären Stränge der DNA wurden am 5'-Ende radioaktiv markiert und mit einem nicht-markierten Partnerstrang zusammengelagert. nt, Nukleotide. Siehe **Fig. 17, B** und **C**. (C) Sequenz-Analyse von Spaltprodukten aus **Fig. 10A**. Die Termination der Primer-Extension in der Sequenzierungsreaktion verweist auf die Position der Spaltungsstelle. Der 3'-terminale A-Überhang (Sternchen) ist ein Artefakt der Sequenzierungsreaktion. Siehe **Fig. 19, A** und **C**. (D) Die Spaltprodukte von **Fig. 10B** wurden mit Größenmarkern, die am 5'-Ende markiert waren und von den komplementären und nicht-komplementären Strängen des Ziel-DNA-Duplex abgeleitet waren, analysiert. M, Marker; P, Spaltprodukt. Siehe **Fig. 19, B** und **C**. (E) Schematische Darstellung der Sequenzen von tracrRNA, crRNA-sp2 und Protospacer 2-DNA. Regionen von crRNA-Komplementarität zur tracrRNA (überstrichen) und die Protospacer-DNA (unterstrichen) sind dargestellt. Die PAM-Sequenz ist markiert; Spaltungsstellen, die in (C) und (D) kartiert wurden, sind durch weiß gefüllte Pfeile (C), einen schwarz gefüllten Pfeil [(D), komplementärer Strang] und einen schwarzen Balken [(D), nicht-komplementärer Strang] dargestellt.

[0525] Fig. 15 zeigt den durch RNA vermittelten CRISPR/Cas-Immunweg vom Typ II. Die Expressions- und Interferenzschritte sind in der Zeichnung dargestellt. Die CRISPR/Cas-Loci vom Typ II bestehen aus einem Operon aus 4 Genen, welches für die Proteine Cas9, Cas1, Cas2 und Csn2 kodieren, einem CRISPR-Array, der aus einer Leader-Sequenz gefolgt von identischen Repeats (schwarze Rechtecke) besteht, die von einzigartigen, auf das Genom gerichteten Spacern (Diamanten) unterbrochen sind, und einer Sequenz, die für die trans-aktivierende tracrRNA kodiert. Vorliegend dargestellt ist der Typ II-CRISPR/Cas-Locus von *S. pyogenes* SF370 (Zugriffsnummer NC_002737) (4). Experimentell bestätigte Promotoren und transkriptionale Terminatoren in diesem Locus sind dargestellt (4). Der CRISPR-Array wird als Vorläufer-CRISPR-RNA (prä-crRNA) -Molekül transkribiert und durchläuft einen Reifungsprozess, der spezifisch für Systeme vom Typ II ist (4). In *S. pyogenes* SF370 wird die tracrRNA in Form von zwei primären Transkripten mit einer Länge von 171 und 89 nt transkribiert, die Komplementarität zu jedem Repeat der prä-crRNA aufweisen. Das erste Prozessierungs-Ereignis umfasst die Paarung von tracrRNA und prä-crRNA, wodurch ein RNA-Duplex gebildet wird, der erkannt wird und durch die Housekeeping-Endonuklease-RNase III in Gegenwart des Cas9-Proteins gespalten wird. Die RNase III-vermittelte Spaltung des RNA-Duplex erzeugt eine prozessierte tracrRNA aus 75 nt und eine intermediäre crRNA aus 66 nt, die aus einer zentralen Region besteht, welche eine Sequenz eines Spacers umfasst, flankiert von Teilen der Repeat-Sequenz. Ein zweites Prozessierungs-Ereignis, welches von (einer) unbekannten Ribonuklease(n) katalysiert wird, führt zur Bildung von reifen crRNAs mit einer Länge von 39 bis 42 nt, welche aus 5'-terminaler, vom Spacer abgeleiteter Guide-Sequenz und vom Repeat-abgeleiteter 3'-terminaler Sequenz bestehen. Nach den ersten und zweiten Prozessierungs-Ereignissen bleibt die reife tracrRNA mit den reifen crRNAs gepaart und bindet an das Cas9-Protein. In diesem tertiären Komplex agiert die duale tracrRNA:crRNA-Struktur als Guide-RNA, welche die Endonuklease Cas9 an die verwandte Ziel-DNA bringt. Die Zielerkennung durch den Cas9-tracrRNA:crRNA-Komplex wird durch Scannen des eindringenden DNA-Moleküls auf Homologie zwischen der Protospacer-Sequenz in der Ziel-DNA und der vom Spacer abgeleiteten Sequenz in der crRNA ausgelöst. Neben der Komplementarität zwischen DNA-Protospacer und crRNA-Spacer erfordert das DNA-Targeting das Vorhandensein eines kurzen Motivs (NGG, wobei N jedes Nukleotid sein kann) in Nachbarschaft zum Protospacer (Protospacer-benachbartes Motiv, PAM). Nach der Paarung zwischen der dualen RNA und der Protospacer-Sequenz wird eine R-Schleife gebildet, und Cas9 bringt anschließend einen Doppelstrang-Bruch (DSB) in die DNA ein. Die Spaltung von Ziel-DNA durch Cas9 erfordert zwei katalytische Domänen im Protein. An einer spezifischen Stelle relativ zum PAM spaltet die HNH-Domäne den komplementären Strang der DNA, während die RuvC-ähnliche Domäne den nicht-komplementären Strang spaltet.

[0526] Fig. 16(A) *S. pyogenes*-Cas9 wurde in *E. coli* als Fusionsprotein exprimiert, das einen N-terminalen His6-MBP-Tag enthielt, und dieses wurde durch eine Kombination von Affinitäts-, Ionenaustausch- und Größenausschlusschromatographie aufgereinigt. Der Affinitäts-Tag wurde durch TEV-Protease-Spaltung nach dem Affinitätsaufreinigungs-Schritt entfernt. Gezeigt ist ein Chromatogramm des finalen Größenausschlusschromatographie-Schritts auf einer Superdex 200(16/60)-Säule. Cas9 eluiert als einzelner monomerer Peak, der frei von kontaminierenden Nukleinsäuren ist, wie durch das Verhältnis der Absorptionswerte bei 280 und 260 nm abgeschätzt werden kann. Inset; eluierte Fraktionen wurden durch SDS-PAGE auf einem 10%igen Polyacrylamid-Gel getrennt und mit SimplyBlue Safe Stain (Invitrogen) gefärbt. (B) SDS-PAGE-Analyse von gereinigten Cas9-Orthologen. Cas9-Orthologe wurden wie im Teil "Ergänzende Materialien und Methoden" beschrieben, aufgereinigt. 2,5 µg von jedem gereinigten Cas9 wurden auf einem 4–20%igen Polyacrylamid-Gradienten-Gel analysiert und mit SimplyBlue Safe Stain (Invitrogen) gefärbt.

[0527] Fig. 17 (siehe auch **Fig. 10**). Die Protospacer 1-Sequenz stammt von *S. pyogenes* SF370 (M1) SPy_0700, dem Ziel der *S. pyogenes* SF370 crRNA_{sp1} (4). Hier wurde die Sequenz des Protospacers 1 manipuliert, indem das PAM von einer nicht funktionsfähigen Sequenz (TTG) in eine funktionsfähige Sequenz (TGG) geändert wurde. Die Protospacer 4-Sequenz stammt von *S. pyogenes* MGAS10750 (M4) MGAS10750_Spy1285, dem Ziel der *S. pyogenes* SF370 crRNA-sp4 (4). (A) Protospacer 1-Plasmid-DNA-Spaltung, die durch verwandte tracrRNA:crRNA-Duplexe geleitet wurde. Die Spaltprodukte wurden durch Agarose-Gel-Elektrophorese aufgetrennt und durch Ethidiumbromid-Färbung sichtbar gemacht. M, DNA-Marker; Fragment-Größen in Basenpaaren sind angegeben. (B) Protospacer 1-Oligonukleotid-DNA-Spaltung, die durch einen verwandten tracrRNA:crRNA-sp1-Duplex geführt wurde. Die Spaltungsprodukte wurden durch denaturierende Polyacrylamid-Gel-Elektrophorese aufgetrennt und durch Phosphor-Imaging sichtbar gemacht. Die Fragment-Größen in Nukleotiden sind angegeben. (C) Protospacer 4-Oligonukleotid-DNA-Spaltung, die durch einen verwandten tracrRNA:crRNA-sp4-Duplex geführt wurde. Die Spaltungsprodukte wurden durch denaturierende Polyacrylamid-Gel-Elektrophorese getrennt und durch Phosphor-Imaging sichtbar gemacht. Die Fragment-Größen in Nukleotiden sind angegeben. (A, B, C) Die Experimente in (A) wurden wie in **Fig. 10A** durchgeführt; in (B) und in (C) wie in **Fig. 10B**. (B, C) Eine schematische Darstellung der tracrRNA:crRNA Ziel-DNA-Interaktion ist nachfolgend gezeigt. Die Regionen von crRNA-Komplementarität zur tracrRNA und die Protospacer-DNA sind mit Oberstrichen bzw. Unterstrichen dargestellt. Die PAM-Sequenz ist markiert.

[0528] Fig. 18 (siehe auch **Fig. 10**). (A) Protospacer 2-Plasmid-DNA wurde in Gegenwart von verschiedenen Konzentrationen von Mg²⁺, Mn²⁺, Ca²⁺, Zn²⁺, Co²⁺, Ni²⁺ oder Cu²⁺ mit Cas9 inkubiert, das mit tracrRNA:crRNA-sp2 komplexiert war. Die Spaltungsprodukte wurden durch Agarose-Gel-Elektrophorese aufgetrennt und durch Ethidiumbromid-Färbung sichtbar gemacht. Die Plasmid-Formen sind angegeben. (B) Ein Protospacer 4-Oligonukleotid-DNA-Duplex, der ein PAM-Motiv enthielt, wurde zusammengelagert und vor der radioaktiven Markierung an beiden 5'-Enden aus dem Gel gereinigt. Der Duplex (Endkonzentration 10 nM) wurde mit Cas9 inkubiert, das mit tracrRNA (Nukleotide 23–89) und crRNA-sp4 (500 nm Endkonzentration, 1:1) programmiert. Zu den angegebenen Zeitpunkten (min) wurden 10 µl-Aliquots der Spaltungsreaktion mit Formamid-Puffer gequench, der 0,025% SDS und 5 mM EDTA enthielt, und diese wurden durch denaturierende Polyacrylamid-Gel-Elektrophorese wie in **Fig. 10B** dargestellt analysiert. Die Größen in Nukleotiden sind angegeben.

[0529] Fig. 19(A) Kartierung der Protospacer 1-Plasmid-DNA-Spaltung. Spaltprodukte von **Fig. 17A** wurden, wie es in **Fig. 10C** gezeigt ist, durch Sequenzierung analysiert. Es ist zu beachten, dass der 3'-terminale A-Überhang (Sternchen) ein Artefakt der Sequenzierungsreaktion ist. (B) Kartierung der Protospacer 4-Oligonukleotid-DNA-Spaltung. Spaltprodukte von **Fig. 17C** wurden gemeinsam mit 5'-endmarkierten Oligonukleotid-Größenmarkern, die von den komplementären und den nicht-komplementären Strängen der Protospacer 4-Duplex-DNA abgeleitet waren, mittels denaturierender Polyacrylamid-Gel-Elektrophorese analysiert. M, Marker; P, Spaltprodukt. Bahnen 1–2: komplementärer Strang. Bahnen 3–4: nicht-komplementärer Strang. Die Fragment-Größen in Nukleotiden sind angegeben. (C) Schematische Darstellungen von tracrRNA, crRNA-sp1 und Protospacer1-DNA-Sequenzen (oben) und tracrRNA, crRNA_{sp4} und Protospacer 4-DNA-Sequenzen (unten). tracrRNA:crRNA bildet eine duale RNA-Struktur, die auf komplementäre Protospacer-DNA gerichtet ist, indem eine crRNA-Protospacer-DNA-Paarung erfolgt. Die Region der crRNA, die komplementär zur tracrRNA und zur Protospacer-DNA sind, werden mit Oberstrichen bzw. Unterstrichen dargestellt. Die Spaltstellen in den komplementären und nicht-komplementären DNA-Strängen, welche in (A) (oben) und (B) (unten) kartiert sind, werden mit Pfeilen dargestellt (A und B, komplementärer Strang) bzw. mit einem schwarzen Balken (B, nicht komplementärer Strang) über den Sequenzen.

[0530] Fig. 20(A) Einzel-Turnover-Kinetik von Cas9 unter verschiedenen Bedingungen von prä-Annealing und Protein-RNA-Vorinkubation. Protospacer 2-Plasmid-DNA wurde entweder mit Cas9 inkubiert, das mit zuvor zusammengelagerten tracrRNA:crRNA-sp2 prä-inkubiert worden war (o), mit Cas9, das nicht mit zuvor zu-

sammengelagerten tracrRNA:crRNA-sp2 prä-inkubiert worden war (•), Cas9, das mit zuvor nicht zusammen-
gelagerten tracrRNA:crRNA-sp2 prä-inkubiert worden war (□) oder Cas9, das nicht mit nicht zuvor zusammen-
gelagerten RNAs prä-inkubiert worden war (■). Die Spaltungsaktivität wurde in einer zeitabhängigen Weise
verfolgt und durch Agarose-Gel-Elektrophorese mit anschließender Ethidiumbromid-Färbung analysiert. Der
durchschnittliche Prozentwert der Spaltung aus drei unabhängigen Experimenten ist gegen die Zeit (min) auf-
getragen und durch eine nicht-lineare Regression angepasst. Die berechneten Spaltungsraten (k_{obs}) sind in der
Tabelle gezeigt. Die Ergebnisse legen nahe, dass die Bindung von Cas9 an die RNAs unter den untersuchten
Bedingungen nicht geschwindigkeitslimitierend ist. Die Plasmid-Formen sind angegeben. Die erhaltenen (k_{obs})-
Werte sind vergleichbar mit denen von Restriktionsendonukleasen und sind üblicherweise in der Größenord-
nung von 1–10 pro Minute (45–47). (B) Cas9 ist eine mehrfach-Turnover-Endonuklease. Cas9, das mit Duplex-
tracrRNA:crRNA-sp2 (1 nM, 1:1:1 – mit grauer Linie in der Graphik dargestellt) beladen war, wurde mit einem
5-fachen Überschuss nativer Protospacer 2-Plasmid-DNA inkubiert. Die Spaltung wurde durch Entnahme von
Proben aus der Reaktion in bestimmten Zeitabständen (0–120 min) verfolgt, gefolgt von einer Analyse mittels
Agarose-Gel-Elektrophorese (oben) und einer Bestimmung der Menge des Spaltprodukts (nM) (unten). Die
Standardabweichungen von drei unabhängigen Experimenten sind angegeben. Im untersuchten Zeitintervall
war 1 nM Cas9 in der Lage, etwa 2,5 nM Plasmid-DNA zu spalten.

[0531] Cas9 enthält Domänen, die homolog sowohl zur HNH- als auch zur RuvC-Endonuklease homolog
sind (**Fig. 11A** und **Fig. 3**) (21–23, 27, 28). Wir haben Cas9-Varianten erstellt und gereinigt, die inaktivierende
Punkt-Mutationen in den katalytischen Resten entweder der HNH-Domäne oder der RuvC-ähnlichen Domäne
aufweisen (**Fig. 11A** und **Fig. 3**) (23, 27). Die Inkubation dieser Cas9-Varianten-Proteine mit nativer Plasmid-
DNA zeigte, dass von dualer RNA geführte Cas9-Mutanten-Proteine zu offenkettigen zirkulären (nicked) Plas-
miden führten, während der WT-Cas9-Protein-tracrRNA:crRNA-Komplex ein lineares DNA-Produkt erzeugte
(**Fig. 10A** und **Fig. 11A** und **Fig. 17A** und **Fig. 25A**). Dieses Ergebnis zeigt, dass die HNH- und RuvC-ähn-
lichen Domänen von Cas9 jeweils einen Plasmid-DNA-Strang spalten. Um zu bestimmen, welcher Strang
der Ziel-DNA durch die jeweilige katalytische Domäne von Cas9 gespalten wird, inkubierten wir Mutanten-
Cas9-tracrRNA:crRNA-Komplexe mit kurzen dsDNA-Substraten, in denen entweder der komplementäre oder
der nicht-komplementäre Strang an seinem 5'-Ende radioaktiv markiert war. Die resultierenden Spaltprodukte
verwiesen darauf, dass die HNH-Domäne von Cas9 den komplementären DNA-Strang spaltet, während die
RuvC-ähnliche Domäne von Cas9 den nicht-komplementären DNA-Strang spaltet (**Fig. 11B** und **Fig. 21B**).

[0532] **Fig. 11A** (oben) Schematische Darstellung der Cas9-Domänen-Struktur, die die Positionen der Do-
mänen-Mutationen zeigt. D10A, Asp10 → A1a10; H840A, His840 → A1a840. Komplexe von WT- oder Mutan-
ten-Nuklease-Cas9-Protein mit tracrRNA:crRNA-sp2 wurden auf Endonuklease-Aktivität untersucht, wie es in
Fig. 10A gezeigt ist. (B) Komplexe von WT-Cas9- oder Nuklease-Domäne-Mutanten mit tracrRNA und crRNA-
sp2 wurden auf Aktivität getestet, wie es in **Fig. 10B** gezeigt ist.

[0533] **Fig. 3.** Die Aminosäuresequenz von Cas9 von *S. pyogenes* (SEQ ID NO: 8) ist gezeigt. Cas9/Csn1-
Proteine von verschiedenen Spezies haben zwei Domänen, die Motive einschließen, welche sowohl zur HNH-
als auch zur RuvC-Endonuklease homolog sind. (A) Motive 1–4 (die Motivnummern sind an der linken Seite
der Sequenz dargestellt) sind für das Cas9/Csn1 von *S. pyogenes* gezeigt. Die drei vorhergesagten RuvC-
ähnlichen Motive (1, 2, 4) und das vorhergesagte HNH-Motiv (3) sind mit Oberstrichen dargestellt. Die Reste
Asp10 und His840, die in der Studie gegen Alanin ausgetauscht wurden, sind mit Sternen oberhalb der Se-
quenz hervorgehoben. Die unterstrichenen Reste sind zwischen den Cas9-Proteinen aus unterschiedlichen
Spezies hochgradig konserviert. Mutationen in den unterstrichenen Resten haben wahrscheinlich funktionale
Auswirkungen auf die Cas9-Aktivität. Es ist zu beachten, dass in der vorliegenden Studie die Kopplung von
zwei Nuklease-ähnlichen Aktivitäten experimentell gezeigt ist (**Fig. 11** und **Fig. 21**). (B) Domänen 1 (Amino-
säuren 7–166) und 2 (Aminosäuren 731–1003), welche die Motive 1–4 umfassen, sind für das Cas9/Csn1 von
S. pyogenes gezeigt. Für zusätzliche Informationen wird auf Tabelle 1 und **Fig. 5** verwiesen.

[0534] **Fig. 21.** Protospacer-DNA-Spaltung durch verwandte tracrRNA:crRNA-gesteuerte Cas9-Mutanten, die
Mutationen in der HNH- oder der RuvC-ähnlichen Domäne aufweisen. (A) Protospacer1-Plasmid-DNA-Spal-
tung. Dieses Experiment wurde wie in **Fig. 11A** dargestellt durchgeführt. Die Konformationen und Größen der
Plasmid-DNA sind in Basenpaaren angegeben. (B) Protospacer 4-Oligonukleotid-DNA-Spaltung. Dieses Ex-
periment wurde wie in **Fig. 11B** dargestellt durchgeführt. Die Größen in Nukleotiden sind angegeben.

Anforderungen an die duale RNA für die Bindung und Spaltung der Ziel-DNA

[0535] Eine tracrRNA kann für die Bindung von Ziel-DNA und/oder für die Stimulation der Nuklease-Aktivi-
tät von Cas9 stromabwärts der Zielerkennung erforderlich sein. Um zwischen diesen Möglichkeiten zu unter-

scheiden, haben wir einen elektrophoretischen Mobilitäts-Shift-Assay verwendet, um die Bindung der Ziel-DNA durch katalytisch inaktive Cas9 in Gegenwart oder in Abwesenheit von crRNA und/oder tracrRNA zu verfolgen. Die Zugabe von tracrRNA verstärkte die Bindung der Ziel-DNA durch Cas9 wesentlich, während wir mit Cas9 allein oder mit Cas9-crRNA eine geringe spezifische DNA-Bindung beobachteten (**Fig. 22**). Dies verweist darauf, dass die tracrRNA für die Erkennung der Ziel-DNA erforderlich ist, möglicherweise indem sie die crRNA für die Interaktion mit dem komplementären Strang der Ziel-DNA in geeigneter Weise ausrichtet. Die vorhergesagte tracrRNA:crRNA-Sekundärstruktur umfasst eine Basenpaarung zwischen den 22 Nukleotiden am 3'-Terminus der crRNA und einem Segment nahe des 5'-Endes der reifen tracrRNA (**Fig. 10E**). Diese Interaktion erzeugt eine Struktur, in der die 5'-terminalen 20 Nukleotide der crRNA, welche hinsichtlich ihrer Sequenz in unterschiedlichen crRNAs abweichen, für die Bindung der Ziel-DNA verfügbar sind. Der Großteil der tracrRNA stromabwärts der crRNA-Basenpaarungs-Region ist frei, um (eine) zusätzliche RNA-Struktur(en) zu bilden und/oder mit Cas9 oder der Ziel-DNA-Stelle zu interagieren. Um zu bestimmen, ob die gesamte Länge der tracrRNA für die ortsspezifische Cas9-katalysierte DNA-Spaltung erforderlich ist, haben wir Cas9-tracrRNA:crRNA-Komplexe getestet, indem wir Komplexe rekonstituierten, wobei reife crRNA vollständiger Länge (42 Nukleotide) und verschiedene trunkierte Formen von tracrRNA verwendet wurden, der Sequenzen an ihren 5'- oder 3'-Ende fehlte. Diese Komplexe wurden auf die Spaltung untersucht, wobei eine kurze Ziel-dsDNA verwendet wurde. Eine erheblich trunkierte Version der tracrRNA, welche die Nukleotide 23–48 der native Sequenz umfasste, war in der Lage, eine robuste, von einer dualen RNA geführte, durch Cas9 katalysierte DNA-Spaltung zu fördern (**Fig. 12, A und C, und Fig. 23, A und B**). Die Trunkierung der crRNA an einem beliebigen Ende zeigte, dass die Cas9-katalysierte Spaltung in Gegenwart von tracrRNA mit crRNAs ausgelöst werden konnte, denen die 3'-terminalen 10 Nukleotide fehlten (**Fig. 12, B und C**). Im Gegensatz dazu verhinderte eine Deletion von 10 Nukleotiden am 5'-Ende der crRNA die DNA-Spaltung durch Cas9 (**Fig. 12B**). Wir haben ferner Cas9-Orthologe aus verschiedenen bakteriellen Spezies auf ihre Fähigkeit zur Unterstützung der *S. pyogenes*-tracrRNA:crRNA-geführten DNA-Spaltung untersucht. Im Gegensatz zu eng verwandten *S. pyogenes*-Cas9-Orthologen waren stärker entfernt verwandte Orthologe in der Spaltungsreaktion nicht funktionsfähig (**Fig. 24**). In ähnlicher Weise war ein Cas9 von *S. pyogenes*, die von tracrRNA:crRNA-Duplexen geführt wurde, welche von stärker entfernten Systemen abstammten, nicht in der Lage, DNA wirksam zu spalten (**Fig. 24**). Die Spezies-Spezifität einer DNA-Spaltung, die von einer dualen RNA geführt wurde, verweist auf eine Co-Evolution von Cas9, tracrRNA und crRNA-Repeat sowie auch auf das Vorhandensein einer weiterhin unbekannten Struktur und/oder Sequenz in der dualen RNA, die entscheidend für die Bildung des tertiären Komplexes mit den spezifischen Cas9-Orthologen ist.

[0536] Um die Anforderungen an die Protospacer-Sequenz für eine CRISPR/Cas-Immunität vom Typ II in bakteriellen Zellen zu untersuchen, haben wir eine Reihe von Protospacer-enthaltenden Plasmid-DNAs, die Einzelnukleotid-Mutationen umfassten, im Hinblick auf ihre Beibehaltung nach Transformationen in *S. pyogenes* und ihre Fähigkeit, von Cas9 in vitro gespalten zu werden, untersucht. Im Gegensatz zu Punktmutationen, die am 5'-Ende des Protospacers eingebracht wurden, wurden Mutationen in der Region nahe des PAM und der Cas9-Spaltstellen in vivo nicht toleriert und führten zu einer verringerten Plasmid-Spaltungseffizienz in vitro (**Fig. 12D**). Unsere Ergebnisse stimmten mit einem vorherigen Bericht über Protospacer-Escape-Mutanten, die in vivo im CRISPR-System vom Typ II aus *S. thermophilus* selektiert wurden, überein (27, 29). Darüber hinaus verweisen Beibehaltung und Spaltung der Plasmide auf das Vorliegen einer "Seed"-Region, die am 3'-Ende der Protospacer-Sequenz lokalisiert ist und entscheidend für die Interaktion mit crRNA und der nachfolgenden Spaltung durch Cas9 ist. Cas9 verstärkte die Hybridisierung des komplementären DNA-Strangs an die crRNA, was diese Vermutung unterstützte; diese Verstärkung war am stärksten in der 3'-terminalen Region der crRNA-Targeting-Sequenz (**Fig. 25A–C**). Als weiterer Beleg für diesen Befund ist ein zusammenhängender Abschnitt von mindestens 13 Basenpaaren zwischen der crRNA und der Ziel-DNA-Stelle neben dem PAM für eine effiziente Spaltung des Ziels erforderlich, während bis zu sechs zusammenhängende Fehlpaarungen in der 5'-terminalen Region des Protospacers toleriert werden (**Fig. 12E**). Diese Befunde erinnern an die zuvor beobachteten Seed-Sequenz-Anforderungen für eine Ziel-Nukleinsäure-Erkennung in Argonauten-Proteinen (30, 31) und an die Kaskaden- und Csy-CRISPR-Komplexe (13, 14).

[0537] **Fig. 12(A)** Cas9-tracrRNA:crRNA-Komplexe wurden rekonstituiert, wobei crRNA-sp2 aus 42 Nukleotiden und trunkierte tracrRNA-Konstrukte verwendet wurden und auf eine Spaltungs-Aktivität untersucht wurden, wie es in **Fig. 10B** beschrieben ist. **(B)** Cas9, das mit tracrRNA vollständiger Länge und crRNA-sp2-Trunkierungen programmiert worden war, wurde auf eine Aktivität untersucht, wie es in **(A)** beschrieben ist. **(C)** Minimal-Regionen von tracrRNA und crRNA, die in der Lage sind, eine Cas9-vermittelte DNA-Spaltung zu führen (schattierte Region). **(D)** Plasmide, die WT- oder Mutanten-Protospacer-2-Sequenzen mit den angegebenen Punktmutationen enthalten, wurden in vitro durch programmierte Cas9 gespalten, wie es in **Fig. 10A** dargestellt ist, und für Transformations-Assays von WT- oder prä-crRNA-defizienten *S. pyogenes*-Mutanten verwendet. Die Transformationseffizienz wurde als Kolonie-bildende Einheiten (colony forming units, CFU) pro

mg Plasmid-DNA berechnet. Die Fehlerbalken stellen Standardabweichungen (SD) für drei biologische Replikate dar. (E) Plasmide, die WT- und Mutanten-Protospacer-2-Inserts enthielten, welche ein unterschiedliches Ausmaß von crRNA-Ziel-DNA-Fehlpaarungen umfassten (unten), wurden in vitro durch ein programmiertes Cas9 gespalten (oben). Die Spaltungsreaktionen wurden mit XmnI weiter verdaut. Die Fragmente von 1880 und 800 bp sind durch Cas9 erzeugte Spaltprodukte. M, DNA-Marker.

[0538] Fig. 22 Elektrophoretische Mobilitäts-Shift-Assays wurden unter Verwendung von Protospacer 4-Ziel-DNA-Duplex und Cas9 (das die Mutationen D10A und H840 enthielt, welche die Nuklease-Domäne inaktivieren) allein oder in Gegenwart von crRNA-sp4, tracrRNA (75 nt) oder beidem durchgeführt. Der Ziel-DNA-Duplex wurde an beiden 5'-Enden radioaktiv markiert. Cas9 (D10/H840A) und die Komplexe wurden von 1 nM bis 1 µM titriert. Die Bindung wurde durch 8%ige native Polyacrylamid-Gel-Elektrophorese analysiert und durch Phosphor-Imaging sichtbar gemacht. Es ist zu beachten, dass Cas9 allein die Ziel-DNA mit moderater Affinität bindet. Diese Bindung wird durch Zugabe von crRNA nicht beeinflusst, was darauf verweist, dass dies eine Sequenz-unspezifische Interaktion mit der dsDNA darstellt. Darüber hinaus kann diese Interaktion in Abwesenheit der crRNA durch tracrRNA allein vermieden werden. In Gegenwart von sowohl crRNA als auch tracrRNA ist die Bindung von Ziel-DNA wesentlich verstärkt und ergibt eine Spezies mit unterschiedlicher elektrophoretischer Mobilität, was auf eine spezifische Erkennung von Ziel-DNA verweist.

[0539] Fig. 23 Ein Fragment von tracrRNA, das einen Teil der crRNA-gepaarten Region und einen Teil der stromabwärts gelegenen Region umfasst, ist ausreichend, um eine Spaltung von Protospacer-Oligonukleotid-DNA durch Cas9 zu steuern. (A) Protospacer-1-Oligonukleotid-DNA-Spaltung und (B) Protospacer 4-Oligonukleotid-DNA-Spaltung durch Cas9, das von einer reifen verwandten crRNA und verschiedenen tracrRNA-Fragmenten geführt wird. (A, B) Die Größe in Nukleotiden sind angegeben.

[0540] Fig. 24 Wie Cas9 von *S. pyogenes* spalten die eng verwandten Cas9-Orthologe der Grampositiven Bakterien *L. innocua* und *S. thermophilus* Protospacer-DNA, wenn sie von tracrRNA: crRNA von *S. pyogenes* geleitet werden. Unter denselben Bedingungen jedoch wird keine DNA-Spaltung durch die weniger eng verwandte Cas9-Orthologe aus den Gramnegativen Bakterien *C. jejuni* und *N. meningitidis* beobachtet. Spy, *S. pyogenes* SF370 (Zugriffsnummer NC_002737); Sth, *S. thermophilus* LMD-9 (STER_1477 Cas9-Ortholog; Zugriffsnummer NC_008532); Lin, *L. innocua* Clip11262 (Zugriffsnummer NC_003212); Cje, *C. jejuni* NCTC 11168 (Zugriffsnummer NC_002163); Nme, *N. meningitidis* A Z2491 (Zugriffsnummer NC_003116). (A) Spaltung von Protospacer-Plasmid-DNA. Protospacer 2-Plasmid-DNA (300 ng) wurde einer Spaltung durch unterschiedliche Cas9-Orthologe (500 nM) unterzogen, wobei diese durch hybride tracrRNA:crRNA-sp2-Duplexe (500 nM, 1:1) aus unterschiedlichen Spezies geführt wurde. Um die RNA-Duplexe zu erstellen, haben wir tracrRNA-Sequenzen von *L. innocua* und *N. meningitidis* basierend auf zuvor publizierten Northern-Blot-Daten (4) vorhergesagt. Die dualen Hybrid-RNA-Duplexe bestehen aus Spezies-spezifischer tracrRNA und einer heterologen crRNA. Die heterologe crRNA-Sequenz wurde so erstellt, dass sie eine auf DNA gerichtete sp2-Sequenz von *S. pyogenes* an ihrem 5'-Ende fusioniert an eine tracrRNA-bindende Repeat-Sequenz von *L. innocua* oder *N. meningitidis* am 3'-Ende enthielt. Cas9-Orthologe von *S. thermophilus* und *L. innocua*, jedoch nicht von *N. meningitidis* oder *C. jejuni*, können durch tracrRNA:crRNA-sp2 von *S. pyogenes* geleitet werden, um Protospacer 2-Plasmid-DNA zu spalten, wobei dies jedoch mit geringfügig verringerter Effizienz erfolgt. In ähnlicher Weise kann das aus *L. innocua* stammende Hybrid-tracrRNA:crRNA-sp2 ein Cas9 von *S. pyogenes* leiten, so dass die Ziel-DNA mit einer hohen Effizienz gespalten wird, während das aus *N. meningitidis* stammende Hybrid-tracrRNA:crRNA-sp2 nur eine geringe DNA-Spaltungs-Aktivität durch Cas9 von *S. pyogenes* auslöst. Als Kontrollen spalten Cas9-Orthologe von *N. meningitidis* und *L. innocua* Protospacer 2-Plasmid-DNA wenn sie durch das verwandte Hybrid-tracrRNA:crRNA-sp2 geleitet werden. Es ist zu beachten, dass wie oben erwähnt die tracrRNA-Sequenz von *N. meningitidis* lediglich vorausgesagt und noch nicht durch RNA-Sequenzierung bestätigt wurde. Somit könnte die geringe Spaltungseffizienz entweder das Ergebnis einer geringen Aktivität der Cas9-Orthologe sein oder das Ergebnis der Verwendung einer nicht optimal erstellten tracrRNA-Sequenz. (B) Spaltung von Protospacer-Oligonukleotid-DNA. Ein am 5'-Ende radioaktiv markiertes Oligonukleotid des komplementären Strangs (10 nM), das mit einem nicht markierten Oligonukleotid des nicht-komplementären Strangs vorab zusammengelagert worden war (Protospacer 1) (10 nM) (links) oder ein am 5'-Ende radioaktiv markiertes Oligonukleotid des nicht-komplementären Strangs (10 nM), das mit einem nicht markierten Oligonukleotid des komplementären Strangs vorab zusammengelagert worden war (10 nM) (rechts) (Protospacer 1), wurde einer Spaltung mit verschiedenen Orthologen von Cas9 (500 nM), die von einem tracrRNA:crRNA-sp1-Duplex von *S. pyogenes* (500 nM, 1:1) geführt wurden, unterzogen. Cas9-Orthologe von *S. thermophilus* und *L. innocua*, jedoch nicht von *N. meningitidis* oder *C. jejuni*, können von der verwandten Dual-RNA von *S. pyogenes* geführt werden, sodass die Protospacer-Oligonukleotid-DNA gespalten wird, auch wenn dies mit verringerter Effizienz erfolgt. Es ist zu beachten, dass die Spaltungsstelle auf dem komplementären DNA-Strang für alle drei Orthologe identisch ist. Die Spaltung des nicht-komplementären Strangs erfolgt an unterschiedlichen

Positionen. (C) Aminosäuresequenz-Identität von Cas9-Orthologen. Die Cas9-Orthologe von *S. pyogenes*, *S. thermophilus* und *L. innocua* teilen eine hohe prozentuale Aminosäuresequenz-Identität. Im Gegensatz dazu unterscheiden sich die Cas9-Proteine aus *C. jejuni* und *N. meningitidis* hinsichtlich der Sequenz und der Länge (etwa 300–400 Aminosäuren kürzer). (D) Co-Faltungen von künstlich erstellten, Spezies-spezifischen, heterologen crRNA-Sequenzen mit den entsprechenden tracrRNA-Orthologen von *S. pyogenes* (experimentell bestätigt, (4)), *L. innocua* (vorhergesagt) oder *N. meningitidis* (vorhergesagt). tracrRNAs; crRNA-Spacer 2-Fragmente; und crRNA-Repeat-Fragmente werden verfolgt und markiert. Hybrid-tracrRNA:crRNA-sp2-Duplexe von *L. innocua* und *S. pyogenes* weisen sehr wenige gemeinsame strukturelle Eigenschaften auf, wobei sich diese allerdings von Hybrid-tracrRNA:crRNA-sp2-Duplexen von *N. meningitidis* unterscheiden. Gemeinsam mit den oben in (A) und (B) beschriebenen Spaltungsdaten verweisen die Vorhersagen zur Co-Faltung darauf, dass die Spezies-spezifische Spaltung von Ziel-DNA durch Cas9-tracrRNA:crRNA durch ein noch unbekanntes strukturelles Merkmal in dem tracrRNA:crRNA-Duplex, der spezifisch von einem verwandten Cas9-Ortholog erkannt wird, beeinflusst wird. Es wurde vorhergesagt, dass die Spezies-Spezifität der Spaltung, die in (A) und (B) beobachtet wurde, auf der Ebene der Bindung von Cas9 an die Dual-tracrRNA:crRNA erfolgt. Die von der Dual-RNA geleitete Cas9-Spaltung von Ziel-DNA kann Spezies-spezifisch sein. Abhängig vom Grad der Diversität/Evolution zwischen Cas9-Proteinen und tracrRNA:crRNA-Duplexen sind Cas9 und Dual-RNA-Orthologe teilweise austauschbar.

[0541] Fig. 25 Eine Gruppe von DNA-Sonden aus 8 Nukleotide, die komplementär zu Regionen in der crRNA sind, welche die auf DNA gerichtete Region und die tracrRNA-bindende Region umfassen, wurden auf ihre Fähigkeit untersucht, mit der crRNA im Zusammenhang mit einem tracrRNA:crRNA-Duplex und dem aus drei Teilen bestehenden Komplex aus Cas9, tracrRNA und crRNA zu hybridisieren. (A) Schematische Darstellung der Sequenzen von DNA-Sonden, die in dem Assay verwendet werden, und deren Bindungsstellen in crRNA-sp4. (B–C) Elektrophoretische Mobilitäts-Shift-Assays von Ziel-DNA-Sonden mit tracrRNA:crRNA-sp4 oder Cas9-tracrRNA:crRNA-sp4. Das tracrRNA (15–89)-Konstrukt wurde in dem Experiment verwendet. Die Bindung der Duplexe oder Komplexe an Ziel-Oligonukleotid-DNAs wurde auf einem 16%igen nativen Polyacrylamid-Gel analysiert und durch Phosphor-Imaging sichtbar gemacht.

Ein kurzes Sequenz-Motiv bestimmt die Bildung einer R-Schleife

[0542] In zahlreichen CRISPR/Cas-Systemen ist gezeigt worden, dass an der Erkennung von "selbst" gegenüber "nicht selbst" ein kurzes Sequenz-Motiv beteiligt ist, das in dem fremden Genom konserviert ist und als PAM bezeichnet wird (27, 29, 32–34). PAM-Motive sind lediglich einige Basenpaare lang und ihre genaue Sequenz und Position variiert mit dem Typ des CRISPR/Cas-Systems (32). In dem System von *S. pyogenes* vom Typ II entspricht das PAM einer NGG-Konsensus-Sequenz, die zwei G:C Basenpaare enthält, welche ein Basenpaar stromabwärts von der crRNA-Bindungssequenz innerhalb der Ziel-DNA vorkommt (4). Transformations-Assays haben gezeigt, dass das GG-Motiv für die Eliminierung der Protospacer-Plasmid-DNA durch CRISPR/Cas in bakteriellen Zellen essentiell ist (**Fig. 24A**), was mit vorhergehenden Beobachtungen in *S. thermophilus* übereinstimmt (27). Das Motiv ist ferner essentiell für die in vitro-Spaltung eines Protospacer-Plasmids durch ein tracrRNA:crRNA-geführtes Cas9 (**Fig. 26B**). Um die Rolle des PAMs bei der Spaltung der Ziel-DNA durch den Cas9-tracrRNA:crRNA-Komplex zu untersuchen, haben wir eine Reihe von dsDNA-Duplexen untersucht, die Mutationen in der PAM-Sequenz auf dem komplementären oder auf dem nicht-komplementären Strang oder auf beiden umfassten (**Fig. 13A**). Die Spaltungs-Assays unter Verwendung dieser Substrate zeigten, dass die durch Cas9 katalysierte DNA-Spaltung besonders empfindlich in Bezug auf Mutationen in der PAM-Sequenz auf dem nicht-komplementären Strang der DNA ist, im Gegensatz zur Erkennung des PAMs im komplementären Strang durch CRISPR/Cas-Systeme vom Typ I (18, 34). Die Spaltung von einzelsträngigen Ziel-DNAs wurde durch die Mutation des PAM-Motivs nicht beeinflusst. Diese Beobachtung legt nahe, dass das PAM-Motiv lediglich im Zusammenhang mit Ziel-dsDNA erforderlich ist und daher bei der Entwindung des Duplex, der Strang-Invasion und der Bildung einer R-Schleifenstruktur erforderlich sein könnte. Als wir ein anderes crRNA-Ziel-DNA-Paar verwendeten (crRNA-sp4 und Protospacer 4-DNA), das aufgrund des Vorhandenseins eines kanonischen PAMs, das nicht in der Protospacer 2-Ziel-DNA vorhanden war, ausgewählt wurde, haben wir herausgefunden, dass beide G-Nukleotide des PAMs für die wirksame durch Cas9 katalysierte DNA-Spaltung erforderlich waren (**Fig. 13B** und **Fig. 26C**). Um zu bestimmen, ob das PAM eine direkte Rolle bei der Rekrutierung des Cas9-tracrRNA:crRNA-Komplexes an der korrekten Stelle der Ziel-DNA spielt, haben wir die Bindungsaffinitäten des Komplexes für Ziel-DNA-Sequenzen durch native Gel-Mobilitäts-Shift-Assays analysiert (**Fig. 13C**). Die Mutation eines der beiden G in der PAM-Sequenz verringerte die Affinität des Cas9-tracrRNA:crRNA für die Ziel-DNA wesentlich. Dieses Ergebnis veranschaulicht eine Rolle der PAM-Sequenz in der Ziel-DNA-Bindung durch Cas9.

[0543] Fig. 13(A) Durch duale RNA programmiertes Cas9 wurde auf eine Aktivität getestet, wie es in **Fig. 10B** beschrieben ist. WT- und mutierte PAM-Sequenzen in Ziel-DNAs sind mit Strichen dargestellt. **(B)** Protospacer-4-Ziel-DNA-Duplexe (an beiden 5'-Enden markiert), die WT- und mutierte PAM-Motive enthielten, wurden mit Cas9 inkubiert, welche mit tracrRNA:crRNA-sp4 (Nukleotide 23 bis 89) programmiert worden war. Zu den angegebenen Zeitpunkten (in Minuten) wurden Aliquots für die Spaltungsreaktion entnommen und wie in **Fig. 10B** dargestellt analysiert. **(C)** Elektrophoretische Mobilitäts-Shift-Assays wurden unter Verwendung von mittels RNA programmiertem Cas9 (D10A/H840A) und Protospacer 4-Ziel-DNA-Duplexen [dieselben wie in **(B)**], welche WT- und mutierte PAM-Motive enthielten, durchgeführt. Der Cas9(D10A/H840A)-RNA-Komplex wurde von 100 pM bis 1 mM titriert.

[0544] Fig. 26(A) Mutationen der PAM-Sequenz in Protospacer 2-Plasmid-DNA beseitigen die Beeinflussbarkeit der Plasmid-Beibehaltung durch das CRISPR/Cas-System von Typ II in bakteriellen Zellen. Wildtyp-Protospacer 2-Plasmide mit einem funktionsfähigen oder einem mutierten PAM wurden in den Wildtyp (Stamm SF370, auch als EC904 bezeichnet) und in eine prä-crRNA-defiziente Mutante (EC1479) von *S. pyogenes* transformiert, wie es in **Fig. 12D** beschrieben ist. PAM-Mutationen werden in vivo durch das CRISPR/Cas-System vom Typ II nicht toleriert. Die Mittelwerte und Standardabweichungen von drei biologischen Replikaten sind gezeigt. **(B)** Mutationen der PAM-Sequenz in Protospacer-Plasmid-DNA beseitigen die Spaltung durch Cas9-tracrRNA:RNA. Wildtyp-Protospacer 2-Plasmid mit einem funktionsfähigen oder einem mutierten PAM wurden einer Cas9-Spaltung unterzogen, wie es in **Fig. 10A** beschrieben ist. Die PAM-Mutanten-Plasmide werden durch den Cas9-tracrRNA:crRNA-Komplex nicht gespalten. **(C)** Mutationen der kanonischen PAM-Sequenz beseitigen die Beeinflussbarkeit der Plasmid-Beibehaltung durch das CRISPR/Cas-System von Typ II in bakteriellen Zellen. Wildtyp-Protospacer 4-Plasmide mit einem funktionsfähigen oder einem mutierten PAM wurden mit Cas9 gespalten, die mit tracrRNA und crRNA-sp2 programmiert worden war. Die Spaltungsreaktionen wurden in Gegenwart von XmnI-Restriktionsendonuklease durchgeführt, um die Cas9-Spaltprodukte als zwei Fragmente sichtbar zu machen (etwa 1880 und etwa 800 bp). Die Größen der Fragmente sind in Basenpaaren angegeben.

Cas9 kann mit einer einzelnen chimären RNA programmiert werden

[0545] Die Untersuchung einer wahrscheinlichen Sekundärstruktur des tracrRNA:crRNA-Duplex (**Fig. 10E** und **Fig. 12C**) legten die Möglichkeit nahe, dass die Merkmale, welche für eine ortsspezifische durch Cas9 katalysierte DNA-Spaltung erforderlich sind, in einer einzelnen chimären RNA zusammengeführt werden können. Obwohl die Selektionsmechanismen für das tracrRNA:crRNA-Ziel in der Natur wirksam funktionieren, ist die Möglichkeit eines durch eine Einzel-RNA geführten Cas9 aufgrund ihrer möglichen Verwendung für die programmierte DNA-Spaltung und das Genom-Editing attraktiv (**Fig. 1A–B**). Wir erstellten zwei Versionen einer chimären RNA, die eine Zielerkennungs-Sequenz am 5'-Ende aufwies gefolgt von einer Haarnadelschleifen-Struktur, welche die Basenpaarungs-Interaktionen beibehielt, die zwischen der tracrRNA und der crRNA vorkommen (**Fig. 14A**). Dieses einzelne Transkript fusioniert in wirksamer Weise das 3'-Ende der crRNA an das 5'-Ende der tracrRNA, wodurch die Struktur der Dual-RNA, welche erforderlich ist, um die ortsspezifische Spaltung von DNA durch Cas9 zu leiten, nachgestellt wird. In Spaltungs-Assays unter Verwendung von Plasmid-DNA haben wir beobachtet, dass die längere chimäre RNA in der Lage war, die durch Cas9 katalysierte DNA-Spaltung in einer Weise zu leiten, die ähnlich zu der ist, die für den trunkierten tracrRNA:crRNA-Duplex beobachtet wurde (**Fig. 14A** und **Fig. 27A** und **C**). Die kürzere chimäre RNA arbeitete in diesem Assay nicht effizient, was belegt, dass Nukleotide, die 5 bis 12 Positionen jenseits der tracrRNA:crRNA-Basenpaarungs-Interaktion liegen, für die effiziente Cas9-Bindung und/oder Zielerkennung wichtig sind. Wir erhielten ähnliche Ergebnisse in Spaltungs-Assays bei Verwendung von kurzer dsDNA als Substrat, was weiterhin darauf verweist, dass die Position der Spaltungsstelle in der Ziel-DNA identisch zu der ist, die bei Verwendung der Dual-tracrRNA:crRNA als Führung beobachtet wurde (**Fig. 14B** und **Fig. 27 B** und **C**). Um festzustellen, ob die Erstellung von chimärer RNA universell einsetzbar ist, haben wir 5 verschiedene chimäre Guide-RNAs erstellt, und auf einen Teil des Gens gerichtet, welches für das grün fluoreszierende Protein (GFP) kodiert (**Fig. 28, A** und **C**), und wir haben ihre Wirksamkeit gegen ein Plasmid, welches die für GFP kodierende Sequenz trägt, in vitro untersucht. In allen 5 Fällen spaltete Cas9, das mit diesen chimären RNAs programmiert worden war, das Plasmid effizient an der korrekten Zielstelle (**Fig. 14C** und **Fig. 28D**), was darauf verweist, dass die gezielte Erstellung von chimären RNAs robust ist und im Prinzip das Targeting einer beliebigen DNA-Sequenz von Interesse mit wenigen Einschränkungen über das Vorhandensein eines GG-Dinukleotids neben der Ziel-Sequenz hinaus ermöglicht.

[0546] Fig. 1 Eine auf DNA gerichtete RNA umfasst ein einzelsträngiges "auf DNA gerichtetes Segment" und ein "Protein-bindendes Segment", welches einen Abschnitt doppelsträngiger RNA umfasst. **(A)** Eine auf DNA gerichtete RNA kann zwei getrennte RNA-Moleküle umfassen (vorliegend als "Doppelmolekül" oder als auf

DNA gerichtete RNA "aus zwei Molekülen" bezeichnet). Ein auf DNA gerichtetes RNA-Doppelmolekül umfasst eine "Targeter-RNA" und eine "Aktivator-RNA". (B) Eine auf DNA gerichtete RNA kann ein einzelnes RNA-Molekül umfassen (als auf DNA gerichtetes RNA-"Einzelmolekül" bezeichnet). Ein auf DNA gerichtetes RNA-Einzelmolekül umfasst "Linker-Nukleotide".

[0547] Fig. 14(A) Ein Plasmid, das Protospacer 4-Ziel-Sequenz und ein WT-PAM umfasst, wurde einer Spaltung mit Cas9 unterzogen, das mit tracrRNA (4–89):crRNA-sp4-Duplex oder mit in vitro transkribierten chimären RNAs programmiert worden war, die konstruiert wurden, indem mit einer GAAA-Tetraschleife das 3'-Ende der crRNA an das 5'-Ende der tracrRNA gefügt wurde. Die Spaltungsreaktionen wurden durch Restriktionskartierung mit XmnI analysiert. Die Sequenzen der chimären RNAs A und B sind mit den auf DNA gerichteten Sequenzen (unterstrichen), den vom crRNA-Repeat abgeleiteten Sequenzen (mit Oberstrichen) und den von tracrRNA abgeleiteten Sequenzen (gestrichelt unterstrichen) gezeigt. (B) Protospacer 4-DNA-Duplex-Spaltungsreaktionen wurden wie in **Fig. 10B** gezeigt durchgeführt. (C) Fünf chimäre RNAs, die so erstellt wurden, dass sie auf das GFP-Gen gerichtet sind, wurden verwendet, um Cas9 so zu programmieren, dass ein Plasmid gespalten wird, das ein GFP-Gen umfasst. Die Plasmid-Spaltungsreaktion wurden wie in **Fig. 12E** beschrieben durchgeführt, mit der Ausnahme, dass die Plasmid-DNA nach der Cas9-Spaltung mit AvrII restriktionskartiert wurde.

[0548] Fig. 27(A) Eine einzelne chimäre RNA leitet die Cas9-katalysierte Spaltung von verwandter Protospacer-Plasmid-DNA (Protospacer 1 und Protospacer 2). Die Spaltungsreaktionen wurden in Gegenwart der XmnI-Restriktionsendonuklease durchgeführt, um die Cas9-Spaltprodukte als zwei Fragmente sichtbar zu machen (etwa 1880 und etwa 800 bp). Die Größen der Fragmente sind in Basenpaaren angegeben. (B) Eine einzelne chimäre RNA leitet die Cas9-katalysierte Spaltung von verwandter Protospacer-Oligonukleotid-DNA (Protospacer 1 und Protospacer 2). Die Größen der Fragmente sind in Nukleotiden angegeben. (C) Schematische Darstellungen der RNAs, die in diesem Experiment verwendet wurden. Die Sequenzen der chimären RNAs A und B sind mit der 5' auf DNA gerichteten Protospacer-Sequenz gezeigt (unterstrichen) der tracrRNA-bindenden Sequenz der crRNA (mit Oberstrichen) und der von tracrRNA abgeleiteten Sequenz (gestrichelt unterstrichen)

[0549] Fig. 28(A) Schematische Darstellung des GFP-Expressionsplasmids pCFJ127. Der Teil des offenen Leserahmens von GFP, auf den im Rahmen des Targetings abgestellt wurde, ist mit einem schwarzen Pfeil gekennzeichnet. (B) Vergrößerte Darstellung der Region, auf die im Rahmen des Targetings abgestellt wurde. Die Sequenzen, auf die im Rahmen des Targetings mittels chimärer RNAs abgestellt wurde, sind mit grauen Balken gezeigt. PAM-Dinukleotide sind in einem Rechteck dargestellt. Eine einzeln vorkommende Sall-Restriktionsschnittstelle ist 60 bp stromaufwärts vom Ziel-Locus lokalisiert. (C) Links: Ziel-DNA-Sequenzen sind zusammen mit ihren angrenzenden PAM-Motiven gezeigt. Rechts: Sequenzen der chimären Guide-RNAs. (D) pCFJ127 wurde durch Cas9 gespalten, die mit den chimären RNAs GFP 1–5 programmiert war, wie es angegeben ist. Das Plasmid wurde darüber hinaus mit Sall verdaut, und die Reaktionen wurden durch Elektrophorese auf einem 3%igen Agarose-Gel analysiert und durch Färbung mit SYBR Safe sichtbar gemacht.

Schlussfolgerungen

[0550] Es wurde ein DNA-Interferenz-Mechanismus identifiziert, an dem eine Dual-RNA-Struktur beteiligt ist, die eine Cas9-Endonuklease dazu bringt, ortsspezifische Doppelstrangbrüche in Ziel-DNA einzubringen. Das tracrRNA:crRNA-geführte Cas9-Protein verwendet unterschiedliche Endonuklease-Domänen (HNH und RuvC-ähnliche Domäne), um die beiden Stränge in der Ziel-DNA zu spalten. Die Zielerkennung durch Cas9 erfordert sowohl eine Seed-Sequenz in der crRNA als auch eine das GG-Dinukleotid enthaltene PAM-Sequenz, die an die crRNA-bindende Region in der Ziel-DNA grenzt. Wir zeigen ferner, dass die Cas9-Endonuklease mit Guide-RNA programmiert werden kann, die als einzelnes Transkript so erstellt wurde, dass sie auf eine beliebige dsDNA-Sequenz von Interesse gerichtet ist und diese spaltet. Das System ist effizient, vielseitig und programmierbar, indem man die Sequenz, welche das DNA-Ziel bindet, in der chimären Guide-RNA ändert.

[0551] Zinkfinger-Nukleasen und Transkriptions-Aktivator-ähnliche Effektor-Nukleasen haben als künstliche Enzyme, die zur Manipulation von Genomen erzeugt wurden, beträchtliches Interesse hervorgerufen (35–38). Dies stellt eine alternative Methodik dar, die auf einem RNA-programmierten Cas9 basiert und Gen-Targeting- und Genom-Editing-Anwendungen ermöglicht.

Zitierte Referenzen

1. B. Wiedenheft, S. H. Sternberg, J. A. Doudna, *Nature* 482, 331 (2012).
2. D. Bhaya, M. Davison, R. Barrangou, *Annu. Rev. Genet.* 45, 273 (2011).
3. M. P. Terns, R. M. Terns, *Curr. Opin. Microbiol.* 14, 321 (2011).
4. E. Deltcheva et al., *Nature* 471, 602 (2011).
5. J. Carte, R. Wang, H. Li, R. M. Terns, M. P. Terns, *Genes Dev.* 22, 3489 (2008).
6. R. E. Haurwitz, M. Jinek, B. Wiedenheft, K. Zhou, J. A. Doudna, *Science* 329, 1355 (2010).
7. R. Wang, G. Preamplume, M. P. Terns, R. M. Terns, H. Li, *Structure* 19, 257 (2011).
8. E. M. Gesner, M. J. Schellenberg, E. L. Garside, M. M. George, A. M. Macmillan, *Nat. Struct. Mol. Biol.* 18, 688 (2011).
9. A. Hatoum-Aslan, I. Maniv, L. A. Marraffini, *Proc. Natl. Acad. Sci. U.S.A.* 108, 21218 (2011).
10. S. J. J. Brouns et al., *Science* 321, 960 (2008).
11. D. G. Sashital, M. Jinek, J. A. Doudna, *Nat. Struct. Mol. Biol.* 18, 680 (2011).
12. N. G. Lintner et al., *J. Biol. Chem.* 286, 21643 (2011).
13. E. Semenova et al., *Proc. Natl. Acad. Sci. U.S.A.* 108, 10098 (2011).
14. B. Wiedenheft et al., *Proc. Natl. Acad. Sci. U.S.A.* 108, 10092 (2011).
15. B. Wiedenheft et al., *Nature* 477, 486 (2011).
16. C. R. Hale et al., *Cell* 139, 945 (2009).
17. J. A. L. Howard, S. Delmas, I. Ivančić-Baće, E. L. Bolt, *Biochem. J.* 439, 85 (2011).
18. E. R. Westra et al., *Mol. Cell* 46, 595 (2012).
19. C. R. Hale et al., *Mol. Cell* 45, 292 (2012).
20. J. Zhang et al., *Mol. Cell* 45, 303 (2012).
21. K. S. Makarova et al., *Nat. Rev. Microbiol.* 9, 467 (2011).
22. K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, E. V. Koonin, *Biol. Direct* 1, 7 (2006).
23. K. S. Makarova, L. Aravind, Y. I. Wolf, E. V. Koonin, *Biol. Direct* 6, 38 (2011).
24. S. Gottesman, *Nature* 471, 588 (2011).
25. R. Barrangou et al., *Science* 315, 1709 (2007).
26. J. E. Garneau et al., *Nature* 468, 67 (2010).
27. R. Saprunauskas et al., *Nucleic Acids Res.* 39, 9275 (2011).
28. G. K. Taylor, D. F. Heiter, S. Pietrovski, B. L. Stoddard, *Nucleic Acids Res.* 39, 9705 (2011).
29. H. Deveau et al., *J. Bacteriol.* 190, 1390 (2008).
30. B. P. Lewis, C. B. Burge, D. P. Bartel, *Cell* 120, 15 (2005).
31. G. Hutvagner, M. J. Simard, *Nat. Rev. Mol. Cell Biol.* 9, 22 (2008).
32. F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, *Microbiology* 155, 733 (2009).
33. L. A. Marraffini, E. J. Sontheimer, *Nature* 463, 568 (2010).
34. D. G. Sashital, B. Wiedenheft, J. A. Doudna, *Mol. Cell* 46, 606 (2012).
35. M. Christian et al., *Genetics* 186, 757 (2010).
36. J. C. Miller et al., *Nat. Biotechnol.* 29, 143 (2011).
37. F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang, P. D. Gregory, *Nat. Rev. Genet.* 11, 636 (2010).
38. D. Carroll, *Gene Ther.* 15, 1463 (2008).
39. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 2, 1989).
40. M. G. Caparon, J. R. Scott, Genetic manipulation of pathogenic streptococci. *Methods Enzymol.* 204, 556 (1991). doi: 10.1016/0076-6879(91)04028-M Medline
41. C. Frøkjær-Jensen et al., Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* 40, 1375 (2008). doi: 10.1038/ng.248 Medline
42. R. B. Denman, Using RNAFOLqD to predict the activity of small catalytic RNAs. *Biotechniques* 15, 1090 (1993). Medline
43. I. L. Hofacker, P. F. Stadler, Memory efficient folding algorithms for circular RNA secondary structures. *Bioinformatics* 22, 1172 (2006). doi: 10.1093/bioinformatics/bt1023 Medline
44. K. Darty, A. Denise, Y. Ponty, VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25, 1974 (2009). doi: 10.1093/bioinformatics/btp250 Medline

Beispiel 2: RNA-programmiertes Genom-Editing in humanen Zellen

[0552] Die unten dargestellten Daten zeigen, dass Cas9 exprimiert und in den Nukleus von humanen Zellen gebracht werden kann, und dass Cas9 sich in humanen Zellen mit single-Guide-RNA ("sgRNA"; umfasst die Merkmale, die sowohl für die Cas9-Bindung als auch für die Erkennung der Ziel-DNA-Stelle erforderlich sind) zusammenlagert. Diese Komplexe können Doppelstrangbrüche erzeugen und an einer Stelle in genom-

scher DNA, die komplementär zur Sequenz der sgRNA ist, eine Reparatur durch nicht-homologes End-Joining (NHEJ) stimulieren, wobei diese Aktivität sowohl Cas9 als auch die sgRNA erfordert. Die Extension der RNA-Sequenz an ihrem 3'-Ende verstärkt die DNA-Targeting-Aktivität in lebenden Zellen. Ferner zeigen Experimente unter Verwendung von Extrakten aus transfizierten Zellen, dass die Zusammenlagerung von sgRNA mit Cas9 der limitierende Faktor für die durch Cas9 vermittelte DNA-Spaltung ist. Diese Ergebnisse zeigen, dass RNA-programmiertes Genom-Editing in lebenden Zellen und auch in vivo funktioniert.

MATERIAL UND METHODEN

Plasmid-Design und Konstruktion

[0553] Die Sequenz, welche für das Cas9 aus *Streptococcus pyogenes* kodiert (Reste 1–1368), und die an ein HA-Epitop (Aminosäuresequenz DAYPYDVPDYASL (SEQ ID NO: 274)) sowie ein nukleares Ortsbestimmungssignal (Aminosäuresequenz PKKKRKVEDPKKKRKVD (SEQ ID NO: 275)) fusioniert war, wurde in Bezug auf eine Expression im Menschen Kodon-optimiert und von GeneArt synthetisiert. Die DNA-Sequenz ist SEQ ID NO: 276, und die Protein-Sequenz ist SEQ ID NO: 277. Ligation-unabhängige Klonierung (LIC) wurde verwendet, um die Sequenz in einen von pcDNA3.1-abgeleiteten GFP-Vektor und einen mCherry LIC-Vektor (Vektor 6D bzw. 6B, erhalten von UC Berkeley MacroLab) zu insertieren, was zu einer Cas9-HA-NLS-GFP und einer Cas9-HA-NLS-mCherry-Fusion führe, die unter der Kontrolle des CMV-Promoters exprimiert wurden. Guide-sgRNAs wurden unter Verwendung des Expressionsvektors pSilencer 2.1-U6 puro (Life Technologies) und pSuper (Oligoengine) exprimiert. RNA-Expressionskonstrukte wurden erzeugt, indem man komplementäre Oligonukleotide zusammenlagerte, um die RNA-kodierende DNA-Sequenz zu bilden, und das zusammengelagerte DNA-Fragment zwischen die Schnittstellen BamHI und HindIII in pSilencer 2.1-U6 puro und BgIII und HindIII in pSuper ligierte.

Zellkultur-Bedingungen und DNA-Transfektionen

[0554] HEK293T-Zellen wurden in Dulbecco modifiziertem Eagle Medium (DMEM), das mit 10% fötalem Kälberserum (FBS) supplementiert war, in einem befeuchteten Inkubator bei 37°C mit 5% CO₂ gehalten. Die Zellen wurden transient mit Plasmid-DNA transfiziert, wobei entweder X-tremeGENE DNA-Transfektions-Reagenz (Roche) oder Turbofect Transfektions-Reagenz (Thermo Scientific) gemäß den empfohlenen Protokollen verwendet wurde. Kurz gesagt wurden HEK293T-Zellen bei 60–80% Konfluenz in Platten mit 6 Vertiefungen unter Verwendung von 0,5 µg des Cas9-Expressions-Plasmids und 2 µg des RNA-Expressions-Plasmids transfiziert. Die Transfektionseffizienzen wurden basierend auf dem Anteil von GFP-positiven Zellen, die durch Fluoreszenz-Mikroskopie beobachtet wurden, auf 30–50% für Turbofect (**Fig. 29E** und **Fig. 37A–B**) und auf 80–90% für X-tremeGENE (**Fig. 31B**) geschätzt. 48 Stunden nach der Transfektion wurden die Zellen mit Phosphatgepufferter Salzlösung (PBS) gewaschen und lysiert, indem 250 µl Lysepuffer (20 mM Hepes pH 7.5, 100 mM Kaliumchlorid (KCl), 5 mM Magnesiumchlorid (MgCl₂), 1 mM Dithiothreitol (DTT), 5% Glycerol, 0,1% Triton X-100, supplementiert mit Roche-Protease-Inhibitor-Cocktail) verwendet wurden, und die Zellen wurden für 10 min bei 4°C geschüttelt. Das resultierende Zelllysate wurde für die weitere Analyse in Aliquots unterteilt. Die genomische DNA wurde aus 200 µl Zelllysate unter Verwendung des DNeasy Blood and Tissue Kits (Qiagen) gemäß dem Protokoll des Herstellers isoliert.

Western-Blot-Analyse der Cas9-Expression

[0555] HEK293T, die mit dem Cas9-HA-NLS-GFP-Expressionsplasmid transfiziert worden waren, wurden geerntet und 48 Stunden nach der Transfektion wie oben beschrieben lysiert. 5 µl des Lysats wurden auf einem 10%igen SDS-Polyacrylamid-Gel getrennt, mittels Blot auf eine PVDF-Membran überführt und mit einem HRP-konjugiertem Anti-HA-Antikörper untersucht (Sigma, 1:1000 Verdünnung in 1 × PBS).

Surveyor-Assay

[0556] Der Surveyor-Assay wurde wie zuvor beschrieben durchgeführt [10, 12, 13]. Kurz gesagt wurde der Locus der leichten A-Kette des humanen Clathrins (clathrin light chain A, CLTA) mittels PCR ausgehend von 200 ng genomischer DNA unter Verwendung einer High Fidelity-Polymerase, Herkulase II-Fusions-DNA-Polymerase (Agilent Technologies) und dem Forward-Primer 5'-GCAGCAGAAGAAGCCTTTGT-3' (SEQ ID NO: //) und dem Reverse-Primer 5'-TTCCTCCTCTCCCTCCTCTC-3' (SEQ ID NO: //) amplifiziert. 300 ng des 360 bp-Amplicons wurden anschließend durch Erhitzen auf 95°C denaturiert und langsam wieder zusammengelagert, wobei ein Hitze-Block verwendet wurde, um Wildtyp-DNA-Stränge und Mutanten-DNA-Stränge auf Zufallsbasis wieder zu hybridisieren. Die Proben wurden anschließend mit Cel-1-Nuklease (Surveyor Kit, Transgenomic)

für 1 Stunde bei 42°C inkubiert. Cel-1 erkennt und spaltet DNA-Helices, die Fehlpaarungen enthalten (Wildtyp: Mutanten-Hybridisierung). Die Produkte aus dem Verdau mit Cel-1-Nuklease wurden auf einem 10%igem Acrylamid-Gel getrennt und durch Färbung mit SYBR Safe (Life Technologies) sichtbar gemacht. Die Quantifizierung der Spaltungs-Banden wurde unter Verwendung der ImageLab-Software (Bio-Rad) durchgeführt. Die prozentuale Spaltung wurde bestimmt, indem man die durchschnittliche Intensität der Spaltprodukte (160–200 bps) durch die Summe der Intensitäten des nicht gespaltenen PCR-Produkts (360 bp) und des Spaltprodukts teilt.

In vitro Transkription

[0557] Guide-RNA wurde in vitro unter Verwendung einer rekombinanten T7-RNA-Polymerase und einer DNA-Matrize, die durch Zusammenlagerung komplementärer synthetischer Oligonukleotids hergestellt wurde, wie es zuvor beschrieben wurde [14]. Die RNAs wurden durch Elektrophorese auf einem denaturierenden Acrylamid-Gel mit 7M Harnstoff elektrophoretisch getrennt, mit Ethanol präzipitiert und in DEPC-behandeltem Wasser gelöst.

Northern-Blot-Analyse

[0558] RNA wurde aus HEK293T-Zellen unter Verwendung des mirVana-Small-RNA-Isolations-Kits (Ambion) gereinigt. Für jede Probe wurden 800 ng RNA nach Denaturierung für 10 min bei 70°C in RNA-Ladepuffer (0,5 × TBE (pH 7.5), 0,5 mg/ml Bromphenolblau, 0,5 mg Xylencyanol und 47% Formamid) auf einem PAGE-Gel mit 10% Harnstoff getrennt. Die Elektrophorese wurde bei 10 W in 0,5 × TBE-Puffer durchgeführt, bis der Bromphenolblau-Farbstoff den Boden des Gels erreichte; anschließend wurden die Proben per Elektroblot bei 20 Volt für 1,5 Stunden in 0,5 × TBE auf eine Nytran-Membran überführt. Die überführten RNAs wurden auf der Nytran-Membran in einem UV-Crosslinker (Stratagene) quervernetzt und bei 45°C für 3 h in einem Puffer prähybridisiert, der 40 Formamid, 5 × SSC, 3 × Denhardt (jeweils 0,1% Ficoll, Polyvinylpyrrolidon und BSA) und 200 µg/ml Herings-Sperma-DNA enthält. Die prähybridisierten Membranen wurden über Nacht in dem Prähybridisierungs-Puffer inkubiert, der mit einer 5'-³²P-markierten Antisense-DNA-Oligo-Sonde bei 1 Million cpm/ml markiert worden war. Nach mehreren Waschschritten in SSC-Puffer (finaler Waschschrift in 0,2 × SSC) wurden die Membranen durch Phosphor-Imaging sichtbar gemacht.

In vitro-Spaltungsassay

[0559] Zelllysate wurden wie oben beschrieben hergestellt und mit CLTA-RFP-Donor-Plasmid [10] inkubiert. Die Spaltungsreaktionen wurden in einem Gesamtvolumen von 20 µl durchgeführt und enthielten 10 µl Lysat, 2 µl 5 × Spaltungspuffer (100 mM HEPES pH 7,5, 500 mM KCl, 25 mM MgCl₂, 5 mM DTT, 25% Glycerol) und 300 ng Plasmid. Wo es angegeben ist, wurden die Reaktionen mit 10 pmol in vitro transkribierter CLTA1-sgRNA supplementiert. Die Reaktionen wurden bei 37°C für eine Stunde inkubiert und anschließend mit 10 U XhoI (NEB) für zusätzliche 30 min bei 37°C verdaut. Die Reaktionen wurden durch Zugabe von Proteinase K (Thermo Scientific) abgestoppt und für 15 min bei 37°C inkubiert. Die Spaltprodukte wurden durch Elektrophorese auf einem 1%igen Agarosegel analysiert und mit SYBR Safe gefärbt. Die Gegenwart von Fragmenten von etwa 2230 bp und etwa 3100 bp verweist auf das Vorliegen einer Cas9-vermittelten Spaltung.

ERGEBNISSE

[0560] Um zu prüfen, ob Cas9 programmiert werden könnte, um genomische DNA in lebenden Zellen zu spalten, wurde Cas9 gemeinsam mit einer sgRNA co-exprimiert, die so erstellt wurde, dass sie auf das Gen für die leichte Kette des humanen Clathrins (CLTA) gerichtet war. Der genomische CLTA-Locus ist bereits zuvor für das Targeting verwendet und unter Verwendung von ZFNs editiert worden [10]. Wir haben zunächst die Expression einer im Hinblick auf humane Kodons optimierten Version des Cas9-Proteins von *Streptococcus pyogenes* und sgRNA in humanen HEK293T-Zellen untersucht. Das 160 kDa-Cas9-Protein wurde als Fusionsprotein exprimiert, welches ein HA-Epitop, ein nukleares Ortsbestimmungssignal (NLS) und ein grün fluoreszierendes Protein (GFP) am C-Terminus von Cas9 aufweist (**Fig. 29A**). Die Analyse von Zellen, die mit einem Vektor transfiziert waren, der für das GFP-fusionierte Cas9 kodierte, ergab eine deutliche Cas9-Expression und eine Lokalisierung im Kern (**Fig. 29B**). Ein Western-Blot bestätigte, dass das Cas9-Protein in Extrakten aus diesen Zellen im Wesentlichen intakt exprimiert wird (**Fig. 29A**). Um Cas9 zu programmieren, exprimierten wir sgRNA, die eine 5'-terminale Sequenz aus 20 Nukleotiden umfasste, welche komplementär zu der Sequenz der Ziel-DNA ist, sowie eine 3'-terminale Haarnadelschleifen-Struktur aus 42 Nukleotiden, die für die Cas9-Bindung erforderlich ist (**Fig. 29C**). Diese 3'-terminale Sequenz entspricht der minimalen Haarnadelschleifen-Struktur, die zuvor verwendet worden war, um Cas9 in vitro zu programmieren [8]. Die Expres-

sion dieser sgRNA wurde durch den humanen U6(RNA Polymerase III)-Promoter gesteuert [11]. Northern-Blot-Analysen von RNA, die aus Zellen extrahiert wurde, welche mit einem sgRNA-Plasmid-Expressionsvektor transfiziert waren, der durch den U6-Promoter gesteuert wurde, zeigten, dass die sgRNA tatsächlich exprimiert wird, und dass ihre Stabilität in Gegenwart von Cas9 verstärkt wird (**Fig. 29D**).

[0561] **Fig. 29** zeigt, dass die Co-Expression von Cas9 und Guide-RNA in humanen Zellen Brüche in Doppelstrang-DNA am Ziel-Locus erzeugt. (A) Oben; schematisches Diagramm des Cas9-HA-NLS-GFP-Expressionskonstrukts. Unten; ein Lysat aus HEK293T-Zellen, die mit dem Cas9-Expressionsplasmid transfiziert worden waren, wurde per Western-Blot unter Verwendung des Anti-HA-Antikörpers analysiert. (B) Fluoreszenzmikroskopie von HEK293T-Zellen, die Cas9-HA-NLS-GFP exprimieren. (C) Aufbau einer Single-Guide-RNA (sgRNA, d. h. ein auf DNA gerichtetes RNA-Einzelmolekül), die auf den humanen CLTA-Locus gerichtet ist. Oben; schematisches Diagramm der sgRNA-Zielstelle in Exon 7 des humanen CLTA-Gens. Die Ziel-Sequenz, die mit dem Guide-Segment der CLTA1-sgRNA hybridisiert, ist als "CLTA1-sgRNA" angegeben. Das GG-Dinukleotid-Protospacer-benachbarte Motiv (PAM) ist durch einen Pfeil gekennzeichnet. Schwarze Linien verweisen auf DNA-bindende Regionen des ZFN-Kontroll-Proteins. Das Translations-Stopp-Kodon des offenen Leserahmens von CLTA ist zum Vergleich mit einer gestrichelten Linie markiert. Mitte; schematisches Diagramm des sgRNA-Expressionsprodukts. Die RNA wird unter der Kontrolle des U6-Pol III-Promoters und eines Poly (T)-Teils, der als Pol III-Transkriptions-Terminator-Signal dient, exprimiert. Unten; sgRNA-geführte Spaltung von Ziel-DNA durch Cas9. Die sgRNA besteht aus einem 5'-terminalen Guide-Segment aus 20 nt, dem eine Haarnadelschleifen-Struktur aus 42 nt folgt, welche für die Cas9-Bindung erforderlich ist. Die Cas9-vermittelte Spaltung der zwei Ziel-DNA-Stränge tritt bei Entwindung der Ziel-DNA und Bildung eines Duplex zwischen dem Guide-Segment der sgRNA und der Ziel-DNA auf. Dies ist vom Vorhandensein eines PAM-Motivs (für das verwendete Cas9 geeignet, z. B. GG: Dinukleotid, siehe obiges Beispiel 1) stromabwärts von der Ziel-Sequenz in der Ziel-DNA abhängig. Es ist zu beachten, dass die Ziel-Sequenz relativ zu dem oberen Diagramm invertiert dargestellt ist. (D) Northern-Blot-Analyse der sgRNA-Expression in HEK293T-Zellen. (E) Surveyor-Nuklease-Assay von genomischer DNA, die aus HEK293T-Zellen, welche Cas9 und/oder CLTA-sgRNA exprimieren, isoliert wurde. Ein ZFN-Konstrukt, das zuvor verwendet worden war, um ein Targeting des CLTA-Locus durchzuführen [10] wurde als Positiv-Kontrolle für den Nachweis der DSB-induzierten DNA-Reparatur durch nicht-homologes End Joining verwendet.

[0562] Als nächstes untersuchten wir, ob die ortsspezifischen DSBs in HEK293T-Zellen erzeugt werden, die mit Cas9-HA-NLS-mCherry und der CLTA1-sgRNA transfiziert worden waren. Um dies durchzuführen, haben wir unter Verwendung des Surveyor-Nuklease-Assays [12] in dem Locus nach geringfügigen Insertionen und Deletionen gesucht, die von einer nicht perfekten Reparatur durch DSB-induziertes NHEJ resultierten. Die Region der genomischen DNA, auf die beim Targeting mit Cas9:sgRNA abgestellt wird, wird durch PCR amplifiziert, und die resultierenden Produkte werden denaturiert und wieder zusammengelagert. Die rehybridisierten PCR-Produkte werden mit der Fehlpaarungs-Erkennungs-Endonuklease Cel-1 inkubiert und auf einem Acrylamid-Gel getrennt, um die Cel-1-Spaltungsbanden zu identifizieren. Da die DNA-Reparatur durch NHEJ üblicherweise durch einen DSB induziert wird, verweist ein positives Signal im Surveyor-Assay darauf, dass die Spaltung der genomischen DNA stattgefunden hat. Unter Verwendung dieses Assays haben wir die Spaltung des CLTA-Locus an einer Position nachgewiesen, auf die beim Targeting mit der CLTA1-sgRNA abgestellt wurde (**Fig. 29E**). Ein Paar aus ZFNs, die auf eine benachbarte Stelle im CLTA-Locus gerichtet waren, stellte in diesen Experimenten eine Positiv-Kontrolle bereit [10].

[0563] Um zu bestimmen, ob entweder die Expression von Cas9 oder die Expression von sgRNA ein limitierender Faktor für die beobachteten Genoms-Editing-Reaktionen ist, wurden ausgehend von transfizierten Zellen hergestellte Lysate mit Plasmid-DNA inkubiert, die ein Fragment des CLTA-Gens umfasste, auf das die CLTA1-sgRNA abstellte. Eine Spaltung der Plasmid-DNA wurde bei Inkubation mit Lysaten, die ausgehend von Zellen hergestellt worden waren, welche man mit dem Cas9-HA-NLS-GFP-Expressionsvektor allein transfiziert hatte, nicht beobachtet, was mit den Ergebnissen des Surveyor-Assays übereinstimmt. Eine robuste Plasmid-Spaltung wurde jedoch nachgewiesen, wenn die Lysate mit in vitro transkribierter CLTA1-sgRNA supplementiert waren (**Fig. 30A**). Darüber hinaus unterstützten Lysate, die ausgehend von Zellen hergestellt wurden, welche sowohl mit Cas9 als auch mit sgRNA-Expressionsvektoren transfiziert worden waren, die Plasmid-Spaltung, während Lysate von Zellen, die allein mit dem sgRNA-kodierenden Vektor transfiziert worden waren, dies nicht taten (**Fig. 30A**). Diese Ergebnisse zeigen, dass ein limitierender Faktor für die Cas9-Funktion in humanen Zellen der Zusammenbau mit der sgRNA sein könnte. Wir haben diese Möglichkeit direkt untersucht, indem die Plasmid-Spaltung in Lysaten von Zellen analysiert wurde, welche wie zuvor in Gegenwart oder in Abwesenheit von zugegebener exogener sgRNA transfiziert worden waren. Es ist zu beachten, dass bei Zugabe von exogener sgRNA zu Lysaten von Zellen, die sowohl mit dem Cas9- als auch mit dem sgRNA-Expressionsvektor transfiziert worden waren, eine deutliche Erhöhung der DNA-Spaltungsaktivität beobachtet

wurde (**Fig. 30B**). Dieses Ergebnis verweist darauf, dass der limitierende Faktor für die Cas9-Funktion in HEK 293T-Zellen die Expression der sgRNA oder die Beladung des Cas9 mit selbiger ist.

[0564] Fig. 30 zeigt, dass Zelllysate aktive Cas9:sgRNA umfassen und die ortsspezifische DNA-Spaltung unterstützen. (A) Lysate von Zellen, die mit dem/den links gezeigten Plasmid/Plasmiden transfiziert worden waren, wurden mit Plasmid-DNA inkubiert, die ein PAM und die zu der CLTA1-sgRNA komplementäre Ziel-Sequenz enthielten; wo es angegeben ist, wurde die Reaktion mit 10 pmol in vitro-transkribierter CLTA1-sgRNA supplementiert; die sekundäre Spaltung mit XhoI erzeugte Fragmente von etwa 2230 bp und etwa 3100 bp, die auf eine Cas9-vermittelte Spaltung verweisen. Eine Kontrollreaktion unter Verwendung von Lysat aus Zellen, die mit einem ZFN-Expressionskonstrukt transfiziert worden waren, zeigte Fragmente von leicht unterschiedlicher Größe, was den Abstand der ZFN-Zielstelle relativ zur CLTA1-Zielstelle widerspiegelt. (B) Lysate von Zellen, die mit Cas9-GFP-Expressions-Plasmid transfiziert worden waren, und, wo es angegeben ist, mit dem CLTA1-sgRNA-Expressions-Plasmid, wurden mit Ziel-Plasmid-DNA wie in (A) in Abwesenheit oder in Gegenwart von in vitro-transkribierten CLTA1-sgRNA inkubiert.

[0565] Als Mittel zur Verstärkung des Zusammenbaus von Cas9:sgRNA in lebenden Zellen haben wir als nächstes die Wirkung einer Ausdehnung der mutmaßlichen Cas9-bindenden Region der Guide-RNA untersucht. Zwei neue Versionen von CLTA1-sgRNA wurden so erstellt, dass sie zusätzliche 6 oder 12 Basenpaare in der Helix aufwiesen, welche die Basenpaarungs-Interaktionen zwischen der crRNA und der tracrRNA nachahmt (**Fig. 31A**). Darüber hinaus wurde das 3'-Ende der Guide-RNA basierend auf der nativen Sequenz der tracrRNA von *S. pyogenes* um fünf Nukleotide ausgedehnt [9]. Vektoren, die für diese am 3'-Ende ausgedehnten sgRNAs unter der Kontrolle entweder eines U6- oder eines H1-Pol III-Promoters kodierten, wurden gemeinsam mit dem Cas9-HA-NLS-GFP-Expressionsvektor in Zellen transfiziert, und die ortsspezifische Spaltung des Genoms wurde unter Verwendung des Surveyor-Assays untersucht (**Fig. 31B**). Die Ergebnisse bestätigten, dass die Spaltung sowohl Cas9 als auch CLTA1-sgRNA erforderte, wobei die Spaltung jedoch nicht auftrat, wenn entweder Cas9 oder die sgRNA allein exprimiert wurde. Darüber hinaus beobachteten wir wesentlich erhöhte Häufigkeiten von NHEJ, wie durch sie die Cel-1-Nuklease-Spaltung nachgewiesen wurde, während die Häufigkeit von NHEJ-Mutagenese, die mit dem Kontroll-ZFN-Paar erhalten wurde, im wesentlichen unverändert blieb.

[0566] Fig. 31 zeigt, dass die 3'-Ausdehnung von sgRNA-Konstrukten die ortsspezifische NHEJ-vermittelte Mutagenese verstärkt. (A) Das Konstrukt für die CLTA1-sgRNA-Expression (oben) wurde so erstellt, dass es Transkripte erzeugt, welche die ursprüngliche Cas9-bindende Sequenz enthielten (v1.0) oder dsRNA-Duplexe, die um 4 Basenpaare erweitert waren (v2.1) oder um 10 Basenpaare (v2.2). (B) Surveyor-Nuklease-Assay von genomischer DNA, die aus HEK293T-Zellen isoliert wurde, welche Cas9 und/oder CLTA1-sgRNA v1.0, v2.1 oder v2.2 exprimierten. Ein ZFN-Konstrukt, das zuvor für ein Targeting des CLTA-Locus verwendet worden war [10], wurde als Positiv-Kontrolle zum Nachweis einer DSB-induzierten DNA-Reparatur durch nicht-homologes End Joining verwendet.

[0567] Die Ergebnisse stellen daher den Rahmen für die Implementierung des Cas9 als mögliches molekulares Werkzeug für diverse Anwendungen beim Genom-Editing dar. Ein bedeutsames Merkmal dieses Systems ist das Potential zur Programmierung von Cas9 mit mehreren sgRNAs in derselben Zelle, entweder um die Effizienz des Targetings für einen einzelnen Locus zu erhöhen oder als Mittel für das Targeting von mehreren Loci gleichzeitig. Solche Strategien würden breite Anwendung bei Genom-weiten Experimenten und bei großen Forschungsvorhaben finden, z. B. bei der Entwicklung von multigenischen Erkrankungsmodellen.

Beispiel 3: Die tracrRNA- und Cas9-Familien von CRISPR-Cas-Immunitätssystemen vom Typ II

[0568] Wir haben nach allen putativen CRISPR-Cas-Loci vom Typ II gesucht, die gegenwärtig in öffentlich zugänglichen bakteriellen Genomen existieren, indem wir nach Sequenzen gesucht haben, die homolog zu Cas9 sind, dem wichtigsten Protein des Systems vom Typ II. Wir haben aus einem multiplen Sequenz-Alignment der identifizierten Cas9-Orthologe einen phylogenetischen Stammbaum erstellt. Die CRISPR-Repeat-Länge und die Genorganisation von Cas-Operons der assoziierten Typ II-Systeme wurden in den unterschiedlichen Cas9-Subclustern analysiert. Eine Subklassifizierung von Typ II-Loci wurde vorgeschlagen, und diese wurde basierend auf der Auswahl von 75 repräsentativen Cas9-Orthologen in weitere Subgruppen unterteilt. Wir haben anschließend die tracrRNA-Sequenzen vorhergesagt, hauptsächlich indem wir CRISPR-Repeat-Sequenzen gefunden haben, und indem wir innerhalb von oder in der Nähe von Cas-Genen und CRISPR-Arrays von ausgewählten Typ II-Loci nach Anti-Repeats gesucht haben. Es wurde eine vergleichende Analyse von Sequenzen und vorhergesagten Strukturen von ausgewählten tracrRNA-Orthologen durchgeführt.

Schließlich haben wir die Expressions- und Prozessierungsprofile von tracrRNAs und crRNAs aus fünf bakteriellen Spezies bestimmt.

MATERIAL UND METHODEN

Bakterielle Stämme und Kulturbedingungen

[0569] Die folgenden Medien wurden zur Anzucht von Bakterien auf Platten verwendet: TSA (Trypticase-Soja-Agar, Trypticase™ Soja Agar (TSA II) BD BBL, Becton Dickinson), der mit 3% Schafblut für *S. mutans* (UA159) supplementiert war, und BHI (brain heart infusion, BD Bacto™ Brain Heart Infusion, Becton Dickinson)-Agar für *L. innocua* (Clip 11262). Für die Kultivierung in Flüssigmedien wurde THY-Medium (Todd Hewitt Nährlösung (THB, Bacto, Becton Dickinson) verwendet, die mit 0,2% Hefeextrakt (Servabacter®) supplementiert war, für *S. mutans* verwendet, BHI-Nährlösung für *L. innocua*, BHI-Flüssigmedium, das 1% Vitamin-Mix VX (Difco, Becton Dickinson) enthielt, für *N. meningitidis* (A Z2491), MH (Mueller Hinton Nährlösung, Oxoid)-Nährlösung, die 1% Vitamin-Mix enthielt, für *C. jejuni* (NCTC 11168; ATCC 700819) und TSB (Tryptic Soy Broth, BD BBL™ Trypticase™ Soja Nährlösung) für *F. novicida* (U112). *S. mutans* wurde bei 37°C, 5% CO₂ ohne Schütteln inkubiert. Stämme von *L. innocua*, *N. meningitidis* und *F. novicida* wurden aerob bei 37°C unter Schütteln angezüchtet. *C. jejuni* wurde bei 37°C unter mikroaerophilen Bedingungen unter Verwendung einer Campygen (Oxoid)-Atmosphäre angezüchtet. Das bakterielle Zellwachstum wurde durch Messung der optischen Dichte der Kulturen bei 620 nm (OD_{620 nm}) in regelmäßigen Zeitabständen unter Verwendung eines Mikroplatten-Lesegeräts (BioTek PowerWave™) gemessen.

Sequenzierung von bakteriellen kleinen RNA-Bibliotheken

[0570] *C. jejuni* NCTC 11168 (ATCC 700819), *F. novicida* U112, *L. innocua* Clip11262, *N. meningitidis* A Z2491 und *S. mutans* UA 159 wurden bis zur mittleren logarithmischen Wachstumsphase kultiviert, und die Gesamt-RNA wurde mit TRIzol1 (Sigma-Aldrich) extrahiert. 10 µg der Gesamt-RNA aus jedem Stamm wurden mit TURBO™ DNase (Ambion) behandelt, um sämtliche restliche genomische DNA zu entfernen. Ribosomale RNAs wurden durch Verwendung des Ribo-Zero™ rRNA Removal Kits® für Grampositive oder Gram-negative Bakterien (Epicentre) gemäß den Anweisungen des Herstellers entfernt. Nach Aufreinigung mit dem RNA Clean & Concentrator™-5-Kit (Zymo Research) wurden die Bibliotheken unter Verwendung des ScriptMiner™ Small RNA-Seq Library Herstellungs-Kits (Multiplex, Illumina®-kompatibel) gemäß den Anweisungen des Herstellers hergestellt. Die RNAs wurden mit der sauren Pyrophosphatase aus Tabak (tobacco acid pyrophosphatase, TAP) (Epicentre) behandelt. Es wurden Säulen des RNA Clean & Concentrator™-5 (Zymo Research) für die anschließende RNA-Aufreinigung verwendet, und die Phusion® High-Fidelity-DNA-Polymerase (New England Biolabs) wurde für die PCR-Amplifikation verwendet. Spezifische benutzerdefinierte Barcodes wurden jeder Bibliothek zugefügt (RNA-Seq Barcode Primer (Illumina®-kompatibel) (Epicentre), und die Proben wurden durch die Next Generation Sequencing (CSF NGS Unit; im Internet bei "csf." gefolgt von "ac.at")-Einrichtung des Biozentrums Wien, Wien, Österreich, sequenziert (Illumina Single-End-Sequenzierung).

Analyse von tracrRNA- und crRNA-Sequenzierungsdaten

[0571] Die RNA-Sequenzierungsläufe wurden unter Verwendung des Illumina2bam-Werkzeugs zurecht geschnitten und bearbeitet, indem man (i) die Illumina-Adapter-Sequenzen entfernt hat (cutadapt 1.0) und (ii) 15 nt am 3'-Ende entfernt hat, um die Qualität der Läufe zu verbessern. Nach Entfernung von Läufen, die kürzer als 15 nt sind, wurden die cDNA-Auslesungen ihrem entsprechenden Genom unter Verwendung von Bowtie gegenübergestellt, wobei 2 Fehlpaarungen erlaubt wurden: *C. jejuni* (GenBank: NC_002163), *F. novicida* (GenBank: NC_008601), *N. meningitidis* (GenBank: NC_003116), *L. innocua* (GenBank: NC_003212) und *S. mutans* (GenBank: NC_004350). Die Abdeckung der Läufe wurde unter Verwendung der BEDTools-Version-2.15.0 an jeder Nukleotidposition getrennt für beide DNA-Stränge berechnet. Ein normalisierter Wiggle-Datei, die eine Abdeckung in Auslesung pro Million (reads per million, rpm) enthielt, wurde erstellt und unter Verwendung des Integrated Genomics Viewers (IGV)-Werkzeugs sichtbar gemacht ("www." gefolgt von "broadinstitute.org/igv/") (**Fig. 36**). Unter Verwendung von SAMtools flagstat⁸⁰ wurde der Teil der kartierten Auslesungen auf insgesamt 9914184 Auslesungen für *C. jejuni* berechnet, 48205 Auslesungen für *F. novicida*, 13110087 Auslesungen für *N. meningitidis*, 161865 Auslesungen für *L. innocua* und 1542239 Auslesungen für *S. mutans*. Es wurde eine Datei erstellt, welche die Anzahl der Auslesungen beginnend von (5') und endend bei (3') an jeder einzelnen Nukleotidposition enthielt, und diese wurde in IGV sichtbar gemacht. Für jedes tracrRNA-Ortholog und jede crRNA wurde die Gesamtanzahl der erhaltenen Auslesungen unter Verwendung des SAMtools berechnet.

Cas9-Sequenz-Analyse, multiples Sequenz-Alignment und Erstellung eines Guide-Stammbaums

[0572] Es wurde ein Positions-spezifisches wiederholtes (position specific iterated, PSI)-BLAST-Programm verwendet, um Homologe der Cas9-Familie in der nicht-redundanten Datenbank des NCBI zu erhalten. Sequenzen, die kürzer waren als 800 Aminosäuren wurden verworfen. Die BLASTClust-Programm-Einstellung mit einem Ausschlusswert von 0,8 für die Längenabdeckung und einer Score-Abdeckungsschwelle (Bit-Score geteilt durch die Länge des Alignments) von 0,8 wurde verwendet, um die verbleibenden Sequenzen zu gruppieren (**Fig. 38**). Dieses Vorgehen erzeugte 78 Cluster (von denen 48 jeweils durch nur eine Sequenz repräsentiert wurden). Ein Vertreter (oder selten einige wenige Vertreter) von jedem Cluster wurde ausgewählt, und es wurde ein multiples Alignment für diese Sequenzen unter Verwendung des MUSCLE-Programms mit voreingestellten Parametern konstruiert, gefolgt von einer manuellen Korrektur auf Basis von lokalen Alignments, die unter Verwendung der Programme PSI-BLAST und HHpred erhalten wurden. Einige zusätzliche Sequenzen waren einem Alignment nicht zugänglich und wurden ebenfalls von den endgültigen Alignments ausgeschlossen. Die sicher miteinander in einem Alignment verglichenen Blöcke mit 272 informativen Positionen wurden für die Rekonstruktion eines Baumes der maximalen Wahrscheinlichkeit verwendet, wobei das FastTree-Programm mit voreingestellten Parametern verwendet wurde: JTT, evolutionäres Modell, diskretes Gamma-Modell mit 20 Bemessungskategorien. Das gleiche Programm wurde verwendet, um die Bootstrap-Werte zu berechnen.

[0573] **Fig. 38** zeigt Sequenzen, die gemäß dem BLASTClust-Clustering-Programm gruppiert wurden. Lediglich Sequenzen, die länger waren als 800 Aminosäuren wurden für die BLASTClust-Analyse ausgewählt (siehe Material und Methoden). Repräsentative Stämme, die Gene für Cas9-Orthologe enthielten, wurden verwendet. Einige Sequenzen wurden nicht in das Cluster gruppiert, wurden jedoch aufgrund des Vorhandenseins von konservierten Motiven und/oder anderen Cas-Genen in ihrer unmittelbaren Nähe als Cas9-Sequenzen verifiziert.

Analyse von CRISPR-Cas-Loci

[0574] Die CRISPR-Repeat-Sequenzen wurden aus der CRISPRdb-Datenbank erhalten oder unter Verwendung der CRISPR-Finder-Anwendung (Grissa I et al., BMC Bioinformatics 2007; 8: 172; Grissa I et al., Nucleic Acids Res 2007) vorhergesagt. Die cas-Gene wurden unter Verwendung des BLASTp-Algorithmus identifiziert und/oder mit der KEGG-Datenbank (im Internet bei "www." gefolgt von kegg.jp/) verifiziert.

In silicio-Vorhersage und Analyse von tracrRNA-Orthologen

[0575] Die mutmaßlichen Anti-Repeats wurden unter Verwendung der Vektor-NTI®-Software (Invitrogen) identifiziert, indem man auf beiden Strängen des jeweiligen Genoms nach zusätzlichen, degenerierten Repeat-Sequenzen gesucht hat, die nicht zu der Repeat-Spacer-Anordnung gehörten, wobei bis zu 15 Fehlpaarungen erlaubt wurden. Die transkriptionalen Promotoren und rho-unabhängigen Terminatoren wurden unter Verwendung des Neural-Netzwerk-Promoter-Vorhersageprogramms BDGP vorhergesagt ("www." gefolgt von fruitfly.org/seq_tools/promoter.html) bzw. der TransTermHP-Software. Die multiplen Sequenz-Alignments wurden unter Verwendung des MUSCLE-Programms mit voreingestellten Parameter durchgeführt. Die Alignments wurden unter Verwendung der RNAalifold-Algorithmus des Vienna-RNA-Pakets 2.0 auf das Vorhandensein von konservierten Strukturmotiven analysiert.

ERGEBNISSE

CRISPR-Cas-Systeme vom Typ II sind in Bakterien weit verbreitet

[0576] Zusätzlich zu der tracrRNA-kodierenden DNA und dem Repeat-Spacer-Array sind CRISPR-Cas-Loci vom Typ II üblicherweise aus 3–4 Cas-Genen zusammengesetzt, die in einem Operon organisiert sind (**Fig. 32A–B**). Cas9 ist das charakteristische Protein, welches für den Typ II kennzeichnend ist, und es ist an den Schritten der Expression und der Interferenz beteiligt. Cas1 und Cas2 sind Kern-Proteine, die in allen CRISPR-Cas-Systemen vorkommen und an der Spacer-Erfassung beteiligt sind. Csn2 und Cas4 sind lediglich in einer Untergruppe von Systemen vom Typ II vorhanden, und es wird vermutet, dass sie eine Rolle bei der Adaption spielen. Um eine maximale Anzahl von CRISPR-Cas-Loci vom Typ II zu erhalten, die tracrRNA enthalten, haben wir zunächst öffentlich zugängliche Genome auf Sequenzen durchsucht, die homolog zu bereits bekannten Cas9-Proteinen sind. 235 Cas9-Orthologe wurden in 203 bakteriellen Spezies identifiziert. Eine Gruppe von 75 verschiedenen Sequenzen, die repräsentativ für alle aufgefundenen Cas9-Orthologe waren, wurde für die weitere Analyse ausgewählt (**Fig. 32, Fig. 38** und Material und Methoden).

[0577] **Fig. 32** zeigt (A) einen phylogenetischen Stammbaum von repräsentativen Cas9-Sequenzen aus verschiedenen Organismen sowie auch (B) den Aufbau eines repräsentativen Cas9-Lokus. Die Bootstrap-Werte, die für jeden Knoten berechnet wurde, sind angegeben. Äste gleicher Farbe repräsentieren ausgewählte Subcluster von ähnlichen Cas9-Orthologen. Die CRISPR-Repeat-Länge in Nukleotide, die durchschnittliche Cas9-Proteingröße in Aminosäuren (aa) und der Konsensus-Aufbau des Locus sind für jedes Subcluster gezeigt. *-gil116628213**-gil116627542 †-gil34557790 ‡-gil34557932. Typ II-A ist gekennzeichnet durch cas9-csx12, cas1, cas2, cas4. Typ II-B ist gekennzeichnet durch cas9, cas1, cas2 gefolgt von einer csn2-Variante. Type II-C ist durch ein konserviertes cas9, cas1, cas2 Operon gekennzeichnet (siehe auch **Fig. 38**).

[0578] Als nächstes haben wir ein multiples Sequenz-Alignment der ausgewählten Cas9-Orthologe durchgeführt. Die vergleichende Analyse zeigte eine hohe Diversität in der Aminosäure-Zusammensetzung und der Proteingröße. Die Cas9-Orthologe teilen lediglich einige identische Aminosäuren, und alle ermittelten Sequenzen hatten den gleichen Domänen-Aufbau, mit einer zentralen HNH-Endonuklease-Domäne und einer gespaltenen RuvC/RNaseH-Domäne. Die Längen der Cas9-Proteine reichen von 984 (*Campylobacter jejuni*) bis 1629 (*Francisella novicida*) Aminosäuren mit typischen Größen von etwa 1100 oder etwa 1400 Aminosäuren. Aufgrund der hohen Diversität der Cas9-Sequenzen, insbesondere hinsichtlich der Länge der Interdomänen-Regionen, haben wir nur gut abgeglichen, informative Positionen des erstellten Alignments ausgewählt, um einen phylogenetischen Stammbaum der analysierten Sequenzen zu rekonstruieren (**Fig. 32** und Material und Methoden). Die Cas9-Orthologe ließen sich in drei monophyletische Cluster gruppieren, wobei einige Ausreißer-Sequenzen auftraten. Die beobachtete Topologie des Cas9-Stammbaums ist in guter Übereinstimmung mit der gegenwärtigen Klassifizierung von Loci des Typs II, bei der die zuvor definierten Typen II-A und II-B getrennte monophyletische Cluster bilden. Um diese Cluster weiter zu charakterisieren, haben wir die Zusammensetzung des Cas-Operons und die CRISPR-Repeat-Sequenzen aller aufgeführten Stämme genau untersucht.

Die Cas9-Subcluster-Bildung spiegelt die Diversität im Aufbau der CRISPR-Cas-Loci vom Typ II wieder

[0579] Eine tiefergehende Analyse von ausgewählten Loci vom Typ II ergab, dass das Clustering von Cas9-Ortholog-Sequenzen mit der Diversität bezüglich der CRISPR-Repeat-Länge korreliert ist. Bei den meisten CRISPR-Cas-Systemen vom Typ II beträgt die Repeat-Länge 36 Nukleotide (nt), mit einigen Abweichungen für zwei der Cas9-Subcluster des Stammbaums. In dem Typ II-A-Cluster (**Fig. 32**), das Loci umfasst, die für das lange Cas9-Ortholog kodieren, welches früher Csx12 genannt wurde, sind die CRISPR-Repeats 37 nt lang. Das kleine Subcluster, welches aus Sequenzen von Bakterien besteht, die zu dem Phylum Bacteroidetes gehört (**Fig. 32**), ist durch ungewöhnlich lange CRISPR-Repeats gekennzeichnet, die bis zu 48 nt groß sind. Des Weiteren haben wir bemerkt, dass das Subclustering von Cas9-Sequenzen mit einem abweichenden Aufbau des cas-Operons korreliert, wie es in **Fig. 32** dargestellt ist. Das dritte Hauptcluster (**Fig. 32**) und die Ausreißer-Loci (**Fig. 32**) bestehen im Wesentlichen aus dem Minimum-Operon, das aus den Genen cas9, cas1 und cas2 zusammengesetzt ist, mit Ausnahme von einigen unvollständigen Loci, die später diskutiert werden. Alle anderen Loci der ersten beiden Hauptcluster sind mit einem vierten Gen assoziiert, hauptsächlich mit cas4, was spezifisch für Typ II-A ist, oder mit einem csn2-ähnlichen, was spezifisch für Typ II-B ist (**Fig. 32**). Wir haben Gene identifiziert, die für kürzere Varianten des Csn2-Proteins kodieren, Csn2A, und die innerhalb von Loci liegen, welche ähnlich zum Typ II-B *S. pyogenes* CRISPR01 und *S. thermophilus* CRISPR3 sind (**Fig. 32**). Die längere Variante von Csn2, Csn2b, wurde assoziiert mit Loci vorgefunden, die ähnlich zum Typ II-B *S. thermophilus* CRISPR1 sind (**Fig. 32**). Interessanterweise haben wir zusätzliche putative cas-Gene gefunden, die für Proteine mit keiner offensichtlichen Sequenz-Ähnlichkeit zu den zuvor beschriebenen Csn2-Varianten kodieren. Eines dieser nicht charakterisierten Proteine ist exklusiv mit Loci vom Typ II von *Mycoplasma*-Spezies assoziiert (**Fig. 32** und **Fig. 33**). Es wurde herausgefunden, dass zwei weitere in Loci des Typs II-B von *Staphylococcus*-Spezies kodiert werden (**Fig. 33**). In allen Fällen stimmt die Diversität des Aufbaus des cas-Operons mit dem Subclustering von Cas9-Sequenzen überein. Diese Eigenschaften gemeinsam mit der allgemeinen Topologie des Cas9-Stammbaums, der in drei verschiedene monophyletische Haupt-Cluster eingeteilt ist, haben uns dazu gebracht, eine neue Einteilung des CRISPR-Cas-Systems vom Typ II in drei Subtypen vorzunehmen. Typ II-A ist mit Csx12-ähnlichem Cas9 und Cas4 assoziiert, Typ II-B ist mit Csn2-ähnlichem assoziiert und Typ II-C enthält lediglich die minimale Gruppe der Gene cas9, cas1 und cas2, wie es in **Fig. 32** dargestellt ist.

[0580] **Fig. 33** zeigt den Aufbau von CRISPR-Cas vom Typ II aus ausgewählten bakteriellen Spezies. Die vertikalen Balken gruppieren die Loci, die für Cas9-Orthologe kodieren, welche zum gleichen Stammbaum-Subcluster gehören (vergleiche mit **Fig. 32**). Horizontale schwarze Balken, Leader-Sequenz; schwarze Rechtecke und Diamanten, Repeat-Spacer-Array. Vorhergesagte Anti-Repeats werden durch Pfeile dargestellt, welche die Richtung der mutmaßlichen Transkription des tracrRNA-Orthologs angeben. Es ist zu beachten, dass bei

den Loci, die nicht experimentell verifiziert wurden, vorliegend angenommen wird, dass die CRISPR-Repeat-Spacer vom selben Strang wie das Cas-Operon transkribiert werden. Die Transkriptionsrichtung des mutmaßlichen tracrRNA-Orthologs ist entsprechend angegeben.

In silicio-Vorhersagen neuer tracrRNA-Orthologe

[0581] Typ II-Loci, die zuvor basierend auf den 75 repräsentativen Cas9-Orthologen ausgewählt worden waren, wurden auf das Vorhandensein putativer tracrRNA-Orthologe untersucht. Unsere vorherige Analyse, die basierend auf einer begrenzten Anzahl von tracrRNA-Sequenzen durchgeführt wurde, ergab, dass weder die Sequenzen der tracrRNAs noch deren Lokalisierung innerhalb der CRISPR-Cas-Loci konserviert zu sein scheinen. Wie jedoch oben erwähnt ist, sind tracrRNAs ferner durch eine Anti-Repeat-Sequenz charakterisiert, die in der Lage ist, mit jedem der prä-crRNA-Repeats eine Basenpaarung zu bilden, wobei tracrRNA:prä-crRNA-Repeat-Duplexe gebildet werden, die durch RNase III in Gegenwart von Cas9 gespalten werden. Um neue tracrRNAs vorherzusagen, haben wir diese Eigenschaft ausgenutzt und den folgenden Arbeitsablauf verwendet: (i) Screening nach potentiellen Anti-Repeats (Sequenz-Basenpaarung mit CRISPR-Repeats) innerhalb der CRISPR-Cas-Loci, (ii) Auswahl von Anti-Repeats, die in den intergenischen Regionen lokalisiert sind, (iii) Validierung der CRISPR-Anti-Repeat:Repeat-Basenpaarung, und (iv) Vorhersage von Promotoren und Rho-unabhängigen transkriptionalen Terminatoren, die mit den identifizierten tracrRNAs assoziiert sind.

[0582] Um nach putativen Anti-Repeats zu suchen, haben wir Repeat-Sequenzen aus der CRISPRdb-Datenbank erhalten oder wir haben, wenn diese Informationen nicht verfügbar waren, die Repeat-Sequenzen unter Verwendung der CRISPRfinder-Software vorhergesagt. In unserer früheren Untersuchung haben wir experimentell gezeigt, dass die Richtung der Transkription des Repeat-Spacer-Arrays im Vergleich zu der des cas-Operons zwischen den Loci variierte. Diese Beobachtung wurde vorliegend durch RNA-Sequenzierungs-Analysen bestätigt. In einigen der analysierten Loci, nämlich in *F. novicida*, *N. meningitidis* und *C. jejuni*, wird der Repeat-Spacer-Array in entgegengesetzter Richtung zum cas-Operon transkribiert (siehe Abschnitt "Deep RNA-Sequenzierung bestätigt die Expression von neuen tracrRNA-Orthologen" sowie **Fig. 33** und **Fig. 34**), während in *S. pyogenes*, *S. mutans*, *S. thermophilus* und *L. innocua* der Array und das cas-Operon in derselben Richtung transkribiert werden. Dies sind die einzigen Typ II-Repeat-Spacer-Array-Expressionsdaten, die bis dato verfügbar sind. Um die Transkriptionsrichtung von anderen Repeat-Spacer-Arrays vorherzusagen, haben wir die frühere Beobachtung berücksichtigt, nach der die letzten Repeats der Arrays normalerweise mutiert sind. Diese Bemerkung steht im Einklang mit dem gegenwärtigen Spacer-Übernahme-Modell, in dem üblicherweise der erste Repeat des Arrays bei Insertion einer Spacer-Sequenz während der Adaptionsphase dupliziert wird. Für 37 Repeat-Spacer-Arrays waren wir in der Lage, den mutierten Repeat am mutmaßlichen Ende des Arrays zu identifizieren. Wir beobachteten, dass die vorhergesagte Richtung der Transkription für die Repeat-Spacer-Arrays von *N. meningitidis* und *C. jejuni* gegensätzlich zur experimentell bestimmten Richtung wäre (RNA-Sequenzierung und Northern Blot-Analyse). Da die vorhergesagte Richtung innerhalb der Cluster nicht übereinstimmte und wir in den meisten Fällen potentielle Promotoren an beiden Enden des Arrays nachweisen konnten, haben wir angenommen, dass die Transkription der Repeat-Spacer-Arrays in dieselbe Richtung verläuft, wie die Transkription des cas-Operons, sofern nichts anderes bestätigt wurde.

[0583] **Fig. 34** zeigt die Co-Prozessierung von tracrRNA und prä-crRNA in ausgewählten CRISPR-Cas-Systemen vom Typ II. Der Aufbau von CRISPR-Loci mit verifizierten Positionen und Richtungen von tracrRNA- und prä-crRNA-Transkription sind gezeigt. Obere Sequenzen, prä-crRNA-Repeats; untere Sequenzen, tracrRNA-Sequenzen, die eine Paarung mit den crRNA-Repeats ausbilden. Putative Stellen für die RNA-Prozessierung, die durch RNA-Sequenzierung ermittelt wurden, sind mit Pfeilköpfen angegeben. Für jeden Locus stellen die Größen der Pfeilköpfe die relativen Mengen der erhaltenen 5'- und 3'-Enden dar (siehe auch **Fig. 37**).

[0584] **Fig. 37** führt alle tracrRNA-Orthologe und reifen crRNAs auf, die durch Sequenzierung für die untersuchten bakteriellen Spezies ermittelt wurden, einschließlich Koordinaten (Regionen von Interesse) und entsprechender cDNA-Sequenzen (5' zu 3'). Die Pfeile stellen die Transkriptionsrichtung (Strang) dar. Die Anzahl von cDNA-Auslesungen (unter Verwendung von SAMtools berechnet), die Zahlen zur Abdeckung (prozentualer Anteil der kartierten Auslesungen) und die mit jedem Transkript vorwiegend assoziierten Enden sind angegeben. Die Anzahl der Auslesungen, die bei jeder Nukleotidposition im Bereich des 5'- und des 3'-Endes eines jeden Transkripts beginnen oder enden, sind dargestellt. Die Größen jeder reifen crRNA-Form sind angegeben. Die Zahl, die jeder crRNA-Spezies zugewiesen wurde, entspricht der Spacer-Sequenz-Position in der prä-crRNA, gemäß der CRISPRdb. Die Zahl, die jeder tracrRNA-Spezies zugewiesen wurde, entspricht unterschiedlichen Formen desselben Transkripts.

[0585] Wir haben anschließend die ausgewählten CRISPR-Cas-Loci einschließlich der Sequenzen, die 1 kb stromaufwärts und stromabwärts auf beiden Strängen lokalisiert waren, nach möglichen Repeat-Sequenzen untersucht, die nicht zu dem Repeat-Spacer-Array gehörten, wobei 15 Fehlpaarungen erlaubt wurde. Wir haben durchschnittlich 1–3 degenerierte Repeat-Sequenzen pro Locus gefunden, die den Anti-Repeats von tracrRNA-Orthologen entsprechen würden, und wir haben die Sequenzen ausgewählt, die innerhalb der intergenischen Region lokalisiert sind. Die putativen Anti-Repeats wurden an 4 typischen Stellen gefunden: stromaufwärts vom cas9-Gen, in der Region zwischen cas9 und cas1, und stromaufwärts oder stromabwärts vom Repeat-Spacer-Array (**Fig. 33**). Für jede erhaltene Sequenz haben wir das Ausmaß der Basenpaarung validiert, die zwischen dem Repeat und dem Anti-Repeat ausgebildet wurde (**Fig. 44**), indem die mögliche RNA:RNA-Interaktion vorhergesagt wurde, wobei ein besonderer Fokus auf Kandidaten mit längerer oder perfekter Komplementaritätsregion lag, welche eine optimale doppelsträngige Struktur für die RNase III-Prozessierung bildete. Um Promotoren und transkriptionale Terminatoren vorherzusagen, die den Anti-Repeat flankieren, haben wir den putativen Transkription-Start und die Terminations-Stellen basierend auf unseren früheren Beobachtungen²⁶ so festgelegt, dass sie von einer Region, die maximal 200 nt stromaufwärts und 100 nt stromabwärts von der Anti-Repeat-Sequenz lokalisiert ist, eingeschlossen waren. Wie oben erwähnt, fehlen bei meisten Repeat-Spacer-Arrays von Systemen des Typs II experimentelle Informationen zur Transkriptionsrichtung. Die in silicio-Algorithmen zur Promoter-Vorhersage geben oftmals falsch-positive Ergebnisse und verweisen auf putative Promotoren, die zur Transkription von Repeat-Spacer-Arrays von beiden Strängen führen würden. In einigen Fällen konnten wir keine transkriptionale Terminatoren bestimmen, obwohl die Expression des tracrRNA-Orthologs experimentell bestätigt werden konnte, wie beispielsweise beim Locus von *C. jejuni* (siehe Abschnitt "Deep RNA-Sequenzierung bestätigt die Expression von neuen tracrRNA-Orthologen"). Wir schlagen vor, die Vorhersagen zu Promoter und Transkriptions-Terminator lediglich als unterstützenden Schritt, nicht jedoch als wesentlichen Schritt der oben beschriebenen Richtlinie anzusehen.

[0586] **Fig. 44** zeigt die vorhergesagte prä-crRNA-Repeat:tracrRNA-Anti-Repeat-Basenpaarung in ausgewählten bakteriellen Spezies. "Die CRISPR-Loci gehören zum Typ II (Nmeni/CASS4)-CRISPR-Cas-System. Die Nomenklatur erfolgt gemäß der CRISPR-Datenbank (CRISPRdb). Es ist zu beachten, dass *S. thermophilus* LMD-9 und *W. succinogenes* zwei Loci vom Typ II enthalten. ^aObere Sequenz, prä-crRNA-Repeat-Konsensus-Sequenz (5' nach 3'); untere Sequenz, tracrRNA-Homolog-Sequenz, die sich mit dem Repeat zusammenlagert (Anti-Repeat; 3' nach 5'). Es ist zu beachten, dass die angegebene Repeat-Sequenz auf der Annahme beruht, dass der CRISPR-Repeat-Spacer-Array vom selben Strang transkribiert wird wie das cas-Operon. Für die Sequenzen, die in dieser Untersuchung experimentell bestätigt wurden, wurden Daten aus der RNA-Sequenzierung berücksichtigt, um die Basenpaarung zu bestimmen (siehe **Fig. 33**). ^dZwei Mögliche Anti-Repeats wurden in den Loci vom Typ II-A von *F. tularensis* subsp. *novicida*, *W. succinogenes* und *Gamma-Proteobakterium* HTCC5015 gefunden. Obere Sequenzpaarung, Anti-Repeat innerhalb der putativen Leader-Sequenz; untere Sequenzpaarung; Anti-Repeat stromabwärts vom Repeat-Spacer-Array. Siehe **Fig. 33**. ^eZwei Mögliche Anti-Repeats wurden im Locus vom Typ II-A von *S. wadsworthensis* gefunden. Obere Sequenzpaarung, Anti-Repeat; untere Sequenzpaarung, Anti-Repeat innerhalb der putativen Leader-Sequenz. Siehe **Fig. 33**. ^fZwei Mögliche Anti-Repeats wurden im Locus vom Typ II-B von *L. gasseri* gefunden. Obere Sequenzpaarung, Anti-Repeat stromaufwärts von cas9; untere Sequenzpaarung; Anti-Repeat zwischen den Genen cas9 und cas1. Siehe **Fig. 33**. ^gZwei mögliche Anti-Repeats wurden in den Loci vom Typ II-C von *C. jejuni* gefunden. Obere Sequenzpaarung, Anti-Repeat stromaufwärts von cas9; untere Sequenzpaarung, Anti-Repeat stromabwärts des Repeat-Spacer-Arrays. Siehe **Fig. 33**. ^hZwei mögliche Anti-Repeats wurden im Locus vom Typ II-C von *R. rubrum* gefunden. Obere Sequenzpaarung, Anti-Repeat stromabwärts vom Repeat-Spacer-Array; untere Sequenzpaarung, Anti-Repeat stromaufwärts von cas1. Siehe **Fig. 33**.

Eine Fülle von tracrRNA-Orthologen

[0587] Wir haben putative tracrRNA-Orthologe für 56 der zuvor ausgewählten 75 Loci vorhergesagt. Die Ergebnisse der Vorhersagen sind in **Fig. 33** dargestellt. Wie bereits erwähnt, ist die Richtung der tracrRNA-Transkription, auf die in dieser Figur verwiesen wird, hypothetisch und basiert auf der angegebenen Richtung der Transkription des Repeat-Spacer-Arrays. Wie zuvor erwähnt, wurden Sequenzen, die für putative tracrRNA-Orthologe kodieren, stromaufwärts vom cas-operon, im cas-Operon und stromabwärts vom cas-Operon gefunden, sowie auch stromabwärts von den Repeat-Spacer-Arrays, einschließlich in den putativen Leader-Sequenzen, die üblicherweise in Loci des Typs II-A gefunden werden (**Fig. 33**). Wir haben jedoch beobachtet, dass Anti-Repeats mit ähnlicher Lokalisierung innerhalb der CRISPR-Cas-Loci in unterschiedlichen Richtungen transkribiert werden können (wie es beobachtet wurde, als z. B. *Lactobacillus rhamnosus* und *Eubacterium rectale* oder *Mycoplasma mobile* und *S. pyogenes* oder *N. meningitidis* verglichen wurden) (**Fig. 33**). Es ist festzustellen, dass Loci, die innerhalb eines gleichen Subclusters des Cas9-Guide-Stammbaums gruppiert sind, einen gemeinsamen Aufbau in Bezug auf die Position des Gens, welches für die tracrRNA kodiert, auf-

weisen. Wir haben Anti-Repeats im Bereich des Repeat-Spacer-Arrays in Loci vom Typ II identifiziert, meistens stromaufwärts vom cas9-Gen in den Typen II-B und II-C, mit einigen bemerkenswerten Ausnahmen, bei denen die putative tracrRNA zwischen cas9 und cas1 in drei verschiedenen Subclustern vom Typ II-B lokalisiert war.

Einige CRISPR-Cas-Loci vom Typ II haben defekte Repeat-Spacer-Arrays und/oder tracrRNA Orthologe

[0588] Für sechs Loci vom Typ II (*Fusobacterium nucleatum*, *Aminomonas paucivorans*, *Helicobacter mustelae*, *Azospirillum* sp., *Prevotella ruminicola* und *Akkermansia muciniphila*) haben wir potentielle Anti-Repeats mit schwacher Basenpaarung mit der Repeat-Sequenz identifiziert oder solche, die im offenen Leserahmen lokalisiert waren. Es ist anzumerken, dass in diesen Loci folgendes identifiziert wurde: ein schwacher Anti-Repeat innerhalb des offenen Leserahmens des Gens, welches für eine putative ATPase in *A. paucivorans* kodiert, ein starker Anti-Repeat innerhalb der ersten 100 nt des cas9-Gens in *Azospirillum* sp. B510 und ein starker Anti-Repeat, der sowohl mit cas9 als auch mit cas1 in *A. muciniphila* überlappt (**Fig. 33**). Für 12 zusätzliche Loci (*Peptoniphilus duerdenii*, *Coprococcus catus*, *Acidaminococcus intestini*, *Catenibacterium mitsuokai*, *Staphylococcus pseudintermedius*, *Ilyobacter polytropus*, *Elusimicrobium minutum*, *Bacteroides fragilis*, *Acidothermus cellulolyticus*, *Corynebacterium diphtheriae*, *Bifidobacterium longum* und *Bifidobacterium dentium*) konnten wir keinen putativen Anti-Repeat nachweisen. Es gibt keine verfügbaren Informationen zur Expression der prä-crRNA und zur Prozessierung in diesen CRISPR-Cas-Loci. Daher muss die Funktionalität von Typ II-Systemen in Abwesenheit eines klar definierten tracrRNA-Orthologs weiter untersucht werden. Für sieben analysierte Loci konnten wir keinen Repeat-Spacer-Array identifizieren (*Parasutterella excrementihominis*, *Bacillus cereus*, *Ruminococcus albus*, *Rhodopseudomonas palustris*, *Nitrobacter hamburgensis*, *Bradyrhizobium* sp. und *Prevotella micans*) (**Fig. 33**), und in dreien davon (*Bradyrhizobium* sp. BTAi1, *N. hamburgensis* und *B. cereus*) haben wir cas9 als einzelnes Gen nachgewiesen, wobei keine anderen cas-Gene in der Umgebung lagen. Für drei dieser Loci konnten wir keine kleine RNA-Sequenz stromaufwärts oder stromabwärts des cas9-Gens vorhersagen. Im Falle von *R. albus* und *P. excrementihominis* ist der cas9 enthaltende genomische Contig zu kurz, um eine Vorhersage des Repeat-Spacer-Arrays zu ermöglichen

Deep-RNA-Sequenzierung bestätigt die Expression von neuen tracrRNA-Orthologen

[0589] Um die in silicio-Vorhersagen zur tracrRNA zu verifizieren und die Muster der tracrRNA:prä-crRNA-Co-Prozessierung zu bestimmen, wurden RNAs aus ausgewählten Gram-positiven (*S. mutans* und *L. innocua*) und Gram-negativen (*N. meningitidis*, *C. jejuni* und *F. novicida*) Bakterien durch Deep-Sequenzierung analysiert. Die Sequenzen von tracrRNA-Orthologen und prozessierten crRNAs wurden erhalten (**Fig. 36** und **Fig. 37**). In Übereinstimmung mit zuvor veröffentlichten Daten einer differenziellen tracrRNA-Sequenzierung in *S. pyogenes*²⁶ waren tracrRNA-Orthologe in den Bibliotheken stark vertreten, mit einem Bereich von 0,08 bis 6,2% der gesamt kartierten Auslesungen. Prozessierte tracrRNAs waren ebenfalls stärker vertreten als primäre Transkripte und reichten von 66% bis 95% der Gesamtzahl der tracrRNA-Auslesungen (**Fig. 36** und **Fig. 37**).

[0590] **Fig. 36** stellt die Expression von bakteriellen tracrRNA-Orthologen und crRNAs dar, die durch Deep-RNA-Sequenzierung erhalten wurden. Die Expressionsprofile von tracrRNA-Orthologen und crRNAs aus ausgewählten bakteriellen Stämmen sind gemeinsam mit den entsprechenden Genomen in Säulendiagrammen dargestellt (Abbildungen vom Integrative Genomics Viewer (IGV)-Tool entnommen). *Campylobacter jejuni* (GenBank: NC_002163), *Francisella novicida* (GenBank: NC_008601), *Neisseria meningitidis* (GenBank: NC_003116), *Listeria innocua* (GenBank: NC_003212) und *Streptococcus mutans* (GenBank: NC_004350). Die genomischen Koordinaten sind angegeben. ^aSequenzabdeckung, wie sie unter Verwendung der BEDTools-Version-2.15.0 berechnet wurde (Abmessung in Auslesungen (reads) pro Millionen angegeben). ^bDie Verteilung der Auslesungen beginnend (5') und endend (3') ist an jeder Nukleotidposition angegeben (Abmessungen als Anzahl von Auslesungen angegeben). Die oberen Teilabbildungen entsprechen Transkripten vom positiven Strang, und die unteren Teilabbildungen entsprechen Transkripten vom negativen Strang. Die negativen Abdeckungswerte und Peaks, die unter den Achsen dargestellt sind, verweisen auf eine Transkription vom Negativ-Strang des Genoms. Vorwiegende 5'- und 3'-Enden der Auslesungen sind für alle RNAs aufgetragen. Es ist zu beachten, dass im Hinblick auf die geringe Qualität der *L. innocua* cDNA-Bibliothek die Auslesungen für crRNA gekürzt wurden, und eine Ansammlung der Auslesungen am 3'-Ende von tracrRNA beobachtet wurde, vermutlich aufgrund von RNA-Abbau.

[0591] Um die 5'-Enden von primären tracrRNA-Transkripten zu ermitteln, haben wir das Vorhandensein aller Auslesungen am 5'-Ende von tracrRNA analysiert, und wir haben die am stärksten hervorgehobenen Auslesungen stromaufwärts oder in der Nähe des 5'-Endes der vorhergesagten Anti-Repeat-Sequenz erhalten. Die 5'-Enden von tracrRNA-Orthologen wurden ferner unter Verwendung des Promoter-Vorhersage-Algorithmus bestätigt. Die identifizierten 5'-Enden von tracrRNAs von *S. mutans*, *L. innocua* und *N. meningitidis* korrelieren

sowohl mit den in silicio-Vorhersagen als auch mit der Northern-Blot-Analyse der tracrRNA-Expression²⁶. Das am stärksten hervorgehobene 5'-Ende der tracrRNA von *C. jejuni* wurde in der Mitte der Anti-Repeat-Sequenz identifiziert. Fünf Nukleotide stromaufwärts wurde ein zusätzliches putatives 5'-Ende nachgewiesen, welches mit der in silicio-Vorhersage korreliert und eine längere Sequenz für die Interaktion mit der CRISPR-Repeat-Sequenz bereitstellt. Wir erhielten relativ geringe Mengen von Auslesungen für die Bibliothek von *F. novicida*, die fast ausschließlich prozessierten Transkripten entsprachen. Die Analyse der sehr geringen Menge von Auslesungen von primären Transkripten ergab ein 5'-Ende, das den starken in silicio-Promotor-Vorhersagen entsprach. Northern-Blot-Untersuchungen von tracrRNA aus *F. novicida* bestätigten ferner die Gültigkeit der Vorhersagen und zeigten ein geringes Vorhandensein von Transkripten einer Länge von etwa 90 nt. Die Ergebnisse sind in Tabelle 2 aufgeführt. Für alle untersuchten Spezies, außer für *N. meningitidis*, wurden primäre tracrRNA-Transkripte als einzelne kleine RNA-Spezies mit einer Länge von 75 bis 100 nt identifiziert. Im Falle von *N. meningitidis* haben wir eine vorherrschende primäre tracrRNA-Form von etwa 110 nt sowie ein putatives, längeres Transkript von etwa 170 nt gefunden, das von einer sehr geringen Menge von Auslesungen repräsentiert wird und zuvor als schwache Bande in der Northern Blot-Analyse nachgewiesen wurde.

Tabelle 2 ausgewählte tracrRNA-Orthologe

Stämme ^a	Transkript	5'-Ende ^b			3'-Ende ^c	Länge (nt)
		RNA-seq		berechnet		
		erste Auslesung	am deutlichsten			
<i>S. pyogenes</i> SF370	<i>Primär</i>	-	854 546	-	854 376	171
	<i>Primär</i>	-	<u>854 464</u>	-		89
	<i>Prozessiert</i>	-	854 450	-		~75
<i>C. jejuni</i> NCTC 11168	<i>Primär</i>	<u>1 455 497</u>	1 455 502	<u>1 455 497</u>	1 455 570	~75
	<i>Prozessiert</i>	-	1 455 509	-		~60
<i>L. innocua</i> Clip11262	<i>Primär</i>	<u>2 774 774</u>	<u>2 774 774</u>	2 774 773	2 774 863	~90
	<i>Prozessiert</i>	-	2 774 788	-		~75
<i>S. mutans</i> UA159	<i>Primär</i>	<u>1 335 040</u>	<u>1 335 040</u>	1 355 039	1 335 141	~100
	<i>Prozessiert</i>	-	1 335 054 1 335 062	-		~85 ~80
<i>N. meningitidis</i> A Z2491	<i>Primär</i>	614 158	614 162	614 154	614 333	~175
	<i>Primär</i>	614 223	614 225	<u>614 223</u>		~110
	<i>Prozessiert</i>	-	614 240	-		~90
<i>F. novicida</i> U112	<i>Primär</i>	817 144	-	817 145 <u>817 154</u>	817 065	~80
	<i>Prozessiert</i>	-	817 138 817 128	-		~75 ~65
<i>S. thermophilus</i> LMD-9	<i>Primär</i>	-	-	<u>1 384 330</u>	1 384 425	~95
	<i>Primär</i>	-	-	<u>646 654</u>	646 762	~110
<i>P. multocida</i> Pm70	<i>Primär</i>	-	-	<u>1 327 287</u>	1 327 396	~110
<i>M. mobile</i> 163K	<i>Primär</i>	-	-	<u>49 470</u>	49 361	~110

^atracrRNA-Orthologe von *S. thermophilus*, *P. multocida* und *M. mobile* wurden in silicio vorhergesagt

^bRNA-seq ermittelt durch RNA-Sequenzierung (Tabelle S3); erste Auslesung, erste durch Sequenzierung ermittelte Position am 5'-Ende; am stärksten hervorgehoben, im Überfluss vorhandenes 5'-Ende gemäß RNA-seq-Daten; vorhergesagt, in silicio-Vorhersage der Transkriptions-Start-Stelle; unterstrichen; 5q'-Ende ausgewählt als primäre tracrRNA für ein Alignment.

geschätztes 3'-Ende gemäß RNA-seq-Daten und Vorhersage des transkriptionalen Terminators.

Die tracrRNA und prä-crRNA-Co-Prozessierungsstellen liegen im Bereich des Anti-Repeat:Repeat

[0592] Wir haben die prozessierten tracrRNA-Transkripte untersucht, indem wir reichlich vorhandene tracrRNA-5'-Enden innerhalb der vorhergesagten Anti-Repeat-Sequenz und reichlich vorhandene, reife crRNA-3'-Enden analysiert haben (**Fig. 34** und **Fig. 45**). In allen Spezies haben wir die markanten 5'-Enden von tracrRNA-Orthologen identifiziert, die aus der Co-Prozessierung der tracrRNA:prä-crRNA-Repeat-Duplexe durch RNase III resultieren könnten. Wir haben ferner die prozessierten 5'-Enden von crRNAs identifiziert, die höchstwahrscheinlich aus einem zweiten Reifungs-Ereignis durch ein vermeintliches Zurechtschneiden resultieren, was mit vorherigen Beobachtungen übereinstimmen würde. Es ist anzumerken, dass wir in den eng verwandten RNA-Paaren von *S. pyogenes*, *S. mutans* und *L. innocua* die gleiche Prozessierungs-Stelle um das G:C-Basenpaar in der Mitte der Anti-Repeat-Sequenz beobachtet haben. Sowohl in *S. mutans* als auch in *L. innocua* haben wir zusätzliche markante tracrRNA-5'-Enden und crRNA-3'-Enden nachgewiesen, die auf ein weiteres Zurechtschneiden des tracrRNA-crRNA-Duplex verweisen könnten, wobei das 3'-Ende der crRNA nach dem durch RNase III katalysierten ersten Prozessierungs-Ereignis zusätzlich zu dem bereits erwähnten Zurechtschneiden des 5'-Endes gekürzt wird. In ähnlicher Weise haben wir in *C. jejuni* lediglich eine kleine Menge von crRNA-3'-Enden gefunden, die zu den RNase III-Prozessierungsmustern passen würden, und wir haben die entsprechenden 5'-Enden der prozessierten tracrRNA erhalten. Somit würde das vermeintliche Zurechtschneiden der tracrRNA:crRNA-Duplexe nach anfänglicher Spaltung durch RNase III zu einem kürzeren Teil in den reifen crRNAs führen, der von dem Repeat abgeleitet ist, was kürzere tracrRNA:crRNA-Duplexe erzeugt, die durch eine dreifache G:C-Basenpaarung für eine Interaktion mit der Endonuklease 9 und der nachfolgenden Spaltung von Ziel-DNAs stabilisiert wird. Der RNA-Duplex von *N. meningitidis* scheint an zwei primären Stellen prozessiert zu werden, die weiter am 3'-Ende des CRISPR-Repeats liegen, was zu einem langen, vom Repeat abgeleiteten Teil in der reifen crRNA und zu einer stabilen RNA:RNA-Interaktion trotz der zentralen Wölbung innerhalb des Duplex führt. Interessanterweise scheint der tracrRNA:prä-crRNA-Duplex von *F. novicida* innerhalb der Region geringer Komplementarität gespalten zu werden, und einige der erhaltenen, reichlich vorhandenen 5'-Enden der tracrRNA verweisen auf ein weiteres Zurechtschneiden derselben ohne gleichzeitiges Zurechtschneiden der crRNA. Unterschiede in Bezug auf die Größe der primären Transkripte und in Bezug auf den Ort der Prozessierungsstellen führen zu verschiedenen Längen von prozessierten tracrRNAs, die von etwa 65 bis etwa 85 nt reichen. Die Koordinaten und Größen der markanten, prozessierten tracrRNA-Transkripte sind in Tabelle 2 und **Fig. 37** gezeigt. Die beobachteten Prozessierungsmuster von tracrRNA und crRNA sind in guter Übereinstimmung mit den zuvor vorgeschlagenen Modellen für zwei Reifungs-Ereignisse. Das vermeintliche weitere Zurechtschneiden von einigen der tracrRNA-5'-Enden und crRNA-3'-Enden könnte von einem zweiten Reifungs-Ereignis stammen oder alternativ dazu ein Artefakt der Herstellung der cDNA-Bibliothek oder der RNA-Sequenzierung sein. Der Ursprung dieser Prozessierungen muss weiter untersucht werden.

Die Sequenzen von tracrRNA-Orthologen sind in hohem Maße divers

[0593] Sequenz-Ähnlichkeiten von ausgewählten tracrRNA-Orthologen wurden ebenfalls bestimmt. Wir haben multiple Sequenz-Alignments von primären tracrRNA-Transkripten von *S. pyogenes* (lediglich 89 nt-Form), *S. mutans*, *L. innocua* und *N. meningitidis* (lediglich 110 nt-Form), *S. thermophilus*, *P. multocida* und *M. mobile* durchgeführt (Tabelle 2, **Fig. 35**). Wir haben eine hohe Diversität bei den tracrRNA-Sequenzen beobachtet, jedoch eine signifikante Konservierung bei Sequenzen von eng verwandten CRISPR-Cas-Loci. tracrRNAs aus *L. innocua*, *S. pyogenes*, *S. mutans* und *S. thermophilus* teilen durchschnittlich 77% Identität, und tracrRNAs von *N. meningitidis* und *P. multocida* teilen 82% Identität gemäß eines paarweise erfolgenden Alignments. Die durchschnittliche Identität der analysierten tracrRNA-Sequenzen beträgt 56%, vergleichbar zur Identität von zufälligen RNA-Sequenzen. Diese Beobachtung bestätigt ferner, dass die Vorhersage von tracrRNA-Orthologen, welche auf einer Sequenz-Ähnlichkeit basiert, lediglich im Falle von eng verwandten Loci durchgeführt werden kann. Wir haben ferner mögliche tracrRNA-Struktur-Konservierungen gesucht, konnten jedoch keine signifikante Ähnlichkeit finden, außer einer Co-Variation und einer konservierten Transkription-Terminations-Struktur (**Fig. 35**).

[0594] **Fig. 35** zeigt die Sequenz-Diversität von tracrRNA-Orthologen. Ein multiples Sequenz-Alignment mit tracrRNA. *S. thermophilus* und *S. thermophilus* 2, tracrRNA, die mit den Cas9-Orthologen von SEQ ID NO: 41 und SEQ ID NO: 40 assoziiert ist. Schwarz, hochkonserviert; dunkles grau, konserviert; helles grau, schwach konserviert. Die vorhergesagte Konsensus-Struktur ist oben im Alignment dargestellt. Pfeile verweisen auf Nukleotid-Konservierungen. tracrRNAs von *S. pyogenes* SF370, *S. mutans* UA159, *L. innocua* Clip11262, *C. jejuni* NCTC 11168, *F. novicida* U112 und *N. meningitidis* A Z2491 wurden durch RNA-Sequenzierung und

Northern-Blot-Analyse bestätigt. *S. thermophilus* LMD-9 tracrRNA wurde durch Northern-Blot-Analyse bestätigt. *P. multocida* Pm70-tracrRNA wurde auf Basis der hohen Ähnlichkeit des CRISPR-Cas-Locus mit dem von *N. meningitidis* A Z2491 vorhergesagt. *M. mobile* 163K-tracrRNA wurde in silicio auf Basis der guten Vorhersagen des transkriptionalen Promotors und des Terminators vorhergesagt.

Beispiel 4: Umwidmung von CRISPR zu einer RNA-geleiteten Plattform für die Sequenz-spezifische Kontrolle der Gen-Expression

[0595] Die gezielte Genregulation in einem Genom-weiten Umfang ist eine leistungsstarke Strategie für das Abfragen, die Störung und die Entwicklung von zellulären Systemen. Die Erfinder haben ein neues Verfahren für die Kontrolle der Gen-Expression entwickelt, das auf Cas9, einer RNA-geführten DNA-Endonuklease aus einem CRISPR-System vom Typ II, basiert. Dieses Beispiel zeigt, dass ein katalytisch totes Cas9, dem eine Endonuklease-Aktivität fehlt, bei Co-Expression mit einer Guide-RNA einen DNA-Erkennungskomplex bildet, der spezifisch mit der Elongation der Transkription, der RNA-Polymerase-Bindung oder der Bindung eines Transkriptionsfaktors interferieren kann. Dieses System, das als CRISPR-Interferenz (CRISPRi) bezeichnet wird, kann in *Escherichia coli* die Expression von Genen, auf die abgestellt wird, ohne nachweisbare unspezifische Wirkung wirkungsvoll reprimieren. CRISPRi kann verwendet werden, um mehrere Ziel-Gene gleichzeitig zu reprimieren, und seine Wirkungen sind reversibel. Darüber hinaus kann das System für eine Gen-Repression in Säugetier-Zellen angepasst werden. Diese durch RNA geführte DNA-Erkennungsplattform stellt einen einfachen Ansatz für die selektive Störung der Gen-Expression in einem Genom-weiten Umfang zur Verfügung.

MATERIAL UND METHODEN

Stämme und Medien

[0596] Der *Escherichia coli* K12-Stamm MG1655 wurde als Wirtstamm für die in vivo Fluoreszenzmessungen verwendet. Ein von *E. coli* MG1655 abgeleiteter Stamm, der endogen eine Variante von RNAP mit einem 3x-FLAG-Epitop-Tag exprimiert, welcher an das C-terminale Ende der RpoC-Untereinheit angeheftet war, wurde für alle Sequenzierungsversuche verwendet. EZ-reiches definiertes Medium (EZ-RDM, Teknoka) wurde als Wachstumsmedium für die in vivo Fluoreszenz-Assays verwendet. Die genetische Transformation und die Bestätigung der Transformation wurden unter Verwendung von Standard-Protokollen durchgeführt, wobei AmpR-, CmR- oder KanR-Gene als selektierbare Marker verwendet wurden.

Plasmid-Konstruktion und Genom-Klonierung von *E. coli*

[0597] Die Gene Cas9 und dCas9 wurden ausgehend von den zuvor beschriebenen Vektoren pMJ806 bzw. pMJ841 kloniert. Die Gene wurden mittels PCR amplifiziert und in einen Vektor inseriert, der einen Anhydrotetracyclin(aTc)-induzierbaren Promoter PLtetO-1, einen Chloramphenicol-selektierbaren Marker und einen p15A-Replikationsursprung umfasste. Die sgRNA-Matrize wurde in einen Vektor kloniert, der einen minimalen synthetischen Promotor (J23119) mit einem annotierten Transkriptions-Startpunkt, einen Ampicillin-selektierbaren Marker und einen ColE1-Replikationsursprung umfasste. Eine inverse PCR wurde verwendet, um sgRNA-Kassetten mit 20 bp-komplementären Regionen zu erzeugen. Um Fluoreszenz-Reporter-Gene in *E. coli*-Genome zu inserieren, wurde das Fluoreszenz-Gen zunächst auf einen Eingangsvektor kloniert, der anschließend mittels PCR amplifiziert wurde, um linearisierte DNA-Fragmente zu erzeugen, die nsfA 5'/3'-UTR-Sequenzen, das Fluoreszenz-Gen und einen KanR-selektierbaren Marker enthielten. Der *E. coli*-Stamm MG1655 wurde mit einem Temperatur-sensitiven Plasmid pKD46 transformiert, das die λ -Red Rekombinations-Proteine enthielt (Exo, Beta und Gamma). Zellkulturen wurden bei 30°C bis zu einer OD (600 nm) von etwa 0,5 angezüchtet, und es wurden 0,2% Arabinose zugegeben, um die Expression des λ -Red-Rekombinations-proteins für eine Stunde zu induzieren. Die Zellen wurden bei 4°C geerntet und für die Transformation der linearisierten DNA-Fragmente durch Elektroporation verwendet. Zellen, die die korrekten Genom-Insertionen enthielten, wurden durch Verwendung von 50 µg/ml Kanamycin selektiert.

Durchflusszytometrie und Analyse

[0598] Die Stämme wurden in EZ-RDM kultiviert, das 100 µg/ml Carbenicillin und 34 µg/ml Chloramphenicol in 2 ml-Deep-Well-Platten mit 96 Vertiefungen (Costar 3960) enthält, über Nacht bei 37°C und 1200 UpM kultiviert. 1 µl dieser Übernachtskultur wurde anschließend zu 249 µl frischem EZ-RDM mit den gleichen Antibiotika-Konzentrationen gegeben und mit 2 µM aTc supplementiert, um die Erzeugung des dCas9-Proteins zu induzieren. Wenn die Zellen bis zur mittleren Log-Phase (etwa 4 h) angezüchtet wurden, wurden die Mengen des Fluoreszenz-Proteins unter Verwendung des LSRII-Durchflusszytometers (BD Biosciences), das mit ei-

nem Hochdurchsatz-Probennehmer ausgerüstet war, bestimmt. Proben der Zellen wurden bei einer geringen Flussrate genommen, bis mindestens 20.000 Zellen gesammelt wurden. Daten wurden unter Verwendung von FCS Express (De Novo Software) analysiert, indem eine Begrenzung auf eine polygonale Region erfolgte, die 60% Zellpopulation im Vorwärts-Seitwärts-Streuungs-Plot enthielt. Für jedes Experiment wurden Dreifachansätze der Kulturen gemessen, und ihre Standardabweichung wurde als Fehlerbalken angegeben.

B-Galaktosidase-Assay

[0599] Um einen β -Galaktosidase-Assay durchzuführen, wurde 1 μ l der wie oben beschrieben hergestellten Übernachtskultur zu 249 μ l frischen EZ-RDM gegeben, das die gleichen antibiotischen Konzentrationen mit 2 μ M aTc enthielt, mit oder ohne 1 mM Isopropyl- β -D-1-thiogalactopyranosid (IPTG). Die Zellen wurden bis zur mittleren Log-Phase angezüchtet. Die LacZ-Aktivität von 100 μ l dieser Kultur wurde unter Verwendung des Hefe- β -Galaktosidase-Assay-Kits (Pierce) gemäß den Anweisungen gemessen.

Extraktion und Aufreinigung von Gesamt-RNA

[0600] Für jede Probe wurde eine monoklonale Kultur von *E. coli* bei 37°C von einer OD (600 nm) von 0,1 in 500 ml EZ-RDM bis zur frühen Log-Phase (OD $0,45 \pm 0,05$) angezüchtet, und an diesem Punkt wurden die Zellen durch Filtration über einen 0,22 μ m-Nitrozellulose-Filter (GE) geerntet und in flüssigem Stickstoff eingefroren, um gleichzeitig alle Transkriptions-Fortschritte anzuhalten. Die gefrorenen Zellen (100 μ g) wurden auf einer Qiagen TissueLyser II-Mixer-Mühle 6 mal bei 15 Hz für 3 min in Gegenwart von 500 μ l gefrorenem Lysepuffer (20 mM Tris pH 8, 0,4% Triton X-100, 0,1% NP-40, 100 mM NH_4Cl , 50 U/ml SUPERase-In (Ambion) und 1 \times Protease-Inhibitor-Cocktail (Complete, EDTA-frei, Roche), der mit 10 mM MnCl_2 und 15 μ M Tagetin-Transkriptionsinhibitor (Epicentre) supplementiert war) pulverisiert.

[0601] Das Lysat wurde auf Eis durch Pipettieren resuspendiert. RQ1 DNase I (110 U insgesamt, Promega) wurde zugegeben und für 20 min auf Eis inkubiert. Die Reaktion wurde mit EDTA gequench (25 mM Endkonzentration), und das Lysat wurde bei 4°C durch Zentrifugation bei 20.000 g für 10 min geklärt. Das Lysat wurde auf eine PD-MiniTrap G-25-Säule (GE Healthcare) geladen und mit Lysepuffer eluiert, der mit 1 ml EDTA supplementiert war.

Aufreinigung der Gesamt-mRNA

[0602] Gesamt-RNA wurde ausgehend von den geklärten Lysaten unter Verwendung des miRNeasy-Kits (Qiagen) aufgereinigt. 1 μ g RNA in 20 μ l 10 mM Tris pH 7 wurde mit einem gleichen Volumen 2x alkalischer Fragmentierungslösung (2 mM EDTA, 10 mM Na_2CO_3 , 90 mM NaHCO_3 , pH 9.3) gemischt und für etwa 25 min bei 95°C inkubiert, um Fragmente im Bereich von 30–100 nt zu erzeugen. Die Fragmentierungsreaktion wurde durch Zugabe von 0,56 ml eiskalter Präzipitationslösung (300 mM NaOAc pH 5.5 plus GlycoBlue (Ambion)) abgestoppt, und die RNA wurde durch Standard-Isopropanol-Präzipitation gereinigt. Die fragmentierte mRNA wurde anschließend in einer 50 μ l-Reaktion mit 25 U T4 PNK (NEB) in 1 \times PNK-Puffer (ohne ATP) plus 0,5 U SUPERase-In, dephosphoryliert und mit GlycoBlue mittels Standard-Isopropanol-Präzipitationsverfahren präzipitiert.

Aufreinigung der entstehenden RNA

[0603] Für die Aufreinigung entstehender RNA wurde das geklärte Lysat zu 0,5 ml Anti-FLAG M2-Affinitätsgel (Sigma Aldrich) gegeben, wie es zuvor beschrieben wurde. Das Affinitätsgel wurde zweimal mit Lysepuffer gewaschen, der mit 1 ml EDTA supplementiert war, bevor die Inkubation mit den geklärten Lysaten bei 4°C für 2,5 h bei Nutation durchgeführt wurde. Die Immun-Präzipitation wurde mit 4 \times 10 ml Lysepuffer gewaschen, der mit 300 mM KCL supplementiert war, und gebundenes RNAP wurde zweimal mit Lysepuffer eluiert, der mit 1 mM EDTA und 2 mg/ml 3 \times FLAG-Peptid (Sigma Aldrich) supplementiert war. Entstehende RNA wurde aus dem Eluat unter Verwendung des miRNeasy-Kits (Qiagen) gereinigt und in DNA umgewandelt, wobei ein zuvor etabliertes Protokoll zur Erzeugung einer Bibliothek verwendet wurde.

Herstellung der DNA-Bibliothek und DNA-Sequenzierung

[0604] Die DNA-Bibliothek wurde auf einem Illumina HiSeq 2000 sequenziert. Die Auslesungen wurden unter Verwendung des HTSeq Python-Pakets und anderer maßgeschneiderter, in Python geschriebener Software verarbeitet. Das 3'-Ende des sequenzierten Transkripts wurde unter Verwendung von Bowtie ("bowtie-

bio" vorhergehend ".sourceforge.net") in einem Alignment gegenübergestellt, und die RNAP-Profile wurden in MochiView ("johnsonlab.ucsf" vorhergehend ".edu/mochi.html") erzeugt.

Plasmid-Design und Plasmid-Konstruktion für CRISPRi in humanen Zellen

[0605] Die Sequenz, die für ein in Bezug auf Säugetier-Kodons optimiertes Cas9 aus *Streptococcus pyogenes* (DNA 2.0) kodierte, wurde mit drei C-terminalen nuklearen SV40-Ortsbestimmungs-Sequenzen (NLS) fusioniert oder mit tagBFP, das von zwei NLS flankiert wurde. Unter Verwendung einer standardmäßigen Ligations-unabhängigen Klonierung haben wir diese beiden Fusionsproteine in MSCV-Puro (Clontech) kloniert. Guide-sgRNAs wurden unter Verwendung eines lentiviralen auf U6 basierenden Expressionsvektors exprimiert, der von pSico abgeleitet ist und mCherry ausgehend von einem CMV-Promoter co-exprimiert. Die sgRNA-Expressionsplasmide wurden durch Insertierung von zusammengefügten Primern in den Expressionsvektor, der auf einem lentiviralen U6 basiert und mit BstXI und XhoI verdaut worden war, kloniert.

Zellkultur, DNA-Transfektion und Fluoreszenz-Messungen für CRISPRi in humanen Zellen

[0606] HEK293-Zellen wurden in Dulbecco modifiziertem Eagle Medium (DMEM) in 10% FBS, 2 mM Glutamin, 100 Einheiten/ml Streptomycin und 100 µg/ml Penicillin gehalten. HEK293 wurden mit einem GFP-exprimierenden MSCV-Retrovirus unter Verwendung von Standardprotokollen infiziert und mittels Durchflusszytometrie im Hinblick auf eine stabile GFP-Expression sortiert, wobei ein BD FACS Aria2 verwendet wurde. GFP-exprimierende HEK293-Zellen wurden transient unter Verwendung von TransIT-LTI-Transfektionsreagenz (Mirus) transfiziert, wobei das empfohlene Protokoll des Herstellers in Platten mit 24 Vertiefungen unter Verwendung von 0,5 µg des dCas9-Expressionsplasmids und 0,5 µg des RNA-Expressionsplasmids (mit 0,25 µg des GFP-Reporterplasmids für **Fig. 45B**) verwendet wurde. 72 h nach der Transfektion wurden die Zellen mit Trypsin in eine Suspension aus Einzelzellen überführt. Der U6-Vektor enthält einen konstitutiven CMV-Promoter, der ein mCherry-Gen steuert. Die GFP-Expression wurde unter Verwendung eines BD LSR II FACS-Geräts analysiert, indem die mCherry-positiven Populationen (> 10-fach heller bei mCherry gegenüber Zellen der Negativ-Kontrolle) eingegrenzt wurde.

Erstellte RNAs

[0607] Die sgRNA-Designs, die in den Figuren verwendet wurden: lediglich die 20 Nukleotiden der passenden Region (auf DNA gerichtetes Segment) sind aufgeführt (sofern nicht anderweitig angegeben). Die in **Fig. 40C** verwendeten, auf mRFP-gerichteten sgRNAs (SEQ ID NOS: 741–746); Die in **Fig. 40D** verwendeten, auf den Promoter gerichteten sgRNAs (SEQ ID NOS: 747–751); Ziel-Promoter-Sequenz in **Fig. 40D** (SEQ ID NO: 752); Die in **Fig. 43B** verwendeten, auf mRFP gerichteten sgRNAs (SEQ ID NOS: 753–760); Die in **Fig. 42B** verwendete, auf sfGFP gerichtete sgRNA (gfp) (SEQ ID NO: 761); Die in **Fig. 43B** verwendeten, auf sfGFP gerichteten sgRNAs (SEQ ID NOS: 762–769); Die Doppel-sgRNA-Targeting-Experimente in **Fig. 43F** und **Fig. 51** (SEQ ID NOS: 770–778); Die in **Fig. 44B** verwendeten, auf das lac-Operon gerichteten sgRNAs (SEQ ID NOS: 779–787); und die in **Fig. 45** verwendeten auf EGFP gerichteten sgRNAs (SEQ ID NOS: 788–794).

Tabelle 3: Sequenzen die in den Figuren von Beispiel 4 verwendet wurden (oben aufgeführt)

Sequenz	SEQ ID NO:	Sequenz	SEQ ID NO:	Sequenz	SEQ ID NO:
T1	741	R2	771	crp	783
T2	742	R3	772	cya	784
T3	743	R4	773	A-Stelle	785
NT1	744	R5	774	O-Stelle	786
NT2	745	R6	775	P-Stelle	787
NT3	746	R7	776	eT1	788
P1	747	R8	777	eT2	789
P2	748	R9	778	eNT1	790
P3	749	lacZ	779	eNT2	791

P4	750	lacI	780	eNT3	792
P5	751	lacY	781	eNT4	793
R1	770	lacA	782	eNT5	794

ERGEBNISSE

[0608] Das CRISPR(clustered regulatory interspaced short palindromic repeats)-System stellt eine neue potentielle Plattform für die gezielte Gen-Regulation dar. Etwa 40% der Bakterien und 90% der Archaeen besitzen CRISPR/CRISPR-assoziierte(Cas)-Systeme, um Resistenz gegen fremde DNA-Elemente zu erhalten. CRISPR-Systeme verwenden kleine RNAs, die Basenpaarungen ausbilden, um fremde DNA-Elemente in Sequenz-spezifischer Weise aufzuspüren und zu spalten. Es gibt unterschiedliche CRISPR-Systeme in verschiedenen Organismen und eines der einfachsten ist das CRISPR-System vom Typ II aus *Streptococcus pyogenes*: lediglich ein einzelnes Gen, das für das Cas9-Protein kodiert und zwei RNAs, eine reife CRISPR-RNA (crRNA) und eine partiell komplementäre trans-wirkende RNA (tracrRNA) sind für das durch RNA geführte Silencing von fremden DNAs erforderlich und ausreichend (**Fig. 46**). Die Reifung von crRNA erfordert tracrRNA und RNase III. Dieses Erfordernis kann jedoch umgangen werden, indem eine künstlich erstellte, kleine Guide-RNA (sgRNA) verwendet wird, die eine künstlich erstellte Haarnadelschleife umfasst, welche den tracrRNA-crRNA-Komplex nachahmt. Eine Basenpaarung zwischen der sgRNA und der Ziel-DNA verursacht aufgrund der Endonuklease-Aktivität von Cas9 Doppelstrang-Brüche (DSBs). Die Bindungsspezifität wird sowohl durch sgRNA-DNA-Basenpaarung als auch durch ein kurzes DNA-Motiv (Protospacer-benachbartes Motiv oder PAM, Sequenz: NGG), welches neben der komplementären DNA-Region angeordnet ist, bestimmt. Somit erfordert das CRISPR-System lediglich eine minimale Gruppe von zwei Molekülen, dem Cas9-Protein und der sgRNA, und besitzt somit das Potential als Wirts-unabhängige Gen-Targeting-Plattform verwendet zu werden. Es wurde gezeigt, dass die Cas9/CRISPR für ein ortsselektives, von RNA geführtes Genom-Editing verwendet werden kann (**Fig. 39A**).

[0609] **Fig. 46** zeigt den Mechanismus des CRISPR-Systems vom Typ II aus *S. pyogenes*. Das System besteht aus einer Gruppe von CRISPR-assoziierten(Cas)-Proteinen und einem CRISPR-Locus, der einen Array von Repeat-Spacer-Sequenzen enthält. Alle Repeats sind gleich und alle Spacer sind unterschiedlich und komplementär zu den Sequenzen der Ziel-DNA. Wenn die Zelle mit fremden DNA-Elementen infiziert wird, wird der CRISPR-Locus zu einem langen Vorläufer-Transkript transkribiert, das in kleinere Fragmente gespalten wird. Die Spaltung wird durch eine trans-wirkende Antisense-RNA (tracrRNA) und die RNase III des Wirts vermittelt. Nach der Spaltung erkennt ein einzelnes Protein, Cas9, die gespaltene Form der crRNA und bindet an diese. Cas9 führt crRNA zur DNA und rastert das DNA-Molekül. Der Komplex wird durch Basenpaarung zwischen der crRNA und dem DNA-Ziel stabilisiert. In diesem Fall verursacht Cas9 aufgrund ihrer Nuklease-Aktivität Brüche in der doppelsträngigen DNA. Dies entfernt normalerweise verwandte DNA-Moleküle, und die Zellen gewinnen Immunität gegen bestimmte DNA-Populationen.

[0610] **Fig. 39** zeigt das Design des CRISPR-Interferenz(CRISPRi)-Systems. (A) Das minimale Interferenz-System besteht aus einem einzelnen Protein und einer künstlich erstellten sgRNA-Chimäre. Die sgRNA-Chimäre besteht aus drei Domänen (eingerahmte Region): einer komplementären Region aus 20 Nukleotiden (nt) für die spezifische DNA-Bindung, eine Haarnadelschleife aus 42 nt für die Cas9-Bindung (Cas9-Handle) und ein Transkription-Terminator aus 40 nt, der von *S. pyogenes* abgeleitet wurde. Das Wildtyp-Cas9-Protein enthält die Nuklease-Aktivität. Das dCas9-Protein ist hinsichtlich der Nuklease-Aktivität defekt. (B) Das Wildtyp-Cas9-Protein bindet an die sgRNA und bildet einen Protein-RNA-Komplex. Der Komplex bindet durch Watson-Crick-Basenpaarung an spezifische DNA-Ziele zwischen der sgRNA und dem DNA-Ziel. Im Fall von Wildtyp-Cas9 wird die DNA aufgrund der Nuklease-Aktivität des Cas9-Proteins gespalten. Im Falle des Nuklease-defizienten Cas9 unterbricht der Komplex eine geeignete Transkription. Ein minimales CRISPRi-System besteht aus einem einzelnen Protein und RNA und kann in wirksamer Weise die Initiation und Elongation der Transkription abschwächen

[0611] Um eine solche CRISPRi-Plattform in *E. coli* zu implementieren, wurden das Wildtyp-Cas9-Gen von *S. pyogenes* und eine sgRNA ausgehend von bakteriellen Vektoren exprimiert, um zu bestimmen, ob das System die Gen-Expression an einem ausgewählten Ort stören kann (**Fig. 40A**). Das CRISPR-System von *S. pyogenes* ist orthogonal zu dem nativen System aus *E. coli*. Das Cas9-Protein wird von einem Anhydrotetracyclin(aTc)-induzierbaren Promoter auf einem Plasmid exprimiert, das einen p15A-Replikationsursprung aufweist, und die sgRNA wird von einem minimalen konstitutiven Promoter auf einem Plasmid exprimiert, das einen ColE1-Replikationsursprung aufweist. Als alternative Strategie wurde eine katalytisch tote Cas9-Mutante (dCas9)

verwendet, die in Bezug auf die DNA-Spaltung defekt ist, und es konnte gezeigt werden, dass diese Form des Cas9 weiterhin als ein einfacher, von RNA geführter DNA-bindender Komplex agiert.

[0612] Fig. 40 zeigt, dass CRISPRi in wirksamer Weise die Elongation und Initiation der Transkription abschwächt. (A) Das CRISPRi-System besteht aus einem induzierbaren Cas9-Protein und einer künstlich erstellten sgRNA-Chimäre. Das dCas9 enthält Mutationen der RuvC1- und HNH-Nuklease-Domänen. Die sgRNA-Chimäre enthält drei funktionale Domänen, wie sie in Fig. 1 beschrieben sind. (B) Sequenz der künstlich erstellten sgRNA (NT1) und dem DNA-Ziel. NT1 ist auf den Nicht-Matrizen-DNA-Strang der für mRFP kodierenden Region gerichtet. Lediglich die Region, die das Basenpaarungs-Motiv (20 nt) umgibt, ist gezeigt. Die Nukleotide der Basenpaarung sind nummeriert, und die dCas9-bindende Haarnadelschleife ist mit Oberstrichen dargestellt. Die PAM-Sequenz ist unterstrichen. (C) CRISPRi blockiert die Transkription-Elongation in Strang-spezifischer Weise. Ein synthetisches, auf Fluoreszenz basierendes Reporter-System, das ein für mRFP kodierendes Gen enthält, wurde in das *E. coli* MG1655 Genom insertiert (den *nsfA*-Locus). 6 sgRNAs, die entweder den Matrizen-DNA-Strang oder den Nicht-Matrizen-DNA-Strang binden, wurden mit dem dCas9-Protein co-exprimiert, wobei deren Wirkungen auf das Ziel-mRFP durch einen in vivo-Fluoreszenz-Assay gemessen wurden. Lediglich sgRNAs, die an den Nicht-Matrizen-DNA-Strang binden, zeigten ein Silencing (10~300-fach). Die Kontrolle zeigte eine Fluoreszenz der Zellen mit dem dCas9-Protein, jedoch ohne die sgRNA. (D) CRISPRi blockiert die Initiation der Transkription. 5 sgRNAs wurden künstlich so erstellt, dass sie an verschiedene Regionen um einen *E. coli*-Promoter binden (J23119). Die Transkriptions-Startstelle wurde als +1 markiert. Das gepunktete Oval zeigt den initialen RNAP-Komplex, der eine 75 bp-Region von -55 bis +20 abdeckt. Lediglich sgRNAs, die auf Regionen innerhalb des initialen RNAP-Komplexes gerichtet waren, zeigten eine Repression (P1-P4). Anders als bei der Blockierung der Elongation der Transkription war das Silencing unabhängig vom DNA-Strang, auf den abgestellt wurde. (E) Die CRISPRi-Regulation war reversibel. Sowohl dCas9 als auch sgRNA (NT1) waren unter der Kontrolle eines aTc-induzierbaren Promoters. Die Zellkultur wurde während der exponentiellen Phase beibehalten. Am Zeitpunkt T = 0 wurden 0,1 µM aTc zu Zellen mit einer OD = 0,001 supplementiert. Die Repression des Ziel-mRFP begann innerhalb von 10 Minuten. Das Fluoreszenz-Signal verschwand auf eine Weise, die mit dem Zellwachstum übereinstimmt, was darauf hinweist, dass das Verschwinden auf eine Zellteilung zurückzuführen war. Innerhalb von 240 min erreichte die Fluoreszenz das vollständig reprimierte Ausmaß. Bei T = 370 min ist das aTc aus dem Wachstumsmedium ausgewaschen, und die Zellen wurden auf eine OD = 0,001 zurück verdünnt. Die Fluoreszenz begann innerhalb von 50 min zu steigen, und es brauchte etwa 300 min, um auf dasselbe Ausmaß wie bei der Positiv-Kontrolle zu gelangen. Positiv-Kontrolle: immer ohne den Inducer; Negativ-Kontrolle: immer mit 1 µM aTc-Inducer. Die Fluoreszenz-Ergebnisse in 2C, 2D und 2E stellen den Durchschnitt und den Standardfehler (SEM) von mindestens drei biologischen Replikaten dar. Siehe auch Fig. 47 und Fig. 48.

[0613] Die mit Cas9 co-exprimierten sgRNA-Moleküle bestanden jeweils aus drei Segmenten: einer Ziel-spezifischen komplementären Region aus 20 Nukleotiden (nt), einer Cas9-bindenden Haarnadelschleife (Cas9-Handle) aus 42 nt und einem Transkriptions-Terminator aus 40 nt, der von *S. pyogenes* abgeleitet ist (Fig. 40B). Ein Reporter-System, das auf einem rot fluoreszierenden Protein (mRFP) basiert, wurde in das Genom von *E. coli* MG1655 insertiert.

[0614] Die Co-Expression des Wildtyp-Cas9-Proteins und einer sgRNA (NT), die auf die mRFP-kodierende Sequenz gerichtet ist, verringerte die Transformations-Effizienz dramatisch, vermutlich aufgrund von durch Cas9 induzierten Doppelstrangbrüchen im Genom (Fig. 47A). Die Sequenzierung einiger weniger überlebender Kolonien zeigte, dass sie alle Sequenz-Umlagerungen im Bereich der mRFP-Zielstelle aufwiesen, was darauf verweist, dass es eine starke Selektion gegen die Expression von Wildtyp-Cas9 und eines auf eine Sequenz des Wirts gerichtete sgRNA gab. Das dCas9-Mutanten-Gen (nicht spaltend), welches zwei Silencing-Mutationen in der RuvC1- und der HNH-Nuklease-Domäne aufwies (D10A und H841A), schwächte diese Letalität ab, wie durch die Transformations-Effizienz und die Wachstumsraten von *E. coli* verifiziert wurde (Fig. 47A und B).

[0615] Fig. 47 steht mit Fig. 40 im Zusammenhang und zeigt die Wachstumsraten von *E. coli*-Zellkulturen, die mit dCas9 und sgRNA co-transformiert waren (A) Transformations-Effizienz für die Transformation von *E. coli*-Zellen mit 2 Plasmiden. Ein Plasmid enthält eine sgRNA, die auf eine genomische Kopie von mRFP gerichtet ist, und das andere Plasmid enthält Wildtyp-Cas9 oder dCas9. Die Co-Transformation von Wildtyp-Cas9 und sgRNA ist in hohem Maße toxisch, was durch Verwendung von dCas9 abgeschwächt werden kann. (B) Die sgRNA (NT1) ist so erstellt, dass sie auf die kodierende Sequenz von mRFP gerichtet ist. Die Co-Expression von dCas9 und sgRNA weist nahezu keine Wirkungen auf zelluläre Wachstumsraten auf, was darauf verweist, dass die dCas9-sgRNA-Interaktion mit DNA stark genug ist, um die RNA-Polymerase zu blockieren, jedoch nicht die DNA-Polymerase oder die Zell-Replikation. Die Ergebnisse stellen den Durchschnitt und den Standardfehler (SEM) von mindestens 3 unabhängigen Experimenten dar.

[0616] Um zu überprüfen, ob der dCas9:sgRNA-Komplex eine hocheffiziente Repression der Gen-Expression ausüben könnte, wurden sgRNAs künstlich hergestellt, die komplementär zu unterschiedlichen Regionen der für mRFP-kodierenden Sequenz waren, wobei diese entweder an den Matrizen-DNA-Strang banden oder an den Nicht-Matrizen-DNA-Strang. Die Ergebnisse zeigten, dass sgRNAs, die auf den Nicht-Matrizen-DNA-Strang gerichtet waren, ein wirksames Gen-Silencing zeigten (10 bis 300-fache Repression), während solche, die auf den Matrizen-Strang gerichtet waren, nur wenig Wirkung zeigten (**Fig. 40C**). Das System zeigte ähnliche Regressions-Wirkungen bei Genen, die innerhalb des *E. coli*-Genoms oder auf einem Plasmid mit hoher Kopienzahl lagen (**Fig. 48**). Darüber hinaus führte das Targeting der Promoter-Region ebenfalls zu einem wirksamen Gen-Silencing (**Fig. 40D**). Das Targeting der –35-Box durch die sgRNAs verringerte die Gen-Expression signifikant (P1, etwa 100-fache Repression), während das Targeting von anderen benachbarten Regionen eine gedämpfte Wirkung zeigte (P2–P4). Targeting-Sequenzen etwa 100 bp stromaufwärts des Promoters zeigten keine Wirkung (P5). Anders als beim Targeting der kodierenden Region ist beim Targeting des Promoters die Wirkung des Silencings unabhängig vom DNA-Strang; das Targeting von Matrizen- oder Nicht-Matrizen-Strang ist gleichermaßen wirksam (P2 und P3).

[0617] **Fig. 48** steht im Zusammenhang mit **Fig. 40C** und zeigt, dass CRISPRi die Expression eines Reporter-Gens auf einem Plasmid mit mehreren Kopien abschwächen konnte. Das mRFP-Gen wurde auf ein p15A-Plasmid kloniert. Das Vorhandensein der dCas9 und einer für mRFP-spezifischen sgRNA (NT1) reprimierte mRFP stark (etwa 300-fach). Die Repressionswirkung ist ähnlich zu der, die bei Verwendung des mRFP im Genom beobachtet wurde (**Fig. 40C**). Das Silencing ist nur wirksam, wenn die sgRNA auf den Nicht-Matrizen-DNA-Strang einwirkt, jedoch nicht, wenn sie auf den Matrizen-DNA-Strang einwirkt (T1). Darüber hinaus ist das Silencing hochgradig spezifisch, da eine für GFP spezifische 3 sgRNA (gfp) keine Wirkung auf die mRFP-Expression zeigte. Die Fluoreszenz-Ergebnisse stellen den Durchschnitt und den Standardfehler (SEM) von mindestens drei biologischen Replikaten dar.

Der CRISPRi-Gen-Knock-Down ist induzierbar und reversibel

[0618] Anders als bei Verfahren zum Gen-Knock-Out besteht ein Vorteil bei der Verwendung eines auf CRISPRi basierenden Knock-Downs der Gen-Expression in der Tatsache, dass diese Störung reversibel sein sollte. Um zu überprüfen, ob eine CRISPRi-Regulation induziert und anschließend umgekehrt werden kann, wurden sowohl dCas9 als auch eine mRFP-spezifische sgRNA (NT1) unter Kontrolle des aTc-induzierbaren Promoters gestellt, und es wurden Zeit-Verlaufs-Messungen der CRISPRi-vermittelten Regulation von mRFP in Reaktion auf Inducer durchgeführt (**Fig. 40E**). Am Zeitpunkt 0 wurden Zellkulturen, die bis zur frühen exponentiellen Phase ohne Inducer wuchsen, mit 1 µM aTc supplementiert. Die Daten verweisen darauf, dass das System schnell auf das Vorhandensein eines Inducers reagieren konnte – das fluoreszente Reporter-Protein-Signal begann, innerhalb von 10 min nach Zugabe des Inducer-Moleküls abzunehmen. Da das mRFP-Protein stabil ist, ist die Geschwindigkeit bei der Verringerung des Fluoreszenz-Signals durch die Verdünnung des Proteins aufgrund des Zellwachstums begrenzt, wie durch ähnliche Zeiten für Zellverdopplung und Verlust der Fluoreszenz-Halbwertszeit (beides etwa 36 min) zu erkennen ist. Bei 240 min wurden alle Zellen einheitlich im selben Ausmaß reprimiert wie die Negativ-Kontrolle. Bei 420 min war der Inducer aus dem Wachstumsmedium herausgewaschen, und die Zellen wurden auf eine geringere OD zurück verdünnt. Nach einer Verzögerung von 50 min begann die mRFP-Fluoreszenz zuzunehmen. Es benötigte insgesamt 300 min, bis die Fluoreszenz einer einzelnen Zelle auf das gleiche Ausmaß zunahm, wie die Positiv-Kontrolle. Die Verzögerung von 50 min wird höchstwahrscheinlich durch die Verschiebung der dCas9/sgRNA-Umsatzgeschwindigkeit aufgrund von Verdünnung durch Zellwachstum und Teilung bestimmt. Insgesamt zeigen diese Ergebnisse, dass die Wirkungen des Silencing von dCas9-sgRNA induziert und umgekehrt werden können.

Die Sequenzierung des nativen, elongierten Transkripts (NET-Seq)
bestätigt, dass CRISPRi durch Blockierung der Transkription funktioniert

[0619] dCas9 schien als DNA-bindender Komplex zu funktionieren, der durch RNA geführt wird und die Bindung der RNA-Polymerase (RNAP) während der Elongation der Transkription blockiert. Da der Nicht-Matrizen-DNA-Strang die gleiche Identität aufweist, wie die transkribierte mRNA und lediglich sgRNAs ein Silencing zeigten, die an den Nicht-Matrizen-DNA-Strang banden, blieb die Möglichkeit, dass der dCas9-sgRNA-Komplex mit mRNA interagiert und deren Translation oder Stabilität verändert. Um zwischen diesen Möglichkeiten zu unterscheiden, wurde ein kürzlich beschriebener Ansatz für die Sequenzierung eines nativen elongierten Transkripts (NET-Seq) in *E. coli* angewendet, der verwendet werden konnte, um die Position von elongierenden RNA-Polymerasen global zu erfassen und die Wirkung des dCas9:sgRNA-Komplexes auf die Transkription zu verfolgen. In diesem NET-Seq-Verfahren wurde das CRISPRi-System in einen von *E. coli* MG1655-abgeleiteten Stamm transformiert, der eine mit FLAG versehene RNAP enthielt. Das CRISPRi enthielt eine

sgRNA (NT1), die an die für mRFP kodierende Region bindet. Eine in vitro-Immunaufreinigung der mit Tag versehenen RNAP, gefolgt von der Sequenzierung der entstehenden Transkripten, die mit der Elongation der RNAP assoziiert waren, erlaubte die Unterscheidung der Stellen des Anhaltens der RNAP.

[0620] Diese Experimente zeigten, dass die sgRNA ein starkes Anhalten der Transkription stromaufwärts des sgRNA-Ziel-Locus induzierte (**Fig. 41A**). Der Abstand zwischen der Stelle des Anhaltens und der Ziel-Stelle beträgt 19 bp, was in perfektem Einklang mit dem zuvor berichteten Abstand von etwa 18 bp zwischen dem Nukleotid-Einbau der RNAP und ihrer Stirnseite steht. Dieser Befund stimmt mit einem Mechanismus von CRISPRi überein, bei dem die Blockierung der Transkription durch eine physikalische Kollision zwischen der elongierenden RNAP und dem dCas9:sgRNA-Komplex erfolgt (**Fig. 41B**). Die Bindung des dCas9:sgRNA-Komplexes an den Matrizen-Strang hat geringe repressive Wirkung, was nahelegt, dass die RNAP in der Lage war, in dieser bestimmten Orientierung durch den Komplex hindurch zu lesen. In diesem Fall ist die sgRNA der RNAP zugewandt, welche durch die Helicase-Aktivität der RNAP entfaltet werden könnte. Diese Experimente haben gezeigt, dass CRISPRi RNAs verwendet, um die Transkription direkt zu blockieren. Dieser Mechanismus unterscheidet sich von dem bei RNAi, bei dem der Knock-Down der Gen-Expression die Zerstörung von bereits transkribierten Messenger-RNAs vor deren Translation erfordert.

[0621] **Fig. 41** zeigt, dass CRISPRi durch Blockierung der Elongation der Transkription funktioniert. (A) Mit FLAG als Tag versehene RNAP-Moleküle wurden immunpräzipitiert, und die mit ihr assoziierten entstehenden mRNA-Transkripte wurden sequenziert. Die obere Teilabbildung zeigt die Ergebnisse der Sequenzierung der entstehenden mRFP-Transkripten in Zellen ohne sgRNA, und die untere Teilabbildung zeigt die Ergebnisse in Zellen mit sgRNA. In Gegenwart von sgRNA wurde ein starkes Anhalten der Transkription 19 bp stromaufwärts der Ziel-Stelle beobachtet, nach der die Anzahl der Sequenzierungs-Auslesungen stark abfällt. (B) Ein vorgeschlagener CRISPRi-Mechanismus basiert auf der physikalischen Kollision zwischen RNAP und dCas9-sgRNA. Der Abstand vom Zentrum der RNAP zu ihrem vorderen Rand beträgt etwa 19 bp, was gut mit der gemessenen Distanz zwischen der Stelle des Anhaltens der Transkription und der 3'-sgRNA-Basenpaarungs-region übereinstimmt. Die angehaltene RNAP bricht die Elongation der Transkription bei Begegnung mit dem dCas9-sgRNA-Hindernis ab. Das durch die CRISPRi-sgRNA geleitete Gen-Silencing ist in hohem Maße spezifisch.

[0622] Um die Spezifität von CRISPRi im Genom-weiten Maßstab zu bewerten, wurde eine Gesamt-Transkriptom-Shotgun-Sequenzierung (RNA-seq) von dCas9-transformierten Zellen mit und ohne Co-Expression von sgRNA durchgeführt (**Fig. 42A**). In Gegenwart der sgRNA, die auf mRFP (NT1) gerichtet ist, war das mRFP-Transkript das einzige Gen, welches eine Verringerung der Häufigkeit zeigte. Kein anderes Gen zeigte im Rahmen der Sequenzierungsfehler eine signifikante Änderung der Expression bei Zugabe der sgRNA. Wir haben ferner bei Zellen eine RNA-seq mit unterschiedlichen sgRNAs, die auf unterschiedliche Gene gerichtet sind, durchgeführt. Keines dieser Experimente zeigte signifikante Änderungen von Genen neben den als Ziel ausgewählten Genen (**Fig. 49**). Somit ist das durch sgRNA geführte Gen-Targeting und die Regulation hochgradig spezifisch und weist keine signifikanten unspezifischen Wirkungen auf.

[0623] **Fig. 42** zeigt die Target-Spezifität des CRISPRi-Systems. (A) Eine Genom-weite mRNA-Sequenzierung (RNA-seq) bestätigte, dass das CRISPRi-Targeting keine unspezifischen Wirkungen aufweist. Die sgRNA NT1, die an die für mRFP-kodierende Region bindet, wurde verwendet. Die Gene dCas9, mRFP und sfGFP sind hervorgehoben. (B) Mehrere sgRNAs können unabhängig voneinander bei zwei fluoreszenten Protein-Reportern in derselben Zelle ein Silencing bewirken. Jede sgRNA reprimierte spezifisch ihr verwandtes Gen, jedoch nicht das andere Gen. Wenn beide sgRNAs vorhanden waren, wurde bei beiden Genen ein Silencing ausgeübt. Fehlerbalken stellen den Standardfehler (SEM) von mindestens drei biologischen Replikaten dar. (C) Mikroskopische Abbildungen bei Verwendung von zwei sgRNAs zur Kontrolle von zwei fluoreszenten Proteinen. Die obere Teilabbildung zeigt Breitfeld-Bilder von E. coli-Zellen, die mittlere Teilabbildung zeigt den RFP-Kanal und unten ist das GFP-Bild dargestellt. Die Co-Expression von einer sgRNA und dCas9 führt nur bei dem verwandten fluoreszenten Protein zum Silencing, nicht jedoch bei dem anderen. Die Knock-Down-Wirkung war stark, da fast keine Fluoreszenz von Zellen, bei denen ein bestimmtes fluoreszentes Protein dem Silencing unterworfen war, beobachtet wurde. Maßstabsbalken, 10 µm. Die Kontrolle zeigte Zellen ohne fluoreszente Protein-Reporter. Die Fluoreszenz-Ergebnisse stellen Durchschnittswerte und Standardfehler (SEM) von mindestens 3 biologischen Replikaten dar (siehe auch **Fig. 49**).

[0624] **Fig. 49** steht im Zusammenhang mit **Fig. 42A** und zeigt die RNA-seq-Daten von Zellen mit sgRNAs, die auf unterschiedliche Gene gerichtet sind. (A) (+/-) sgRNA, die auf den Promoter des endogenen lacI-Gens in E. coli gerichtet ist. Es wurde die gleiche auf lacI gerichtete sgRNA wie in **Fig. 44A** verwendet. (B) (+/-) 1 mM IPTG für Zellen ohne auto-inhibierte sgRNA (sgRNA reprimierte ihren eigenen Promoter). (C) (+/-) sgRNA,

die auf das endogene lacZ-Gen in *E. coli* gerichtet ist. Es wurde dieselbe auf lacZ gerichtete sgRNA wie in **Fig. 44A** verwendet. 1 mM IPTG wurde auch mit der auf lacZ gerichteten sgRNA zu Zellen gegeben.

CRISPRi kann verwendet werden, um gleichzeitig mehrere Genen zu regulieren

[0625] Das CRISPRi-System kann die Kontrolle von mehreren Genen unabhängig voneinander ohne Überschneidung ermöglichen. Es wurde ein Zwei-Farben-Fluoreszenzreporter-System verwendet, das auf mRFP und sfGFP basiert. Zwei sgRNAs mit unterschiedlichen komplementären Regionen zu jedem Gen wurden erstellt. Die Expression jeder sgRNA führte lediglich bei dem verwandten Gen zu einem Silencing und hatte keine Wirkung auf das andere Gen. Die Co-Expression von zwei sgRNAs führte zum Knock-Down beider Gene (**Fig. 42B** und **Fig. C**). Diese Ergebnisse legen nahe, dass das durch sgRNA geleitete Targeting spezifisch ist, wobei die Spezifität durch die Sequenz-Identität bestimmt und durch das Vorhandensein anderer sgRNAs nicht beeinflusst wird. Dieses Verhalten sollte eine Multiplex-Kontrolle von mehreren Genen gleichzeitig durch CRISPRi ermöglichen.

Faktoren, die die Effizienz des CRISPRi-Silencings bestimmen

[0626] Um Faktoren für die Effizienz des CRISPRi-Targetings zu finden, wurde die Rolle der Länge, der Sequenz-Komplementarität und der Position auf die Effizienz des Silencings untersucht (**Fig. 43A**). Wie in **Fig. 40C** nahegelegt, war die Lokalisierung der sgRNA-Ziel-Sequenz entlang des Gens für die Effizienz wichtig. Es wurden ferner sgRNAs so erstellt, dass sie die volle Länge der kodierenden Regionen sowohl für mRFP als auch für sfGFP abdeckten (ergänzende Daten für sgRNA-Sequenzen). In allen Fällen war die Repression invers mit dem Abstand des Ziels von der Transkriptions-Startstelle korreliert (**Fig. 43B**). Eine starke lineare Korrelation wurde für mRFP beobachtet. Eine ähnliche, jedoch geringfügig schwächere Korrelation wurde beobachtet, wenn sfGFP als Ziel verwendet wurde, was vielleicht auf abweichende kinetische Eigenschaften der RNA-Polymerase während verschiedener Punkte der Elongation dieses Gens verweist.

[0627] Die sgRNA enthält eine Region von 20 bp, die komplementär zu dem Ziel ist. Um die Bedeutung dieser Basenpaarungs-Region zu identifizieren, wurde die Länge der sgRNA NT1 verändert (**Fig. 43C**). Während die Verlängerung der Region vom 5' Ende das Silencing nicht beeinflusste, verringerte die Trunkierung der Region die Repression erheblich. Die minimale Länge der Basenpaarungs-Region, die für das Gen-Silencing erforderlich war, betrug 12 bp, wobei eine weitere Trunkierung zum vollständigen Verlust der Funktion führte. Einzelne Mutationen wurden in die Basenpaarungs-Region von sgRNA NT1 eingebracht, und die Gesamtwirkung auf das Silencing wurde untersucht. Anhand dieser Ergebnisse konnten drei Sub-Regionen unterschieden werden, jede mit einem unterschiedlichen Beitrag zur Gesamt-Bindung und zum Silencing (**Fig. 43D**). Jede einzelne Mutation der ersten sieben Nukleotide verringerte die Repression erheblich, was darauf hinweist, dass diese Sequenz eine "Seed-Region" für die Bindung darstellt, wie zuvor für CRISPR-Systeme sowohl vom Typ I als auch vom Typ II angemerkt wurde. Angrenzende Nukleotide wurden ebenfalls in Paaren mutiert (**Fig. 43E** und **Fig. 50**). In dem meisten Fällen war die relative Regressions-Aktivität aufgrund einer doppelten Mutation multiplikativ im Vergleich zu den Wirkungen der einzelnen Mutanten, was auf einen unabhängigen Zusammenhang zwischen den Fehlpaarungen hindeutet. Darüber hinaus beseitigte ein inkorrektes PAM das Silencing selbst bei einer perfekten Bindungsregion aus 20 bp (**Fig. 43E**), was in Übereinstimmung mit vorherigen Ergebnissen zur Bedeutung der PAM-Sequenz steht. Somit wird die Spezifität des CRISPRi-Systems gemeinsam durch das PAM (2 bp) und mindestens einem 12 bp-sgRNA-DNA-Abschnitt bestimmt, dessen Raum groß genug ist, um die meisten bakteriellen Genome im Hinblick auf einzigartige Ziel-Stellen abzudecken,.

[0628] Es wurden zwei sgRNAs untersucht, die beide auf das gleiche Gen gerichtet waren (**Fig. 43F** und **Fig. 51**). Abhängig von der relativen Positionierung mehrerer sgRNAs wurden unterschiedliche kombinatorische Wirkungen beobachtet. Die Kombination von zwei sgRNAs, von denen jede eine etwa 300-fache Repression verursachte, erlaubte ein erhöhtes Gesamt-Silencing bis etwa 1000-fach. Die Kombination von zwei schwächeren sgRNAs (etwa 5-fach) zeigte bei gemeinsamer Verwendung multiplikative Wirkungen. Suppressive kombinatorische Wirkungen wurden beobachtet, wenn zwei sgRNAs verwendet wurden, dessen Ziele überlappen. Dies lag vermutlich an der Konkurrenz beider sgRNAs um die Bindung an die gleiche Region.

[0629] **Fig. 43** zeigt die Charakterisierung von Faktoren, die die Effizienz des Silencings beeinflussen. (A) Die Wirkungen des Silencing wurden anhand von sgRNAs mit unterschiedlichen Targeting-Loci auf demselben Gen (Abstand vom Translations-Start-Kodon) und anhand von sgRNAs mit unterschiedlichen Längen der Basenpaarungs-Region zum selben Ziel-Lokus (basierend auf NT1) gemessen. (B) Die Effizienz des Silencings war invers mit dem Ziel-Abstand vom Translations-Start-Kodons korreliert (orange-MRFP und grün-SFGFP). Die relative Repressionsaktivität wurde durch Normalisierung der Repression jeder sgRNA zu der sgRNA mit

der höchsten Repressionsänderung normalisiert. Die Fehlerbalken stellen den Standardfehler (SEM) von 3 biologischen Replikaten dar. (C) Die Länge der Watson-Crick-Basenpaarungs-Region zwischen der sgRNA und der Ziel-DNA beeinflusst die Effizienz der Repression. Verlängerungen der Basenpaarungs-Region zeigten jeweils eine starke Silencing-Wirkung, und Trunkierungen verringerten die Repression erheblich. Die minimale Länge der Basenpaarungs-Region für eine nachweisbare Repression beträgt 12 bp. Die Fehlerbalken stellen den Standardfehler (SEM) von 3 biologischen Replikaten dar. (D) Einzelne Fehlpaarungen wurden in jedes Nukleotid der sgRNA eingebracht (NT1, **Fig. 40B**), und es wurde gemessen, wie diese einzelnen Fehlpaarungen die Regressions-Effizienz beeinflussten. Drei Subregionen von unterschiedlicher Bedeutung auf das gesamte Silencing konnten unterschieden werden. Diese zeigen eine Stufenfunktion. Die Region der ersten 7 Nukleotide war entscheidend für das Silencing und stellt vermutlich eine "Seed"-Region für das Sondieren der Bindung der sgRNAs an das DNA-Ziel dar. Die PAM-Sequenz (NGG) war für das Silencing unerlässlich. Die Fehlerbalken stellen den Standardfehler (SEM) von 3 biologischen Replikaten dar. (E) Die Wirkungen von sgRNAs mit benachbarten Doppel-Fehlpaarungen auf das Silencing. Die relative Repressionsaktivität von einzeln fehlgepaarten sgRNAs ist gezeigt, wobei die Position der Fehlpaarung unten markiert ist. Die experimentell bestimmte Aktivität von doppelt fehlgepaarten sgRNAs ist gezeigt. Die berechnete Aktivität, die durch Multiplikation der Wirkungen von zwei einzelnen fehlgepaarten sgRNAs erhalten wurde, ist in Weiß dargestellt und als "Com" bezeichnet. In den meisten Fällen war die Silencing-Aktivität einer doppelt fehlgepaarten sgRNA einfach eine Multiplikation der Aktivitäten der einzeln fehlgepaarten sgRNAs (außer **Fig. 50B**), was auf einen unabhängigen Zusammenhang zwischen den einzelnen Fehlpaarungen hindeutet. Die Fehlerbalken stellen den Standardfehler (SEM) aus 3 biologischen Replikaten dar. (F) Kombinatorische Silencing-Wirkungen bei Verwendung von doppelten sgRNAs beim Targeting eines einzelnen mRFP-Gens. Bei Verwendung zweier sgRNAs, die auf dasselbe Gen gerichtet sind, kann die Gesamt-Knockdown-Wirkung auf annähernd 1000-fach verbessert werden. Wenn zwei sgRNAs an nicht überlappenden Sequenzen desselben Gens binden, wurde die Repression verstärkt. Wenn zwei sgRNAs auf überlappende Regionen gerichtet sind, wurde die Repression supprimiert. Die Fehlerbalken stellen den Standardfehler (SEM) aus 3 biologischen Replikaten dar.

[0630] **Fig. 50** bezieht sich auf **Fig. 43E** und stellt die Wirkungen des Silencings von sgRNAs mit benachbarten, doppelten Fehlpaarungen dar. Die relative Regressions-Aktivität von einzeln fehlgepaarten sgRNAs ist gezeigt, wobei die Position der Fehlpaarung unten markiert ist. Die experimentell gemessene Aktivität von doppelt fehlgepaarten sgRNAs ist ebenfalls gezeigt. Die Aktivität, die durch Multiplikation der Wirkungen von zwei einzeln fehlgepaarten sgRNAs errechnet wurde, ist in Weiß gezeigt und mit "Com" markiert. Die Fluoreszenz-Ergebnisse stellen den Durchschnitt und den Standardfehler (SEM) aus 3 biologischen Replikaten dar.

[0631] **Fig. 51** bezieht sich auf **Fig. 43F** und stellt die kombinatorischen Silencing-Wirkungen bei Verwendung von zwei sgRNAs zur Regulierung eines einzelnen Gens dar. In allen Fällen zeigten nicht überlappende sgRNAs verstärkende Silencing-Wirkungen, und überlappende sgRNAs zeigten suppressive Wirkungen. Die kombinatorische Wirkung war unabhängig davon, ob die sgRNA auf den Matrizen- oder auf den Nicht-Matrizen-Strang gerichtet war. Die Fluoreszenzergebnisse stellen den Durchschnitt und den Standardfehler (SEM) aus 3 biologischen Replikaten dar.

Abfragen eines endogenen regulatorischen Netzwerks unter Verwendung von CRISPRi-Gen-Knock-Down

[0632] Das CRISPRi-System wurde anschließend als Plattform für ein Gen-Knock-Down verwendet, um endogene Gen-Netzwerke abzufragen. Vorherige Verfahren zum Abfragen mikrobieller Gen-Netzwerke basierten meist auf aufwendigen und kostenintensiven genomischen Konstruktions- und Knock-Out-Methoden. Im Gegensatz dazu erfordert der Gen-Knock-Down mit CRISPRi lediglich die Erstellung und Synthese einer kleinen sgRNA, die eine komplementäre Region von 20 bp zu den gewünschten Genen umfasst. Um dies zu zeigen, wurde CRISPRi verwendet, um Knock-Down-Stämme von *E. coli* zu erzeugen, indem sgRNAs erstellt wurden, um systematisch Gene zu stören, die Teil des gut charakterisierten regulatorischen Wegs von *E. coli* für Laktose sind (**Fig. 44A**). Es wurden β -Galaktosidase-Assays durchgeführt, um die Expression von LacZ in den Knock-Down Stämmen mit und ohne Isopropyl- β -D-1-Thiogalaktopyranosid (IPTG), einer Chemikalie, die den Lac-Repressor (LacI) inhibiert, zu messen. In Wildtyp-Zellen induziert die Zugabe von IPTG die Expression von LacZ. Die Ergebnisse zeigen, dass eine LacZ-spezifische sgRNA, die Expression von LacZ stark reprimieren konnte (**Fig. 44B**). Umgekehrt führte eine sgRNA, die auf das LacI-Gen gerichtet war, zur Aktivierung der LacZ-Expression, selbst in Abwesenheit von IPTG, wie es bei einem Silencing eines direkten Repressors der LacZ-Expression zu erwarten wäre.

[0633] Es ist bekannt, dass cAMP-CRP ein wichtiger Aktivator der LacZ-Expression ist, indem es an eine cis-regulatorische Stelle stromaufwärts des Promotors bindet (A-Stelle). Im Einklang damit führte die sgRNA, die auf das crp-Gen oder auf die A-Stelle im LacZ-Promotor gerichtet war, zur Repression, was ein Mittel zeigt, um

einen Regulator unter Verwendung von CRISPRi-Experimenten an seine cis-regulatorische Sequenz zu koppeln. Das Targeting des Gens für die Adenylatcyclase (*cya*), das erforderlich ist, um das cAMP zu erzeugen, welches CRP am LacZ-Promoter wirksamer macht, führte lediglich zur partiellen Repression. Die Zugabe von 1 mM cAMP zum Wachstumsmedium komplementierte die Wirkungen eines *cya*-Knock-Downs, jedoch nicht die eines *crp*-Knock-Downs, was darauf verweist, dass *cya* ein indirekter Regulator von LacZ ist. Darüber hinaus führte das Targeting der cis-regulatorischen Stelle von LacI (O-Stelle) mit einer sgRNA zu einer Inhibierung, vermutlich weil die Bindung des Cas9-Komplexes an dieser Stelle die RNA-Polymerase sterisch blockiert, was das Verhalten des LacI-Transkriptions-Repressors nachstellt. Das Targeting der bekannten RNAP-Bindungsstelle (P-Stelle) blockierte ebenfalls die Repression. Insgesamt zeigen diese Untersuchungen, dass das auf CRISPRi-basierende Gen-Knock-Down-Verfahren eine schnelle und wirksame Möglichkeit zur Abfrage der regulatorischen Funktion (aktivierend oder reprimierend) von Genen und cis-Elementen in einem komplexen regulatorischen Netzwerk darstellt (**Fig. 44C**).

[0634] **Fig. 44** zeigt das funktionale Profiling eines komplexen regulatorischen Netzwerks unter Verwendung von CRISPRi-Knock-Down. (A) sgRNAs wurden erstellt und zum Knock-Down von Genen (*cya*, *crp*, *lacI*, *lacZ*, *lacY*, *lacA*) im lac-Regulationsweg oder zur Blockierung transkriptionaler Operator-Stellen (A/P/O) verwendet. LacI ist ein Repressor des LacZYA-Operons, indem es an eine Transkriptions-Operator-Stelle (O-Stelle) bindet. Das LacZ-Gen kodiert für ein Enzym, das die Umwandlung von Laktose in Glukose katalysiert. Einige trans-wirkende Gene, wie beispielsweise *cya* und *crp*, sind an der Aktivierung des LacZYA-Systems beteiligt. Der cAMP-CRP-Komplex bindet an eine Transkriptions-Operator-Stelle (A-Stelle) und rekrutiert eine RNA-Polymerase-Bindung an die P-Stelle, was die Transkription von LacZYA auslöst. IPTG, eine Chemikalie, welche die LacI-Funktion inhibiert, induziert die LacZ-Expression. (B) β -Galaktosidase-Assays der Knock-Down-Stämme ohne (weiß) und mit (grau) IPTG. Die Kontrolle zeigt, dass die Wildtyp-Zellen ohne CRISPRi-Störung durch Zugabe von IPTG induziert werden konnten. Die sgRNA, die auf LacZ gerichtet ist, reprimierte die LacZ-Expression stark, selbst in Gegenwart von IPTG. Wenn ein Targeting auf LacI erfolgte, war die LacZ-Expression hoch, selbst ohne IPTG. Bei einem Targeting der Gene *cya* und *crp* kam es zu einem verringerten LacZ-Expressionsniveau in Gegenwart von IPTG. Das Vorhandensein von 1 mM cAMP hob den *cya*-Knock-Down auf, nicht jedoch den *crp*-Knock-Down. Die Blockierung der Transkriptions-Operator-Stellen führte zur Repression von LacZ, was darauf verweist, dass diese wichtige cis-wirkende regulatorische Stellen für LacZ sind. Die verringerte Expression (Pfeile nach unten) und die erhöhte Expression (Pfeile nach oben) bei Störung von LacZ ist angegeben. Die Fehlerbalken stellen den Standardfehler (SEM) aus 3 biologischen Replikaten dar. (C) Die Knock-Down-Experimente ermöglichten es uns, die Rollen der Regulatoren im regulatorischen lac-Kreislauf zu untersuchen. Die Daten sind auf einer 2D-Grafik gezeigt, wobei die X-Achse die LacZ Aktivität ohne IPTG zeigt, und die Y-Achse die Aktivität mit IPTG. Die Verteilung der Ovale entlang jeder Achse zeigt die Standardabweichungen. Die Ergebnisse des β -Galaktosidase-Assays stellen den Durchschnittswert und den Standardfehler (SEM) von 3 biologischen Replikaten dar. Siehe auch **Fig. 49** für RNA-seq-Daten zum LacI- und LacZ-Targeting.

CRISPRi kann die Expression eines Gens, auf das abgestellt wird, in humanen Zellen abschalten

[0635] Um die allgemeine Gültigkeit des CRISPRi-Ansatzes bei der Verwendung des dCas9-sgRNA-Komplexes zur Repression der Transkription zu untersuchen, wurde das System in HEK293-Säugetierzellen getestet. Das dCas9-Protein wurde Kodon-optimiert, an drei Kopien eines nuklearen Ortsbestimmungssignals (NLS) fusioniert und ausgehend von einem murinen, retroviralen Stammzellvirus(MSCV)-Vektor exprimiert. Dasselbe sgRNA-Design, das in **Fig. 40B** gezeigt ist, wurde verwendet, um die Expression von einem RNA-Polymerase III-U6-Promotor zu bewirken. Eine Reporter-HEK293-Zelllinie, die EGFP unter dem SV40-Promotor exprimiert, wurde durch virale Infektion erzeugt. Unter Verwendung einer sgRNA (eNT2), die auf den Nicht-Matrizen-DNA-Strang der für EGFP kodierenden Region gerichtet war, wurde ein moderater, jedoch reproduzierbarer Knock-Down der Gen-Expression beobachtet (46% Repression, **Fig. 45A**). Diese Repression war sowohl vom dCas9-Protein als auch von sgRNA abhängig, was impliziert, dass die Repression durch den dCas9-sgRNA-Komplex und das durch RNA geführte Targeting erfolgte. Die gleiche sgRNA zeigte bei transienter Expression von einem Plasmid eine bessere Repression desselben Gens (63% Repression, **Fig. 52**). In Übereinstimmung mit dem bakteriellen System zeigten lediglich sgRNAs, die auf den Nicht-Matrizen-Strang gerichtet waren, eine Repression. Faktoren, wie der Abstand vom Transkriptions-Start und der Zustand des lokalen Chromatins können entscheidende Parameter bei der Bestimmung der Repressions-Effizienz sein (**Fig. 52**). Die Optimierung von dCas9- und sgRNA-Expression, der Stabilität, der Ortsbestimmung im Kern und der Interaktion wird eine weitere Verbesserung der Effizienz von CRISPRi in Säugetier-Zellen ermöglichen.

[0636] **Fig. 45** zeigt, dass CRISPRi die Gen-Expression in humanen Zellen reprimieren kann. (A) Ein CRISPRi-System in HEK293-Zellen. Die SV40-EGFP-Expressionskassette wurde über retrovirale Infektion in das Ge-

nom insertiert. Das dCas9-Protein wurde Kodon-optimiert und an drei Kopien einer NLS-Sequenz fusioniert. Die sgRNA wurde ausgehend von einem RNA-Polymerase III-U6-Vektor exprimiert. Die Co-Transfektion von dCas9 und einer sgRNA (eNT2), die auf den Nicht-Matrizen-Strang von EGFP gerichtet war, verringerte die Fluoreszenz (etwa 46%) während die Expression von dCas9 allein oder sgRNA allein keine Wirkung zeigte. (B) Die dCas9: sgRNA-vermittelte Repression war abhängig von den Target-Loci. Sieben sgRNAs wurden so erstellt, dass sie auf unterschiedliche Regionen der für EGFP kodierenden Sequenz auf dem Matrizen-Strang oder dem Nicht-Matrizen-Strang gerichtet sind. Lediglich eNT2 und eNT5 zeigten eine moderate Repression. Die Fluoreszenz-Ergebnisse von 7A und 7B stellen den Durchschnitt und den Fehler von zwei biologischen Replikaten dar.

[0637] Fig. 52 steht mit Fig. 45 in Zusammenhang und zeigt, dass die sgRNA-Repression von den Ziel-Loci und dem relativen Abstand vom Transkriptions-Start abhängt. Die gleiche sgRNA wurde verwendet, um dasselbe EGFP-Gen mit unterschiedlichen Promotoren zu reprimieren. Die dCas9/sgRNA-Komplexe reprimierten die Transkription von transient transfizierter Plasmid-DNA. Das Ausmaß der Transkriptions-Repression war geringfügig besser (63%) als die Repression, die für genomische Gene beobachtet wurde, und der prozentuale Anteil von GFP-negativen Zellen erhöhte sich in Gegenwart von sgRNA. Der Ziel-Locus weist einen abweichenden Abstand vom Transkriptionsstartpunkt auf. Während SV40-EGFP eine Repression zeigte, hatte LTR-EGFP keine Wirkung. Die Fluoreszenz-ergebnisse stellen den Durchschnitt und den Fehler von zwei biologischen Replikaten dar.

CRISPRi reprimiert die Transkription von Ziel-Genen effizient und selektive

[0638] Das CRISPRi-System ist eine verhältnismäßige einfache Plattform für die gezielte Gen-Regulation. CRISPRi beruht nicht auf dem Vorliegen von komplexen Wirtsfaktoren, sondern benötigt stattdessen lediglich das dCas9-Protein und Guide-RNAs und ist somit flexibel und in hohem Maße veränderbar. Das System kann in wirksamer Weise Gene in Bakterien unterdrücken. Das Silencing ist sehr wirksam, da keine unspezifischen Wirkungen nachgewiesen wurden. Darüber hinaus kann die Wirksamkeit des Knock-Downs verändert werden, indem die Target-Loci und der Grad der Basenpaarung zwischen der sgRNA und dem Ziel-Gen verändert werden. Dies wird es ermöglichen, allelische Gruppen von Hypomorphen zu erzeugen – ein Merkmal, das insbesondere für die Untersuchung von essentiellen Genen nützlich ist. Das System funktioniert durch direkte Blockierung der Transkription, in einer Weise, die einfach programmiert werden kann, indem sgRNAs erstellt werden. Mechanistisch unterscheidet sich dies von einem RNAi-basierten Silencing, das die Zerstörung von bereits transkribierten mRNAs erfordert.

[0639] Darüber hinaus können diese dCas9:sgRNA-Komplexe die Transkription modulieren, indem sie auf wichtige cis-wirkende Motive innerhalb eines beliebigen Promoters gerichtet werden, und sterisch die Assoziation ihrer verwandten trans-wirkenden Transkriptionsfaktoren blockieren. Somit könnte CRISPRi zusätzlich zu einer Verwendung als Gen-Knock-Down-Werkzeug auch für die funktionale Kartierung von Promotoren und anderen genomischen regulatorischen Modulen verwendet werden

CRISPRi ist der Genom-weiten Analyse und Regulation zugänglich

[0640] Das CRISPRi-Verfahren basiert auf der Verwendung von auf DNA gerichteten RNAs, und lediglich das auf DNA gerichtete Segment muss für spezifische Gen-Ziele erstellt werden. Mit dem Aufkommen der Synthese-Technologie für DNA-Oligonukleotide im großen Maßstab ist die Erzeugung großer Gruppen von Oligonukleotiden, die einzigartige Regionen von 20 bp umfassen, für das Genom-Targeting schnell und kostengünstig geworden. Diese Oligonukleotid-Bibliotheken könnten es uns erlauben, ein Targeting bei einer großen Anzahl von individuellen Genen durchzuführen, um die Gen-Funktion zu ermitteln, oder um ein Targeting bei Gen-Paaren durchzuführen, um genetische Interaktionen zu entschlüsseln. Desweiteren könnte CRISPRi verwendet werden, um gleichzeitig die Expression großer Gruppen von Genen zu modulieren, da die geringe Größe von sgRNAs es ermöglicht, mehrere Elemente im gleichen Expressions-Vektor zu verknüpfen.

CRISPRi stellt neue Werkzeuge für die Manipulation mikrobieller Genome bereit

[0641] Da die CRISPRi-Plattform kompakt und in sich geschlossen ist, kann sie an unterschiedliche Organismen angepasst werden. CRISPRi ist ein wirksames Werkzeug für die Untersuchung von Nicht-Modell-Organismen, für die genetische Modifikationsverfahren nicht hinreichend entwickelt sind, einschließlich pathogener oder industriell nützlicher Organismen. Anders als bei den meisten Eukaryonten fehlt den meisten Bakterien eine RNAi-Maschinerie. Daher ist die Regulation von endogenen Genen unter Verwendung von künstlich er-

stellten, synthetischen RNAs gegenwärtig begrenzt. CRISPRi könnte ein RNAi-ähnliches Verfahren für die Störung von Genen in Mikroben bereitstellen.

CRISPRi als Plattform zur Erstellung von transkriptionalen regulatorischen Netzwerken

[0642] CRISPRi kann als flexibler Rahmen für die Erstellung transkriptionaler regulatorischer Netzwerke verwendet werden. Da die CRISPRi-Plattform im Wesentlichen ein DNA-bindender Komplex ist, die durch eine RNA geleitet wird, stellt diese auch ein flexibles Gerüst für die Leitung unterschiedlicher regulatorischer Maschinerien zu spezifischen Stellen im Genom bereit. Über die einfache Blockierung der Transkription von Ziel-Genen hinaus ist es möglich, das dCas9-Protein mit zahlreichen regulatorischen Domänen zu koppeln, um unterschiedliche biologische Prozesse zu modulieren und unterschiedliche funktionale Ergebnisse zu erreichen (z. B. transkriptionale Aktivierung, Chromatin-Modifikationen).

[0643] Im CRISPRi-System ist es möglich, mehrere sgRNAs zu transkriptionalen Kreisen zu verschalten, in denen eine stromaufwärts positionierte sgRNA die Expression einer anderen, stromabwärts positionierten sgRNA kontrolliert. Da RNA-Moleküle in Mikroorganismen dazu neigen, kurzlebig zu sein, nehmen wir an, dass die genetischen Programme, die durch sgRNAs reguliert werden, eine schnelle Kinetik zeigen könnten, die sich von Kreisen unterscheidet, welche langsame Prozesse, wie Protein-Expression und Proteinabbau umfassen. Zusammenfassend gesagt ist das CRISPRi-System eine allgemeine genetische Programmierungsplattform, die für eine Vielzahl von biomedizinischen Anwendungen in Forschung und Klinik geeignet ist, einschließlich Genom-weitem funktionalem Profiling, metabolischer Modifikation bei Mikroben und Re-Programmierung von Zellen.

Beispiel 5: Ein chimäres ortsspezifisches Polypeptid kann verwendet werden, um die Transkription in humanen Zellen zu modulieren (zu aktivieren oder reprimieren)

[0644] Wir haben in humanen Zellen gezeigt, dass ein Fusionsprotein, welches ein katalytisch inaktives Cas9 und eine Aktivator-Domäne oder eine Repressor-Domäne umfasst, die Transkription von fremder Ziel-DNA erhöhen bzw. verringern kann.

[0645] Fig. 55. Wir haben das humanisierte, katalytisch inaktive Cas9 mit einer Transkriptions-Aktivator-Domäne VP64 fusioniert. (A) Um die Wirksamkeit bei der Gen-Aktivierung unter Verwendung dieses Systems zu untersuchen, haben wir einen induzierbaren GAL4-UAS-Promotor, der ein GFP kontrolliert, in das Genom von HEK293 (humane Gewebekulturzellen) insertiert. (B) Der GAL4-UAS-Promoter kann in Gegenwart des von Hefe abgeleiteten Proteins GAL4 induziert werden. Die dCas9-VP64-Fusion kann in wirksamer Weise GAL4-UAS in Gegenwart einer verwandten Guide-RNA, die an die GAL4-UAS-Region bindet, um den Faktor 20 aktivieren. (C) Mikroskopische Aufnahmen der dCas9-VP64-Aktivierung. (D) Durchflusszytometrie-Daten für die dCas9-VP64-Aktivierung.

[0646] Fig. 56. Wir haben das humanisierte, katalytisch inaktive Cas9 mit einer Transkriptions-Repressor-Domäne KRAP fusioniert. (Oben) Wir haben 10 Guide-RNAs erstellt, die auf einen gut charakterisierten Promoter, den SV40-Early-Promotor gerichtet sind, und eine Guide-RNA, die auf die für EGFP kodierende Region gerichtet ist. (Unten) Unter Verwendung einer nicht chimären dCas9 haben wir eine etwa 2 bis 3-fache Repression für gRNAs von P9 und NT2 beobachtet. Diese Wirksamkeit war bei Verwendung der dCas9-KRAP-Fusion stark verbessert. Beispielsweise zeigten P9 und NT2 mit der dCas9-KRAP-Fusion eine 20-fache bzw. 15-fache Repression. Darüber hinaus zeigten P1–P6 eine signifikante Verringerung der Expression bei Verwendung des Fusionsproteins, jedoch eine begrenzte Repression bei Verwendung eines nicht chimären dCas9.

Beispiel 6: dCas9 kann künstliche Guide-RNAs verwenden, die nicht in der Natur vorkommen, um die Spaltung der Ziel-DNA durchzuführen

[0647] Eine künstliche crRNA und eine künstliche tracrRNA wurden basierend auf dem Protein-bindenden Segment von natürlich vorkommenden Transkripten von crRNA und tracrRNAs aus *S. pyogenes* erstellt und so modifiziert, dass sie die asymmetrische Beule innerhalb des natürlichen crRNA:tracrRNA-Duplex von *S. pyogenes* nachahmen (siehe die Beule in der Protein-bindenden Domäne sowohl im künstlichen (oben) als auch im natürlichen (unten) RNA-Molekül, die in **Fig. 57A** dargestellt sind). Die künstlichen tracrRNA-Sequenzen zeigten weniger als 50% Identität mit der natürlichen tracrRNA. Die vorhergesagte Sekundärstruktur des crRNA:tracrRNA-Proteinbindungs-Duplex ist für beide RNA-Paare gleich, wobei sich jedoch die vorhergesagte Struktur der restlichen Teile der RNAs stark unterscheidet.

[0648] Fig. 57 zeigt, dass künstliche Sequenzen, die eine sehr geringe (etwa 50%) Identität mit natürlich vorkommenden tracrRNAs und crRNAs aufweisen, mit dCas9 zusammenwirken können, um Ziel-DNA zu spalten, solange die Struktur der Protein-bindenden Domäne der auf DNA gerichteten RNA konserviert ist. (A) Co-Faltung von tracrRNA und crRNA aus *S. pyogenes* und künstlicher tracrRNA und crRNA. (B) Es wurden Kombinationen von Cas9 und tracrRNA:crRNA-Orthologen von *S. pyogenes* verwendet, um Spaltungsassays mit Plasmid-DNA durchzuführen. Spy – *S. pyogenes*, Lin – *L. innocua*, Nme – *N. meningitidis*, Pmu – *P. multocida*. Cas9 von *S. pyogenes* kann durch einige, jedoch nicht durch alle tracrRNA-crRNA-Orthologe, die in ausgewählten bakteriellen Spezies natürlich vorkommen, geführt werden. Es ist anzumerken, dass Cas9 von *S. pyogenes* durch das künstliche tracrRNA:crRNA-Paar geführt werden kann, welches basierend auf der Struktur des Protein-bindenden Segments der natürlich vorkommenden, auf DNA gerichteten RNA unter Verwendung von Sequenz erstellt wurde, die vollständig unverwandt zum CRISPR-System ist.

[0649] Die künstliche "tracrRNA" (Aktivator-RNA), die verwendet wurde, war 5'-GUUUUCCCCUUUUUCAAAGAAUCUCCUGGGCACCUAUCUUCUAGGUGCCCUCUUUUAAACCUGACCAGUUAACCGG-CUGGUUAGGUUUUU-3'(SEQ ID NO: 1347). Die künstliche "crRNA" (Targeter-RNA), die verwendet wurde, war 5'-GAGAUUUUAUGAAAAGGGAAAAC-3' (SEQ ID NO: 1348).

Beispiel 7: Erzeugung eines nicht-humanen transgenen Organismus

[0650] Eine transgene Maus, die Cas9 exprimiert, (entweder unmodifiziert, so modifiziert, dass es verringerte enzymatische Aktivität aufweist, modifiziert als Fusionsprotein für einen der oben erwähnten Zwecke) wird erzeugt, wobei ein geeignetes Verfahren verwendet wird, das dem Fachmann bekannt ist (z. B. (i) Gen-Knock-In an einem Ziel-Locus (z. B. ROSA 26) einer embryonalen Stammzelle (ES-Zelle) der Maus, gefolgt von Blastozysten-Injektion und Erzeugung einer chimären Maus; (ii) Injektion eines auf Zufallsbasis integrierenden Transgens in den Pronukleus einer befruchteten Oozyte der Maus, gefolgt durch Ovum-Implantation in ein pseudoschwangeres Weibchen; usw.). Das Cas9-Protein steht unter der Kontrolle eines Promotors, der mindestens in embryonalen Stammzellen exprimiert und zusätzlich unter einer zeitlichen oder gewebespezifischen Kontrolle stehen kann (z. B. Wirkstoff-induzierbar, kontrolliert durch ein auf Cre/Lox-basierendes Promoter-System, usw.). Sobald eine Linie von transgenen Cas9-exprimierenden Mäusen erzeugt worden ist, werden embryonale Stammzellen isoliert und kultiviert, und in einigen Fällen werden ES-Zellen für die weitere künftige Verwendung eingefroren. Weil die isolierten ES-Zellen Cas9 exprimieren (und in einigen Fällen die Expression unter zeitlicher Kontrolle steht (z. B. Wirkstoff-induzierbar)), werden neue Knock-Out- oder Knock-In-Zellen (und damit Mäuse) schnell an jedem gewünschten Locus im Genom erzeugt, indem eine geeignete künstlich erstellte auf DNA gerichtete RNA eingebracht wird, welche Cas9 zu einem bestimmten ausgewählten Locus leitet. Ein solches System, und zahlreiche Variationen davon, werden verwendet, um neue genetisch modifizierte Organismen an einem beliebigen, ausgewählten Locus zu erzeugen. Wenn ein modifiziertes Cas9 verwendet wird, um die Transkription zu modulieren und/oder eine DNA zu modifizieren und/oder Polypeptide, die mit DNA assoziiert sind, zu modifizieren, werden die ES-Zellen selbst (oder beliebige differenzierte Zellen, die von diesen ES-Zellen abgeleitet sind (z. B. eine gesamte Maus, eine differenzierte Zelllinie, usw.) verwendet, um die Eigenschaften eines beliebigen ausgewählten Gens (oder eines beliebigen, ausgewählten Expressionsprodukts oder eines beliebigen, ausgewählten genomischen Locus) zu untersuchen, in dem man einfach eine geeignete auf DNA gerichtete RNA in eine gewünschte Cas9-exprimierende Zelle einbringt.

[0651] Während die vorliegende Erfindung in Bezug auf spezifische Ausführungsformen derselben beschrieben wurde, sollte der Fachmann verstehen, dass zahlreiche Änderungen durchgeführt werden und Äquivalente ausgetauscht werden können, ohne vom Kern und vom Umfang der Erfindung abzuweichen. Darüber hinaus können zahlreiche Modifikationen durchgeführt werden, um eine bestimmte Situation, ein Material, eine Zusammensetzung, ein Verfahren, einen Verfahrensschritt oder Verfahrensschritte an das Ziel, den Gedanken und den Umfang der vorliegenden Erfindung anzupassen. All diese Modifikationen sollen im Schutzbereich der vorliegend beigefügten Ansprüche liegen.

[0652] Die folgenden Dinge stellen bestimmte nicht-einschränkende Ausführungsformen der Erfindung dar:

1. Eine auf DNA-gerichtete RNA, die Folgendes umfasst:
 - (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (ii) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert.
2. Die auf DNA gerichtete RNA nach Ausführungsform 1, wobei das erste Segment 8 Nukleotide umfasst, die 100% Komplementarität zu einer Sequenz in der Ziel-DNA aufweist.
3. Die auf DNA gerichtete RNA nach Ausführungsform 1, wobei das zweite Segment eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nu-

- kleotiden zu einer der in SEQ ID NOs: 563–682 gezeigten Nukleotidsequenzen oder einem Komplement davon aufweist.
4. Die auf DNA gerichtete RNA nach Ausführungsform 1, wobei das zweite Segment eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 431–562 gezeigten Nukleotidsequenzen oder einem Komplement davon aufweist.
5. Die auf DNA gerichtete RNA nach Ausführungsform 1, wobei das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz umfasst, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.
6. Ein DNA-Polynukleotid, das eine Nukleotidsequenz umfasst, welche für die auf DNA gerichtete RNA nach Ausführungsform 1 kodiert.
7. Ein rekombinanter Expressionsvektor, der das DNA-Polynukleotid nach Ausführungsform 6 umfasst.
8. Der rekombinante Expressionsvektor nach Ausführungsform 7, wobei die Nukleotidsequenz, welche für die auf DNA gerichtete RNA kodiert, in funktionsfähiger Weise mit einem Promoter verknüpft ist.
9. Der rekombinante Expressionsvektor nach Ausführungsform 8, wobei der Promoter ein induzierbarer Promoter ist.
10. Der rekombinante Expressionsvektor nach Ausführungsform 7, wobei die Nukleotidsequenz, welche für die auf DNA gerichtete RNA nach Ausführungsform 1 kodiert, ferner eine multiple Klonierungsstelle enthält.
11. Eine in vitro genetisch modifizierte Wirtszelle, die das DNA-Polynukleotid nach Ausführungsform 6 umfasst.
12. Ein rekombinanter Expressionsvektor, der Folgendes umfasst:
- (i) eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:
 - (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und
 - (ii) eine Nukleotidsequenz, die für das ortsspezifisch modifizierende Polypeptid kodiert, und Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitäts-Teil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.
13. Ein rekombinanter Expressionsvektor, der Folgendes umfasst:
- (i) eine Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:
 - (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und
 - (ii) eine Nukleotidsequenz, die für das ortsspezifisch modifizierende Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitäts-Teil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.
14. Eine Variante eines ortsspezifisch modifizierenden Polypeptids, die Folgendes umfasst:
- (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (ii) einen Aktivitäts-Teil, der verringerte ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.
15. Die Variante eines ortsspezifisch modifizierenden Polypeptids nach Ausführungsform 14, umfassend eine H840A-Mutation der Sequenz von *S. pyogenes* von SEQ ID NO: 8 oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen.
16. Die Variante eines ortsspezifisch modifizierenden Polypeptids nach Ausführungsform 14, umfassend eine D10A-Mutation der Sequenz von *S. pyogenes* von SEQ ID NO: 8 oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen.
17. Die Variante eines ortsspezifisch modifizierenden Polypeptids nach Ausführungsform 14, die sowohl (i) eine D10A-Mutation der Sequenz von *S. pyogenes* von SEQ ID NO: 8 oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist, als auch (ii) eine H840A-Mutation der Sequenz von *S. pyogenes* von SEQ ID NO: 8 oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen.

18. Ein chimäres ortsspezifisch modifizierendes Polypeptid, das Folgendes umfasst:
 - (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (ii) einen Aktivitätsteil, der eine ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.
19. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18, umfassend eine Aminosäuresequenz, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.
20. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18, wobei die auf DNA gerichtete RNA ferner eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 563–682 gezeigten Nukleotidsequenzen aufweist.
21. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18, wobei die auf DNA gerichtete RNA ferner eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 431–562 gezeigten Nukleotidsequenzen aufweist.
22. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18, wobei die enzymatische Aktivität die Ziel-DNA modifiziert.
23. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 22, wobei die enzymatische Aktivität eine Nuklease-Aktivität, eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine DNA-Reparatur-Aktivität, eine DNA-Schädigungs-Aktivität, eine Deaminierungs-Aktivität, eine Dismutase-Aktivität, eine Alkylierungs-Aktivität, eine Depurinierungs-Aktivität, eine Oxidations-Aktivität, eine Pyrimidin-Dimer-bildende Aktivität, eine Integrase-Aktivität, eine Transposase-Aktivität, eine Rekombinase-Aktivität, eine Polymerase-Aktivität, eine Ligase-Aktivität, eine Helicase-Aktivität, eine Photolyase-Aktivität oder eine Glykosylase-Aktivität ist.
24. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 23, wobei die enzymatische Aktivität eine Nuklease-Aktivität ist.
25. Das chimäre ortsspezifisch modifizierenden Polypeptid nach Ausführungsform 24, wobei die Nuklease-Aktivität einen Doppelstrangbruch in die Ziel-DNA einführt.
26. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18, wobei die enzymatische Aktivität ein Ziel-Polypeptid modifiziert, das mit der Ziel-DNA assoziiert ist.
27. Das chimäre ortsspezifisch modifizierenden Polypeptid nach Ausführungsform 26, wobei die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität ist.
28. Das chimäre ortsspezifisch modifizierenden Polypeptid nach Ausführungsform 26, wobei das Ziel-Polypeptid ein Histon ist, und die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität oder eine deubiquitinierende Aktivität ist.
29. Ein RNA-Polynukleotid, das eine Nukleotidsequenz umfasst, welche für das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18 kodiert.
30. Ein DNA-Polynukleotid, das eine Nukleotidsequenz umfasst, welche für das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 18 kodiert.
31. Ein rekombinanter Expressionsvektor, der das Polynukleotid nach Ausführungsform 30 umfasst.
32. Der rekombinante Expressionsvektor nach Ausführungsform 31, wobei das Polynukleotid in funktionsfähiger Weise mit einem Promoter verknüpft ist.
33. Der rekombinante Expressionsvektor nach Ausführungsform 32, wobei der Promoter ein induzierbarer Promoter ist.
34. Eine in vitro genetisch modifizierte Wirtszelle, die das Polynukleotid nach Ausführungsform 30 umfasst.
35. Ein chimäres ortsspezifisch modifizierendes Polypeptid, das Folgendes umfasst:
 - (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (ii) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.
36. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 35, wobei der Aktivitätsteil die Transkription innerhalb der Ziel-DNA erhöht.

37. Das chimäre ortsspezifisch modifizierende Polypeptid nach Ausführungsform 35, wobei der Aktivitätsteil die Transkription innerhalb der Ziel-DNA verringert.
38. Eine genetisch modifizierte Zelle, die ein rekombinantes ortsspezifisch modifizierendes Polypeptid umfasst, das einen RNA-bindenden Teil umfasst, der mit einer auf DNA gerichteten RNA interagiert; und einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.
39. Die genetisch modifizierte Zelle nach Ausführungsform 38, wobei das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz umfasst, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.
40. Die genetisch modifizierte Zelle nach Ausführungsform 38, wobei die Zelle ausgewählt ist aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer Bakterien-Zelle, einer Eukaryonten-Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzen-Zelle, einer Algen-Zelle, einer Tier-Zelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fisch-Zelle, einer Frosch-Zelle, einer Vogel-Zelle, einer Säugetier-Zelle, einer Schweine-Zelle, einer Kuh-Zelle, einer Ziegen-Zelle, einer Schaf-Zelle, einer Nagetier-Zelle, einer Ratten-Zelle, einer Maus-Zelle, einer nicht-humanen Primaten-Zelle und einer humanen Zelle.
41. Ein transgener nicht-humaner Organismus, dessen Genom ein Transgen umfasst, welches eine Nukleotidsequenz umfasst, die für ein rekombinantes ortsspezifisch modifizierendes Polypeptid kodiert, das Folgendes umfasst: (i) einen RNA-bindenden Teil, der mit einer auf DNA gerichteten RNA interagiert; und (ii) einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.
42. Der transgene Organismus nach Ausführungsform 41, wobei das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz umfasst, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.
43. Der transgene Organismus nach Ausführungsform 41, wobei der Organismus ausgewählt ist aus der Gruppe bestehend aus: einem Archaeon, einem Bakterium, einem eukaryontischen einzelligen Organismus, einer Alge, einer Pflanze, einem Tier, einem Invertebraten, einer Fliege, einem Wurm, einem Nesseltier, einem Vertebraten, einem Fisch, einem Frosch, einem Vogel, einem Säugetier, einem Huftier, einem Nagetier, einer Ratte, einer Maus und einem nicht-humanen Primaten.
44. Eine Zusammensetzung, die Folgendes umfasst:
- (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:
 - (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (b) ein zweites Segment, das mit dem ortsspezifisch modifizierenden Polypeptid interagiert; und
 - (ii) das ortsspezifisch modifizierende Polypeptid, oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.
45. Die Zusammensetzung nach Ausführungsform 44, wobei das erste Segment der auf DNA gerichteten RNA 8 Nukleotide umfasst, die mindestens 100% Komplementarität zu einer Sequenz in der Ziel-DNA aufweisen.
46. Die Zusammensetzung nach Ausführungsform 44, wobei das zweite Segment der auf DNA gerichteten RNA eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 563–682 gezeigten Nukleotidsequenzen aufweist.
47. Die Zusammensetzung nach Ausführungsform 44, wobei das zweite Segment der auf DNA gerichteten RNA eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 431–562 gezeigten Nukleotidsequenzen aufweist.
48. Die Zusammensetzung nach Ausführungsform 44, wobei das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz umfasst, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.

49. Die Zusammensetzung nach Ausführungsform 44, wobei die enzymatische Aktivität die Ziel-DNA modifiziert.
50. Die Zusammensetzung nach Ausführungsform 49, wobei die enzymatische Aktivität eine Nuklease-Aktivität, eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine DNA-Reparatur-Aktivität, eine DNA-Schädigungs-Aktivität, eine Deaminierungs-Aktivität, eine Dismutase-Aktivität, eine Alkylierungs-Aktivität, eine Depurinierungs-Aktivität, eine Oxidations-Aktivität, eine Pyrimidin-Dimer-bildende Aktivität, eine Integrase-Aktivität, eine Transposase-Aktivität, eine Rekombinase-Aktivität, eine Polymerase-Aktivität, eine Ligase-Aktivität, eine Helicase-Aktivität, eine Photolyase-Aktivität oder eine Glykosylase-Aktivität ist.
51. Die Zusammensetzung nach Ausführungsform 50, wobei die enzymatische Aktivität eine Nuklease-Aktivität ist.
52. Die Zusammensetzung nach Ausführungsform 51, wobei die Nuklease-Aktivität einen Doppelstrangbruch in die Ziel-DNA einführt.
53. Die Zusammensetzung nach Ausführungsform 44, wobei die enzymatische Aktivität ein Ziel-Polypeptid modifiziert, das mit der Ziel-DNA assoziiert ist.
54. Die Zusammensetzung nach Ausführungsform 53, wobei die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität ist.
55. Die Zusammensetzung nach Ausführungsform 53, wobei das Ziel-Polypeptid ein Histon ist, und die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität oder eine deubiquitinierende Aktivität ist.
56. Die Zusammensetzung nach Ausführungsform 44, wobei die auf DNA gerichtete RNA ein auf DNA gerichtetes RNA-Doppelmolekül ist, und die Zusammensetzung sowohl eine Targeter-RNA als auch eine Aktivator-RNA umfasst, deren Duplex-bildende Segmente komplementär sind und hybridisieren, sodass das zweite Segment der auf DNA gerichteten RNA gebildet wird.
57. Die Zusammensetzung nach Ausführungsform 56, wobei das Duplex-bildende Segment der Aktivator-RNA eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 431–682 gezeigten Nukleotidsequenzen aufweist.
58. Eine Zusammensetzung, die Folgendes umfasst:
- (i) die auf DNA gerichtete RNA nach Ausführungsform 44 oder ein DNA-Polynukleotid, das für selbige kodiert; und
 - (ii) einen Puffer für die Stabilisierung von Nukleinsäuren.
59. Eine Zusammensetzung, die Folgendes umfasst:
- (i) das ortsspezifisch modifizierende Polypeptid nach Ausführungsform 44 oder ein Polynukleotid, das für selbiges kodiert; und
 - (ii) einen Puffer für die Stabilisierung von Nukleinsäuren und/oder Proteinen.
60. Eine Zusammensetzung, die Folgendes umfasst:
- (i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:
 - (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und
 - (ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.
61. Die Zusammensetzung nach Ausführungsform 60, wobei der Aktivitätsteil die Transkription innerhalb der Ziel-DNA erhöht.
62. Die Zusammensetzung nach Ausführungsform 60, wobei der Aktivitätsteil die Transkription innerhalb der Ziel-DNA verringert.
63. Eine Zusammensetzung, die Folgendes umfasst:
- (i) das ortsspezifisch modifizierende Polypeptid nach Ausführungsform 60, oder ein Polynukleotid, das für selbiges kodiert; und
 - (ii) einen Puffer für die Stabilisierung von Nukleinsäuren und/oder Proteinen.

64. Ein Verfahren zur ortsspezifischen Modifikation einer Ziel-DNA, wobei das Verfahren Folgendes umfasst:
- das Inkontaktbringen der Ziel-DNA mit:
- (i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:
 - (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und
 - (b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und
 - (ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität aufweist.
65. Das Verfahren nach Ausführungsform 64, wobei die Ziel-DNA extrachromosomal vorliegt.
66. Das Verfahren nach Ausführungsform 64, wobei die Ziel-DNA eine PAM-Sequenz des komplementären Strangs umfasst, die 5'-CCY-3' ist, wobei Y ein beliebiges DNA-Nukleotid ist, und Y unmittelbar 5' von der Ziel-Sequenz des komplementären Strangs der Ziel-DNA ist.
67. Das Verfahren nach Ausführungsform 64, wobei die Ziel-DNA Teil eines Chromosoms in vitro ist.
68. Das Verfahren nach Ausführungsform 64, wobei die Ziel-DNA Teil eines Chromosoms in vivo ist.
69. Das Verfahren nach Ausführungsform 64, wobei die Ziel-DNA Teil eines Chromosoms in einer Zelle ist.
70. Das Verfahren nach Ausführungsform 69, wobei die Zelle ausgewählt ist aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer Bakterien-Zelle, einer Eukaryonten-Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzen-Zelle, einer Algen-Zelle, einer Tier-Zelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fisch-Zelle, einer Frosch-Zelle, einer Vogel-Zelle, einer Säugetier-Zelle, einer Schweine-Zelle, einer Kuh-Zelle, einer Ziegen-Zelle, einer Schaf-Zelle, einer Nagetier-Zelle, einer Ratten-Zelle, einer Maus-Zelle, einer nicht-humanen Primaten-Zelle und einer humanen Zelle.
71. Das Verfahren nach Ausführungsform 64, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOS: 563–682 gezeigten Nukleotidsequenzen aufweist.
72. Das Verfahren nach Ausführungsform 64, wobei die auf DNA gerichtete RNA eine Nukleotidsequenz umfasst, die mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOS: 431–562 gezeigten Nukleotidsequenzen aufweist.
73. Das Verfahren nach Ausführungsform 64, wobei das DNA-modifizierende Polypeptid eine Aminosäuresequenz umfasst, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäuren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der in SEQ ID NOS: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.
74. Das Verfahren nach Ausführungsform 64, wobei die enzymatische Aktivität die Ziel-DNA modifiziert.
75. Das Verfahren nach Ausführungsform 74, wobei die enzymatische Aktivität eine Nuklease-Aktivität, eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine DNA-Reparatur-Aktivität, eine DNA-Schädigungs-Aktivität, eine Deaminierungs-Aktivität, eine Dismutase-Aktivität, eine Alkylierungs-Aktivität, eine Depurinierungs-Aktivität, eine Oxidations-Aktivität, eine Pyrimidin-Dimer-bildende Aktivität, eine Integrase-Aktivität, eine Transposase-Aktivität, eine Rekombinase-Aktivität, eine Polymerase-Aktivität, eine Ligase-Aktivität, eine Helicase-Aktivität, eine Photolyase-Aktivität oder eine Glykosylase-Aktivität ist.
76. Das Verfahren nach Ausführungsform 75, wobei die DNA-modifizierende enzymatische Aktivität eine Nuklease-Aktivität ist.
77. Das Verfahren nach Ausführungsform 76, wobei die Nuklease-Aktivität einen Doppelstrangbruch in die Ziel-DNA einführt.
78. Das Verfahren nach Ausführungsform 77, wobei das Inkontaktbringen unter Bedingungen erfolgt, die ein nicht-homologes End Joining oder eine Homologie-gesteuerte Reparatur ermöglichen.
79. Das Verfahren nach Ausführungsform 78, wobei das Verfahren ferner das Inkontaktbringen der Ziel-DNA mit einem Donor-Polynukleotid umfasst, wobei das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA integriert.
80. Das Verfahren nach Ausführungsform 78, wobei das Verfahren nicht das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid umfasst, wobei die Ziel-DNA so modifiziert wird, dass Nukleotide innerhalb der Ziel-DNA deletiert werden.
81. Das Verfahren nach Ausführungsform 64, wobei die enzymatische Aktivität ein Ziel-Polypeptid modifiziert, das mit der Ziel-DNA assoziiert ist.
82. Das Verfahren nach Ausführungsform 81, wobei die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-

Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität, eine deubiquitinierende Aktivität, eine Adenylierungs-Aktivität, eine Deadenylierungs-Aktivität, eine SUMOylierende Aktivität, eine deSUMOylierende Aktivität, eine Ribosylierungs-Aktivität, eine Deribosylierungs-Aktivität, eine Myristoylierungs-Aktivität oder eine Demyristoylierungs-Aktivität ist.

83. Das Verfahren nach Ausführungsform 81, wobei das Ziel-Polypeptid ein Histon ist und die enzymatische Aktivität eine Methyltransferase-Aktivität, eine Demethylase-Aktivität, eine Acetyltransferase-Aktivität, eine Deacetylase-Aktivität, eine Kinase-Aktivität, eine Phosphatase-Aktivität, eine Ubiquitin-Ligase-Aktivität oder eine deubiquitinierende Aktivität ist.

84. Das Verfahren nach Ausführungsform 64, wobei der Komplex ferner eine Aktivator-RNA umfasst.

85. Das Verfahren nach Ausführungsform 84, wobei die Aktivator-RNA eine Nukleotidsequenz umfasst, die mit mindestens 60% Identität über einen Abschnitt von mindestens 8 zusammenhängenden Nukleotiden zu einer der in SEQ ID NOs: 431–682 gezeigten Nukleotidsequenzen aufweist.

86. Ein Verfahren zur Modulierung der ortsspezifischen Transkription in einer Ziel-DNA, wobei das Verfahren das Inkontaktbringen der Ziel-DNA mit folgendem umfasst:

(i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der die Transkription moduliert,

wobei das Inkontaktbringen zur Modulierung der Transkription innerhalb der Ziel-DNA führt.

87. Das Verfahren nach Ausführungsform 86, wobei die Transkription innerhalb der Ziel-DNA erhöht wird.

88. Das Verfahren nach Ausführungsform 86, wobei die Transkription innerhalb der Ziel-DNA verringert wird.

89. Ein Verfahren zur ortsspezifischen Modifikation an einer Ziel-DNA, wobei das Verfahren Folgendes umfasst:

das Inkontaktbringen der Ziel-DNA mit:

(i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert.

90. Das Verfahren nach Ausführungsform 86, wobei das ortsspezifisch modifizierende Polypeptid die Transkription der Ziel-DNA erhöht.

91. Das Verfahren nach Ausführungsform 86, wobei das ortsspezifisch modifizierende Polypeptid die Transkription der Ziel-DNA verringert.

92. Ein Verfahren zur Förderung der ortsspezifischen Spaltung und Modifikation einer Ziel-DNA in einer Zelle, wobei das Verfahren das Einbringen in die Zelle von folgendem umfasst:

(i) einer auf DNA gerichteten RNA oder einem DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der eine Nuklease-Aktivität aufweist, die einen Doppelstrangbruch in der Ziel-DNA erzeugt;

wobei die Stelle des Doppelstrangbruchs durch die auf DNA gerichtete RNA bestimmt wird, das Inkontaktbringen unter Bedingungen erfolgt, die ein nicht-homologes End Joining oder eine Homologie-gesteuerte Reparatur ermöglichen, und die Ziel-DNA gespalten und wieder miteinander verbunden wird, sodass eine modifizierte DNA-Sequenz erzeugt wird.

93. Das Verfahren nach Ausführungsform 92, wobei das Verfahren ferner das Inkontaktbringen der Ziel-DNA mit einem Donor-Polynukleotid umfasst, wobei das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA integriert.

94. Das Verfahren nach Ausführungsform 92, wobei das Verfahren nicht das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid umfasst, wobei die Ziel-DNA so modifiziert wird, dass Nukleotide innerhalb der Ziel-DNA deletiert werden.

95. Das Verfahren nach Ausführungsform 92, wobei die Zelle ausgewählt ist aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer Bakterien-Zelle, einer Eukaryonten-Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzen-Zelle, einer Algen-Zelle, einer Tier-Zelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fisch-Zelle, einer Frosch-Zelle, einer Vogel-Zelle, einer Säugetier-Zelle, einer Schweine-Zelle, einer Kuh-Zelle, einer Ziegen-Zelle, einer Schaf-Zelle, einer Nagetier-Zelle, einer Ratten-Zelle, einer Maus-Zelle, einer nicht-humanen Primaten-Zelle und einer humanen Zelle.

96. Das Verfahren nach Ausführungsform 92, wobei die Zelle in vitro vorliegt.

97. Das Verfahren nach Ausführungsform 92, wobei die Zelle in vivo vorliegt.

98. Ein Verfahren zur Herstellung einer genetisch modifizierten Zelle in einem Subjekt, wobei das Verfahren Folgendes umfasst:

(I) das Einbringen in eine Zelle von folgendem:

(i) einer auf DNA gerichtete RNA oder einem DNA-Polynukleotid, das für selbige kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) einem ortsspezifisch modifizierenden Polypeptid oder einem Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der eine Nuklease-Aktivität aufweist, die einen Doppelstrangbruch in der Ziel-DNA erzeugt;

wobei die Stelle des Doppelstrangbruchs durch die auf DNA gerichtete RNA bestimmt wird, das Inkontaktbringen unter Bedingungen erfolgt, die ein nicht-homologes End Joining oder eine Homologie-gesteuerte Reparatur ermöglichen, und die Ziel-DNA gespalten und wieder miteinander verbunden wird, sodass eine modifizierte DNA-Sequenz erzeugt wird; wodurch die genetisch modifizierte Zelle erzeugt wird; und (II) das Transplantieren der genetisch modifizierten Zelle in das Subjekt.

99. Das Verfahren nach Ausführungsform 98, wobei das Verfahren ferner das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid umfasst, wobei das Donor-Polynukleotid, ein Teil des Donor-Polynukleotids, eine Kopie des Donor-Polynukleotids oder ein Teil der Kopie des Donor-Polynukleotids in die Ziel-DNA integriert.

100. Das Verfahren nach Ausführungsform 98, wobei das Verfahren nicht das Inkontaktbringen der Zelle mit einem Donor-Polynukleotid umfasst, wobei die Ziel-DNA so modifiziert wird, dass Nukleotide innerhalb der Ziel-DNA deletiert werden.

101. Das Verfahren nach Ausführungsform 98, wobei die Zelle ausgewählt ist aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer Bakterien-Zelle, einer Eukaryonten-Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzen-Zelle, einer Algen-Zelle, einer Tier-Zelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fisch-Zelle, einer Amphibien-Zelle, einer Vogel-Zelle, einer Säugetier-Zelle, einer Huftier-Zelle, einer Nagetier-Zelle, einer nicht-humanen Primaten-Zelle und einer humanen Zelle.

102. Ein Verfahren zur Modifizierung von Ziel-DNA in einer genetisch modifizierten Zelle, die eine Nukleotidsequenz umfasst, welche für ein exogenes ortsspezifisch modifizierendes Polypeptid kodiert, wobei das Verfahren die Einbringung einer auf DNA gerichteten RNA oder eines DNA-Polynukleotids, das für selbige kodiert, in die genetisch modifizierte Zelle umfasst, wobei:

(i) die auf DNA gerichtete RNA Folgendes umfasst:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in der Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der eine Nuklease-Aktivität aufweist.

103. Das Verfahren nach Ausführungsform 102, wobei das ortsspezifisch modifizierende Polypeptid eine Aminosäuresequenz umfasst, die mindestens etwa 75% Aminosäuresequenz-Identität zu den Aminosäu-

ren 7–166 oder 731–1003 der in **Fig. 3** gezeigten Aminosäuresequenz von Cas9/Csn1 oder zu den entsprechenden Teilen in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.

104. Das Verfahren nach Ausführungsform 102, wobei die Zelle ausgewählt ist aus der Gruppe bestehend aus: einer Archaeen-Zelle, einer Bakterien-Zelle, einer Eukaryonten-Zelle, einem eukaryontischen einzelligen Organismus, einer somatischen Zelle, einer Keimzelle, einer Stammzelle, einer Pflanzen-Zelle, einer Algen-Zelle, einer Tier-Zelle, einer Invertebraten-Zelle, einer Vertebraten-Zelle, einer Fisch-Zelle, einer Amphibien-Zelle, einer Vogel-Zelle, einer Säugetier-Zelle, einer Huftier-Zelle, einer Nagetier-Zelle, einer nicht-humanen Primaten-Zelle und einer humanen Zelle.

105. Das Verfahren nach Ausführungsform 102, wobei die Zelle in vivo vorliegt.

106. Das Verfahren nach Ausführungsform 102, wobei die Zelle in vitro vorliegt.

107. Das Verfahren nach Ausführungsform 102, wobei die Expression des ortsspezifisch modifizierenden Polypeptids unter der Kontrolle eines induzierbaren Promoters steht.

108. Das Verfahren nach Ausführungsform 102, wobei die Expression des ortsspezifisch modifizierenden Polypeptids unter der Kontrolle eines Zelltyp-spezifischen Promoters steht.

109. Ein Kit, der Folgendes umfasst:

(i) die auf DNA gerichtete RNA nach Ausführungsform 1 oder ein DNA-Polynukleotid, das für selbige kodiert; und

(ii) ein Reagenz zur Rekonstituierung und/oder Verdünnung.

110. Der Kit nach Ausführungsform 109, ferner umfassend ein Reagenz ausgewählt aus der Gruppe bestehend aus: einem Puffer für die Einbringung der auf DNA gerichteten RNA in Zellen, einem Waschpuffer, einem Kontrollreagenz, einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid, einem Reagenz für die Transkription der auf DNA gerichteten RNA ausgehend von DNA, und Kombinationen davon.

111. Ein Kit, der Folgendes umfasst:

(i) das ortsspezifisch modifizierende Polypeptid nach Ausführungsform 44 oder ein Polynukleotid, das für selbiges kodiert; und

(ii) ein Reagenz zur Rekonstituierung und/oder Verdünnung.

112. Der Kit nach Ausführungsform 111, ferner umfassend ein Reagenz das ausgewählt ist aus der Gruppe bestehend aus: einem Puffer für die Einbringung des ortsspezifisch modifizierenden Polypeptids in Zellen, einem Waschpuffer, einem Kontrollreagenz, einem Kontroll-Expressionsvektor oder einem RNA-Polynukleotid, einem Reagenz für die in vitro Herstellung des ortsspezifisch modifizierenden Polypeptids ausgehend von DNA, und Kombinationen davon.

113. Ein Kit, das Folgendes umfasst:

(i) das ortsspezifisch modifizierende Polypeptid nach Ausführungsform 60 oder ein Polynukleotid, das für selbiges kodiert; und

(ii) ein Reagenz zur Rekonstituierung und/oder Verdünnung.

114. Ein Kit, der Folgendes umfasst:

(i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbiges kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

115. Ein Kit, das Folgendes umfasst:

(i) eine auf DNA gerichtete RNA oder ein DNA-Polynukleotid, das für selbige kodiert, umfassend:

(a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist; und

(b) ein zweites Segment, das mit einem ortsspezifisch modifizierenden Polypeptid interagiert; und

(ii) das ortsspezifisch modifizierende Polypeptid oder ein Polynukleotid, das für selbiges kodiert, umfassend:

(a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

(b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

116. Ein Kit, der Folgendes umfasst:

(i) den rekombinanten Expressionsvektor nach Ausführungsform 12; und

(ii) ein Reagenz für die Rekonstituierung und/oder Verdünnung.

117. Ein Kit, der Folgendes umfasst:

- (i) den rekombinanten Expressionsvektor nach Ausführungsform 13; und
- (ii) ein Reagenz für die Rekonstituierung und/oder Verdünnung.

118. Ein Kit, der Folgendes umfasst:

- (i) den rekombinanten Expressionsvektor nach Ausführungsform 7; und
- (ii) ein Reagenz für die Rekonstituierung und/oder Verdünnung.

119. Ein Kit, der Folgendes umfasst:

- (i) den rekombinanten Expressionsvektor nach Ausführungsform 7; und
- (ii) einen rekombinanten Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitätsteil, der ortsspezifische enzymatische Aktivität aufweist, wobei die Stelle der enzymatischen Aktivität durch die auf DNA gerichtete RNA bestimmt wird.

120. Ein Kit, der Folgendes umfasst:

- (i) den rekombinanten Expressionsvektor nach Ausführungsform 7; und
- (ii) einen rekombinanten Expressionsvektor, der eine Nukleotidsequenz umfasst, die für ein ortsspezifisch modifizierendes Polypeptid kodiert, wobei das ortsspezifisch modifizierende Polypeptid Folgendes umfasst:
 - (a) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und
 - (b) einen Aktivitätsteil, der die Transkription innerhalb der Ziel-DNA moduliert, wobei die Stelle der modulierten Transkription innerhalb der Ziel-DNA durch die auf DNA gerichtete RNA bestimmt wird.

121. Ein Kit für das Targeting einer Ziel-DNA, der Folgendes umfasst: zwei oder mehrere auf DNA gerichtete RNAs nach Ausführungsform 1 oder DNA-Polynukleotide, die für selbige kodieren, wobei sich das erste Segment von mindestens einer der zwei oder mehreren auf DNA gerichteten RNAs in mindestens einem Nukleotid von dem ersten Segment von mindestens einer anderen der zwei oder mehreren auf DNA gerichteten RNAs unterscheidet.

122. Ein Verfahren zur selektiven Modulation der Transkription einer Ziel-DNA in einer Wirtszelle, wobei das Verfahren das Einbringen von folgendem in eine Wirtszelle umfasst:

- (a) einer auf DNA gerichtete RNA oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für die auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:

- (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Ziel-Sequenz in der Ziel-DNA ist;

- (ii) ein zweites Segment, das mit einem ortsspezifischen Polypeptid interagiert; und

- (iii) eine Stabilitäts-Kontroll-Sequenz; und

- (b) einem ortsspezifischen Varianten-Cas9-Polypeptid oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für das ortsspezifische Varianten-Cas9-Polypeptid kodiert, wobei das ortsspezifische Varianten-Cas9-Polypeptid Folgendes umfasst:

- (i) einen RNA-bindenden Teil, der mit der auf DNA gerichteten RNA interagiert; und

- (ii) einen Aktivitätsteil, der eine verringerte Endodesoxyribonuklease-Aktivität aufweist, wobei die auf DNA gerichtete RNA und das Varianten-Cas9-Polypeptid einen Komplex in der Zelle bilden, und wobei der Komplex selektiv die Transkription einer Ziel-DNA in der Zelle moduliert.

123. Ein Verfahren nach Ausführungsform 122, wobei das ortsspezifische Varianten-Cas9-Polypeptid den nicht-komplementären Strang der Ziel-DNA spalten kann, jedoch eine verringerte Fähigkeit aufweist, den komplementären Strang der Ziel-DNA zu spalten.

124. Das Verfahren nach Ausführungsform 123, wobei das ortsspezifische Varianten-Polypeptid eine H840A-Mutation der in **Fig. 39A** gezeigten Aminosäuresequenz oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.

125. Das Verfahren nach Ausführungsform 122, wobei das ortsspezifische Varianten-Polypeptid den komplementären Strang der Ziel-DNA spalten kann, jedoch eine verringerte Fähigkeit aufweist, den nicht-komplementären Strang der Ziel-DNA zu spalten.

126. Das Verfahren nach Ausführungsform 125, wobei das ortsspezifische Varianten-Polypeptid eine D10A-Mutation der in **Fig. 39A** gezeigten Aminosäuresequenz oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen aufweist.

127. Das Verfahren nach Ausführungsform 122, wobei das ortsspezifische Varianten-Polypeptid eine verringerte Fähigkeit hat, den komplementären Strang der Ziel-DNA zu spalten, und eine verringerte Fähigkeit, den nicht-komplementären Strang der Ziel-DNA zu spalten.

128. Das Verfahren nach Ausführungsform 127, wobei das ortsspezifische Varianten-Polypeptid sowohl (i) eine D10A-Mutation der in **Fig. 41** gezeigten Aminosäuresequenz oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen umfasst; als auch (ii) eine H840A-Mutation von SEQ ID NO: 8 oder die entsprechende Mutation in einer der in SEQ ID NOs: 1–256 und 795–1346 gezeigten Aminosäuresequenzen.

129. Das Verfahren nach Ausführungsform 122, wobei die Wirtszelle eine prokaryontische Zelle ist.

130. Das Verfahren nach Ausführungsform 122, wobei die Wirtszelle eine eukaryontische Zelle ist.
131. Das Verfahren nach Ausführungsform 130, wobei die eukaryontische Zelle eine Säugetier-Zelle ist.
132. Das Verfahren nach Ausführungsform 122, wobei die Transkription der Ziel-DNA im Vergleich zur Transkription der Ziel-DNA in Abwesenheit des Komplexes um mindestens etwa 10% inhibiert ist.
133. Das Verfahren nach Ausführungsform 122, wobei das ortsspezifische Varianten-Cas9-Polypeptid ein heterologes Polypeptid umfasst.
134. Das Verfahren nach Ausführungsform 133, wobei das heterologe Polypeptid eine Sequenz für eine subzelluläre Ortsbestimmung, eine nachweisbare Markierung, ein Stabilitäts-Kontroll-Peptid, eine Transkriptions-Modulationssequenz, eine Protein-bindende Sequenz oder eine Kombination davon umfasst.
135. Das Verfahren nach Ausführungsform 134, wobei das heterologe Polypeptid ein Stabilitäts-Kontroll-Peptid umfasst.
136. Das Verfahren nach Ausführungsform 135, wobei das Stabilitäts-Kontroll-Peptid eine Degron-Sequenz umfasst.
137. Das Verfahren nach Ausführungsform 133, wobei das heterologe Polypeptid ausgewählt ist aus der Gruppe bestehend aus: einem transkriptionalen Aktivator, einem transkriptionalen Repressor, einer Histon-Lysin-Methyltransferase, einer Histon-Lysin-Demethylase, einer Histon-Lysin-Acetyltransferase, einer Histon-Lysin-Deacetylase, einer DNA-Methylase, einer DNA-Demethylase, einem Begrenzungselement, einem peripheren Rekrutierungselement, einem Protein-Docking-Element und Kombinationen davon.
138. Das Verfahren nach Ausführungsform 133, wobei die Transkription der Ziel-DNA im Vergleich zur Transkription der Ziel-DNA in Abwesenheit des Komplexes um mindestens Faktor 1,2 erhöht ist.
139. Das Verfahren nach Ausführungsform 138, wobei das heterologe Polypeptid ausgewählt ist aus der Gruppe bestehend aus: einem transkriptionalen Aktivator, einer Histon-Lysin-Methyltransferase, einer Histon-Lysin-Demethylase, einer Histon-Lysin-Acetyltransferase, einer DNA-Demethylase, einem Protein-Dockingelement, und Kombinationen davon.
140. Das Verfahren nach Ausführungsform 122, wobei die auf DNA gerichtete RNA ein auf DNA gerichtetes RNA-Einzelmolekül ist.
141. Das Verfahren nach Ausführungsform 122, wobei die auf DNA gerichtete RNA eine auf DNA gerichtete RNA aus zwei Molekülen ist.
142. Das Verfahren nach Ausführungsform 122, bei dem man ferner eine zweiten auf DNA gerichtete RNA oder eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für die zweite auf DNA gerichtete RNA kodiert, in eine Wirtszelle einbringt, wobei die zweite auf DNA gerichtete RNA ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer zweiten Ziel-Sequenz in der Ziel-DNA ist; ein zweites Segment, das mit dem ortsspezifischen Polypeptid interagiert; und eine Stabilitäts-Kontroll-Sequenz umfasst.
143. Das Verfahren nach Ausführungsform 122, wobei das zweite Segment eine Länge von etwa 30 Nukleotiden bis etwa 60 Nukleotiden aufweist und eine Nukleotidsequenz umfasst, die mindestens etwa 95% Nukleotidsequenz-Identität zu einer Nukleotidsequenz aufweist, die ausgewählt ist aus:
- 1) 5'-GUUUUAGAGCUA-(Linker)-UAGCAAGUAAAA-3';
 - 2) 5'-GUUUUAGAGCUG-(Linker)-CAGCAAGUAAAA-3';
 - 3) 5'-GUUUUAGAGCUG-(Linker)-CAGCGAGUAAAA-3';
 - 4) 5'-GUUUUUGUACUCU-(Linker)-AGAAGCUACAAAGAU-3';
 - 5) 5'-GUUGUAGCUCC-(Linker)-GUUGCUACAAU-3'; und
 - 6) 5'-GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCG-3'.
144. Eine isolierte Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für eine auf DNA gerichtete RNA kodiert, umfassend:
- (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Zielsequenz in der Ziel-DNA ist;
 - (b) ein zweites Segment, das mit einem ortsspezifischen Polypeptid interagiert; und
 - (c) eine Stabilitäts-Kontroll-Sequenz.
145. Die isolierte Nukleinsäure nach Ausführungsform 144, wobei die Nukleotidsequenz, welche für die auf DNA gerichtete RNA kodiert, in funktionsfähiger Weise an einen Promoter geknüpft ist.
146. Die isolierte Nukleinsäure nach Ausführungsform 144, wobei die Nukleinsäure ferner eine Nukleotidsequenz umfasst, welche für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, das eine verringerte Endodesoxyribonuklease-Aktivität relativ zur Wildtyp-Cas9 aufweist.

147. Die isolierte Nukleinsäure nach Ausführungsform 146, wobei die Nukleotidsequenz, welche für das ortsspezifische Varianten-Cas9-Polypeptid kodiert, in funktionsfähiger Weise mit einem Promoter verknüpft ist.
148. Ein rekombinanter Expressionsvektor, der die Nukleinsäure nach Ausführungsform 24 umfasst.
149. Eine genetisch modifizierte Wirtszelle, welche die isolierte Nukleinsäure nach Ausführungsform 24 oder den rekombinanten Expressionsvektor nach Ausführungsform 28 umfasst.
150. Eine Bibliothek aus Nukleinsäuren, die eine Vielzahl von Mitgliedern umfasst, wobei jedes Nukleinsäure-Mitglied eine Nukleotidsequenz umfasst, die für eine auf DNA gerichtete RNA kodiert, umfassend:
- (a) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Zielsequenz in der Ziel-DNA ist;
 - (b) ein zweites Segment, das mit einem ortsspezifischen Polypeptid interagiert; und
 - (c) eine Stabilitäts-Kontroll-Sequenz,
- wobei sich jedes Mitglied von anderen Mitgliedern der Bibliothek in der Nukleotidsequenz des ersten Segments unterscheidet.
151. Eine Bibliothek nach Ausführungsform 150, wobei die Nukleotidsequenz, die für eine auf DNA gerichtete RNA kodiert, in funktionsfähiger Weise an einen Promoter geknüpft ist.
152. Ein Kit, der Folgendes umfasst:
- (a) eine auf DNA gerichtete RNA oder einer Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für die auf DNA gerichtete RNA kodiert, wobei die auf DNA gerichtete RNA Folgendes umfasst:
 - (i) ein erstes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Ziel-Sequenz in der Ziel-DNA ist;
 - (ii) ein zweites Segment, das mit einem ortsspezifischen Polypeptid interagiert; und
 - (iii) eine Stabilitäts-Kontroll-Sequenz; und
 - (b) einen Puffer.
153. Der Kit nach Ausführungsform 152, ferner umfassend ein ortsspezifisches Varianten-Cas9-Polypeptid, das eine verringerte Endodesoxyribonuklease-Aktivität relativ zum Wildtyp-Cas9 aufweist.
154. Der Kit nach Ausführungsform 152, wobei die Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für die auf DNA gerichtete RNA kodiert, ferner eine Nukleotidsequenz umfasst, welche für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, das eine verringerte Endodesoxyribonuklease-Aktivität relativ zum Wildtyp-Cas9 aufweist.
155. Der Kit nach Ausführungsform 152, ferner umfassend eine Nukleinsäure, die eine Nukleotidsequenz umfasst, welche für ein ortsspezifisches Varianten-Cas9-Polypeptid kodiert, das eine verringerte Endodesoxyribonuklease-Aktivität relativ zum Wildtyp-Cas9 aufweist.

Es folgt ein Sequenzprotokoll nach WIPO St. 25. Dieses kann von der amtlichen Veröffentlichungsplattform des DPMA heruntergeladen werden.

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

- US 7169874 [0111, 0483]
- WO 2001/018048 [0111, 0483]
- US 7029913 [0143]
- US 5843780 [0143]
- US 6200806 [0143]
- WO 99/20741 [0143]
- WO 01/51616 [0143]
- WO 03/020920 [0143]
- US 7153684 [0144]
- US 20090047263 [0145]
- US 20090068742 [0145]
- US 20090191159 [0145]
- US 20090227032 [0145]
- US 20090246875 [0145]
- US 20090304646 [0145]
- US 5489677 [0196]
- US 5602240 [0196]
- US 5034506 [0197, 0201]
- US 5539082 [0200]
- US 5714331 [0200]
- US 5719262 [0200]
- WO 98/39352 [0204]
- WO 99/14226 [0204]
- US 3687808 [0208]
- US 7078387 [0218, 0245, 0277, 0471]
- WO 94/12649 [0219, 0246, 0278, 0472]
- WO 93/03769 [0219, 0246, 0278, 0472]
- WO 93/19191 [0219, 0246, 0278, 0472]
- WO 94/28938 [0219, 0246, 0278, 0472]
- WO 95/11984 [0219, 0246, 0278, 0472]
- WO 95/00655 [0219, 0246, 0278, 0472]
- WO 93/09239 [0219, 0246, 0278, 0472]
- US 5222982 [0308, 0312]
- US 5385582 [0308, 0312]
- US 20080081064 [0312]
- US 20090196903 [0312]
- US 5451513 [0376]
- US 5545817 [0376]
- US 5545818 [0376]
- US 5576198 [0376]
- WO 95/16783 [0376]

Zitierte Nicht-Patentliteratur

- Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), insbesondere Kapitel 11 und Tabelle 11.1 [0093]
- Sambrook, J. and Russell, W., Molecular Cloning: A Laboratory Manual, 3. Auflage, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (2001) [0093]
- Sambrook et al., siehe oben, 11.7–11.8 [0094]
- Altschul et al., J. Mol. Biol., 1990, 215, 403–410 [0095]
- Zhang und Madden, Genome Res., 1997, 7, 649–656 [0095]
- Adv. Appl. Math., 1981, 2, 482–489 [0095]
- ncbi.nlm.nih.gov/BLAST [0100]
- ebi.ac.uk/Tools/msa/tcoffee/ [0100]
- ebi.ac.uk/Tools/msa/muscle/ [0100]
- mafft.cbrc.jp/alignment/software/ [0100]
- Altschul et al. (1990), J. Mol. Biol. 215: 403–10 [0100]
- Miyagishi et al., Nature Biotechnology 20, 497–500 (2002) [0105]
- Xia et al., Nucleic Acids Res. 2003 Sep 1; 31 (17) [0105]
- Chef et al. (1987) Cell 51: 7–19 [0108]
- Llewellyn, et al. (2010) Nat. Med. 16(10): 1161–1166 [0108]
- Oh et al. (2009) Gene Ther 16: 437 [0108]
- Sasaoka et al. (1992) Mol. Brain Res. 16: 274 [0108]
- Boundy et al. (1998) J. Neurosci. 18: 9989 [0108]
- Kaneda et al. (1991) Neuron 6: 583–594 [0108]
- Radovick et al. (1991) Proc. Natl. Acad. Sci. USA 88: 3402–3406 [0108]
- Oberdick et al. (1990) Science 248: 223–226 [0108]
- Bartge et al. (1988) Proc. Natl. Acad. Sci. USA 85: 3648–3652 [0108]
- Comb et al. (1988) EMBO J. 17: 3793–3805 [0108]
- Mayford et al. (1996) Proc. Natl. Acad. Sci. USA 93: 13250 [0108]
- Casanova et al. (2001) Genesis 31: 37 [0108]
- Liu et al. (2004) Gene Therapy 11: 52–60 [0108]
- Tozzo et al. (1997) Endocrinol. 138: 1604 [0109]
- Ross et al. (1990) Proc. Natl. Acad. Sci. USA 87: 9590 [0109]
- Pavjani et al. (2005) Nat. Med. 11: 797 [0109]

- Knight et al. (2003) Proc. Natl. Acad. Sci. USA 100: 14725 [0109]
- Kuriki et al. (2002) Biol. Pharm. Bull. 25: 1476 [0109]
- Sato et al. (2002) J. Biol. Chem. 277: 15703 [0109]
- Tabor et al. (1999) J. Biol. Chem. 274: 20603 [0109]
- Mason et al. (1998) Endocrinol. 139: 1013 [0109]
- Chen et al. (1999) Biochem. Biophys. Res. Comm. 262: 187 [0109]
- Kita et al. (2005) Biochem. Biophys. Res. Comm. 331: 484 [0109]
- Chakrabarti (2010) Endocrinol. 151: 2408 [0109]
- Platt et al. (1989) Proc. Natl. Acad. Sci. USA 86: 7490 [0109]
- Seo et al. (2003) Molec. Endocrinol. 17: 1522 [0109]
- Franz et al. (1997) Cardiovasc. Res. 35: 560–566 [0110]
- Robbins et al. (1995) Ann. N. Y. Acad. Sci. 752: 492–505 [0110]
- Linn et al. (1995) Circ. Res. 76: 584–591 [0110]
- Parmacek et al. (1994) Mol. Cell. Biol. 14: 1870–1885 [0110]
- Hunter et al. (1993) Hypertension 22: 608–617 [0110]
- Sartorelli et al. (1992) Proc. Natl. Acad. Sci. USA 89: 4047–4051 [0110]
- Akyürek et al. (2000) Mol. Med. 6: 983 [0111]
- Kim, et al. (1997) Mol. Cell. Biol. 17, 2266–2278 [0111]
- Li, et al., (1996) J. Cell Biol. 132, 849–859 [0111]
- Moessler, et al. (1996) Development 122, 2415–2425 [0111]
- Young et al. (2003) Ophthalmol. Vis. Sci. 44: 4076 [0112]
- Nicoud et al. (2007) J. Gene Med. 9: 1015 [0112]
- Yokoyama et al. (1992) Exp Eye Res. 55: 225 [0112]
- Panyam et al., Adv Drug Deliv Rev. 2012 Sep 13. pii: 50169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023 [0123]
- Ausubel, et al., (Short Protocols in Molecular Biology, 3. Aufl., Wiley & Sons, 1995) [0124]
- Morrison et al. (1997) Cell 88: 287–298 [0140]
- Thomson et. al, Science. 1998 Nov 6; 282 (5391): 1145–7 [0142]
- Takahashi et. al, Cell. 2007 Nov 30; 131(5): 861–72 [0142]
- Takahashi et. al, Nat Protoc. 2007; 2(12): 3081–9 [0142]
- Vu et. al, Science. 2007 Dez 21; 318(5858): 1917–20. Epub 2007 Nov 20 [0142]
- Thomson et al. (1998) Science 282: 1145 [0143]
- Thomson et al. (1995) Proc. Natl. Acad. Sci. USA 92: 7844 [0143]
- Thomson et al. (1996) Biol. Reprod. 55: 254 [0143]
- Shambloott et al., Proc. Natl. Acad. Sci. USA 95: 13726, 1998 [0143]
- Matsui, Y., et al., (1992) Cell 70: 841 [0144]
- Shambloott, M., et al. (2001) Proc. Natl. Acad. Sci. USA 98: 113 [0144]
- Shambloott, M., et al. (1998) Proc. Natl. Acad. Sci. USA, 95: 13726 [0144]
- Koshimizu, U., et al. (1996) Development, 122: 1235 [0144]
- Molecular Cloning: A Laboratory Manual, 3. Aufl. (Sambrook et al., HaRBor Laboratory Press 2001) [0154]
- Short Protocols in Molecular Biology, 4. Aufl. (Ausubel et al. eds., John Wiley & Sons 1999) [0154]
- Protein Methods (Bollag et al., John Wiley & Sons 1996) [0154]
- Nonviral Vectors for Gene Therapy (Wagner et al. Hrsg., Academic Press 1999) [0154]
- Viral Vectors (Kapliff & Loewy Hrsg., Academic Press 1995) [0154]
- Immunology Methods Manual (I. Lefkovits Hrsg., Academic Press 1997) [0154]
- Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998) [0154]
- Nakamura et al., Genes Cells. 2012 Mai; 17 (5): 344–64 [0175]
- Vavalle et al., Future Cardiol. 2012 Mai; 8(3): 371–82 [0175]
- Citartan et al., Biosens Bioelectron. 2012 Apr 15; 34(1): 1–11 [0175]
- Liberman et al., Wiley Interdiscip Rev RNA. 2012 Mai–Juni; 3(3): 369–84 [0175]
- Dwaine A. Braasch und David R. Corey, Biochemistry, 2002, 41(14), 4503–4510 [0201]
- Wang et al., J. Am. Chem. Soc., 2000, 122, 8595–8602 [0202]
- Singh et al., Chem. Commun., 1998, 4, 455–456 [0203]
- Wahlestedt et al., Proc. Natl. Acad. Sci. U.S.A., 2000, 97, 5633–5638 [0203]
- Koshkin et al., Tetrahedron, 1998, 54, 3607–3630 [0204]
- Martin et al., Helv. Chim. Acta, 1995, 78, 486–504 [0205]
- The Concise Encyclopedia Of Polymer Science And Engineering, Seiten 858–859 [0208]
- Kroschwitz, J. I., Hrsg. John Wiley & Sons, 1990 [0208]
- English et al., Angewandte Chemie, International Edition, 1991, 30, 613 [0208]

- Sanghvi, Y. S., Kapitel 15, Antisense Research and Applications, Seiten 289–302 [0208]
- Crooke, S. T. und Lebleu, B., Hrsg., CRC Press, 1993 [0208]
- Sanghvi et al., Hrsg., Antisense Research and Applications, CRC Press, Boca Raton, 1993, Seiten 276–278 [0208]
- Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553–6556 [0210]
- Manoharan et al., Bioorg. Med. Chem. Lett., 1994, 4, 1053–1060 [0210]
- Manoharan et al., Ann. N. Y. Acad. Sci., 1992, 660, 306–309 [0210]
- Manoharan et al., Bioorg. Med. Chem. Lett., 1993, 3, 2765–2770 [0210]
- Oberhauser et al., Nucl. Acids Res., 1992, 20, 533–538 [0210]
- Saison-Behmoaras et al., EMBO J., 1991, 10, 1111–1118 [0210]
- Kabanov et al., FEBS Lett., 1990, 259, 327–330 [0210]
- Svinarchuk et al., Biochimie, 1993, 75, 49–54 [0210]
- Manoharan et al., Tetrahedron Lett., 1995, 36, 3651–3654 [0210]
- Shea et al., Nucl. Acids Res., 1990, 18, 3777–3783 [0210]
- Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969–973 [0210]
- Manoharan et al., Tetrahedron Lett., 1995, 36, 3651–3654 [0210]
- Mishra et al., Biochim. Biophys. Acta, 1995, 1264, 229–237 [0210]
- Crooke et al., J. Pharmacol. Exp. Ther., 1996, 277, 923–937 [0210]
- Inder et al. (2002) Cancer Gene Ther. 9(6): 489–96 [0211]
- Noguchi et al. (2003) Diabetes 52(7): 1732–1737 [0211]
- Trehin et al. (2004) Pharm. Research 21: 1248–1256 [0211]
- Wender et al. (2000) Proc. Natl. Acad. Sci. USA 97: 13003–13008 [0211]
- Aguilera et al. (2009) Integr Biol (Camb) June; 1(5–6): 371–381 [0211]
- Lee et al., Invest Ophthalmol Vis Sci 35: 2543–2549, 1994 [0219]
- Borrás et al., Gene Ther 6: 515–524, 1999 [0219]
- Li and Davidson, PNAS 92: 7700–7704, 1995 [0219]
- Sakamoto et al., H Gene Ther 5: 1088–1097, 1999 [0219]
- Ali et al., Hum Gene Ther 9: 81–86, 1998 [0219]
- Flannery et al., PNAS 94: 6916–6921, 1997 [0219]
- Bennett et al., Invest Ophthalmol Vis Sci 38: 2857–2863, 1997 [0219]
- Jomary et al., Gene Ther 4: 683–690, 1997 [0219]
- Rolling et al., Hum Gene Ther 10: 641–648, 1999 [0219]
- Ali et al., Hum Mol Genet 5: 591–594, 1996 [0219]
- Samulski et al., J. Vir. (1989) 63: 3822–3828 [0219]
- Mendelson et al., Virol. (1988) 166: 154–165 [0219]
- Flotte et al., PNAS (1993) 90: 10613–10617 [0219]
- Miyoshi et al., PNAS 94: 10319–10323, 1997 [0219]
- Takahashi et al., J Virol 73: 7812–7816, 1999 [0219]
- Bitter et al. (1987) Methods in Enzymology, 153: 516–544 [0221]
- Panyam et al., al Adv Drug Deliv Rev. 2012 Sep 13. pii: S0169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023 [0225]
- Li et al., Invest Ophthalmol Vis Sci 35: 2543–2549, 1994 [0246]
- Borrás et al., Gene Ther 6: 515–524, 1999 [0246]
- Li und Davidson, PNAS 92: 7700–7704, 1995 [0246]
- Sakamoto et al., H Gene Ther 5: 1088–1097, 1999 [0246]
- Ali et al., Hum Gene Ther 9: 81–86, 1998 [0246]
- Flannery et al., PNAS 94: 6916–6921, 1997 [0246]
- Bennett et al., Invest Ophthalmol Vis Sci 38: 2857–2863, 1997 [0246]
- Jomary et al., Gene Ther 4: 683–690, 1997 [0246]
- Rolling et al., Hum Gene Ther 10: 641–648, 1999 [0246]
- Ali et al., Hum Mol Genet 5: 591–594, 1996 [0246]
- Samulski et al., J. Vir. (1989) 63: 3822–3828 [0246]
- Mendelson et al., Virol. (1988) 166: 154–165 [0246]
- Flotte et al., PNAS (1993) 90: 10613–10617 [0246]
- Miyoshi et al., PNAS 94: 10319–10323, 1997 [0246]
- Takahashi et al., J Virol 73: 7812–7816, 1999 [0246]
- Bitter et al. (1987) Methods in Enzymology, 153: 516–544 [0248]
- Panyam et al., al Adv Drug Deliv Rev. 2012 Sep 13. pii: S0169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023 [0252]
- Kawai et al., Bioeng Bugs. 2010 Nov–Dec; 1(6): 395–403: "Transformation of Saccharomyces cerevisiae and other fungi: methods and possible underlying mechanism" [0272]

- Tanka et al., *Nature*. 2004 Mar 18; 428(6980): 323–8: "Conformational variations in an infectious protein determine prion strain differences" [0272]
- Lee et al., *Invest Ophthalmol Vis Sci* 35: 2543 2549, 1994 [0278]
- Borrás et al., *Gene Ther* 6: 515 524, 1999 [0278]
- Li and Davidson, *PNAS* 92: 7700 7704, 1995 [0278]
- Sakamoto et al., *H Gene Ther* 5: 1088 1097, 1999 [0278]
- Ali et al., *Hum Gene Ther* 9: 81 86, 1998 [0278]
- Flannery et al., *PNAS* 94: 6916 6921, 1997 [0278]
- Bennett et al., *Invest Ophthalmol Vis Sci* 38: 2857 2863, 1997 [0278]
- Jomary et al., *Gene Ther* 4: 683 690, 1997 [0278]
- Rolling et al., *Hum Gene Ther* 10: 641 648, 1999 [0278]
- Ali et al., *Hum Mol Genet* 5: 591 594, 1996 [0278]
- Samulski et al., *J. Vir.* (1989) 63: 3822–3828 [0278]
- Mendelson et al., *Virology* (1988) 166: 154–165 [0278]
- Flotte et al., *PNAS* (1993) 90: 10613–10617 [0278]
- Miyoshi et al., *PNAS* 94: 10319 23, 1997 [0278]
- Takahashi et al., *J Virol* 73: 7812 7816, 1999 [0278]
- Bitter et al. (1987) *Methods in Enzymology*, 153: 516–544 [0281]
- Beumer et al. (2008) "Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases", *PNAS* 105(50): 19821–19826 [0283]
- Futaki et al. (2003) *Curr Protein Pept Sci*. 2003 Apr; 4(2): 87–9 und 446 [0288]
- Wender et al. (2000) *Proc. Natl. Acad. Sci. U.S.A* 2000 Nov. 21; 97(24): 13003–8 [0288]
- Wender et al. 2000 [0288]
- Uemura et al. 2002 [0288]
- Chang et al. (1987) *Proc. Natl. Acad Sci USA* 84: 4959–4963 [0303]
- Nehls et al. (1996) *Science* 272: 886–889 [0303]
- Remington's Pharmaceutical Sciences, Mace Publishing Company, Philadelphia, Pa., 17. Auflage (1985) [0319]
- Langer, *Science* 249: 1527–1533 (1990) [0319]
- Cho et al., *Curr Protoc Cell Biol*. 2009 Mar; Chapter 19: Unit 19.11: Generation of transgenic mice [0361]
- Gama et al., *Brain Struct Funct*. 2010 Mar; 214(2–3): 91–109. Epub 2009 Nov 25: *Animal transgenesis: an overview* [0361]
- Husaini et al., *GM Crops*. 2011 Jun–Dec; 2(3) : 150–62. Epub 2011 Jun 1: *Approaches for gene targeting and targeted gene expression in plants* [0361]
- Glick und Thompson, (Hrsg.), *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton, Fla.: CRC Press (1993) [0374]
- Klein et al. (*Nature* 327: 70–73 (1987) [0375]
- Weissbach und Weissbach, (1989) *Methods for Plant Molecular Biology* Academic Press [0376]
- Gelvin et al., (1990) *Plant Molecular Biology Manual*, Kluwer Academic Publishers [0376]
- Herrera-Estrella et al. (1983) *Nature* 303: 209, Bevan (1984) *Nucl Acid Res.* 12: 8711–8721 [0376]
- Klee (1985) *Bio/Technology* 3: 637–642 [0376]
- Christou (1991) *Bio/Technology* 9: 957–9 und 4462 [0376]
- Gordon-Kamm (1990) *Plant Cell* 2: 603–618 [0376]
- Weeks et al. (1993) *Plant Physiol* 102: 1077–1084 [0376]
- Vasil (1993) *Bio/Technology* 10: 667–674 [0376]
- Wan und Lemeaux (1994) *Plant Physiol* 104: 37–48 [0376]
- Ishida et al. (1996) *Nature Biotech* 14: 745–750 [0376]
- Danieli et al *Nat. Biotechnol* 16: 345–348, 1998 [0376]
- Staub et al *Nat. Biotechnol* 18: 333–338, 2000 [0376]
- O'Neill et al *Plant J.* 3: 729–738, 1993 [0376]
- Knoblauch et al *Nat. Biotechnol* 17: 906–909 [0376]
- Boynton et al., *Methods in Enzymology* 217: 510–536 (1993) [0376]
- Svab et al., *Proc. Natl. Acad. Sci. USA* 90: 913–917 (1993) [0376]
- McBride et al., *Proc. Nat. Acad. Sci. USA* 91: 7301–7305 (1994) [0376]
- Nakamura et al., *Genes Cells*. 2012 Mai; 17 (5): 344–64 [0396]
- Vavalle et al., *Future Cardiol.* 2012 Mai; 8(3): 371–82 [0396]
- Citartan et al., *Biosens Bioelectron.* 2012 Apr 15; 34(1): 1–11 [0396]
- Liberman et al., *Wiley Interdiscip Rev RNA*. 2012 Mai–Juni; 3(3): 369–84 [0396]
- Dohmen et al., *Science*, 1994. 263(5151): S. 1273–1276: *Heat-inducible degron: a method for constructing temperature-sensitive mutants* [0459]
- Schoeber et al., *Am J Physiol Renal Physiol*. 2009 Jan; 296(1): F204–11: *Conditional fast expression and function of multimeric TRPV5 channels using Shield-1* [0459]

- Chu et al., *Bioorg Med Chem Lett.* 2008 Nov 15; 18(22): 5941–4: Recent progress with FKBP-derived destabilizing domains; Kane-maki, *Pflugers Arch.* 2012 Dec 28: Frontiers of protein expression control with conditional degrons [0459]
- Yang et al., *Mol Cell.* 2012 Nov 30; 48(4): 487–8: Titivated for destruction: the methyl degron [0459]
- Barbour et al., *Biosci Rep.* 2013 Jan 18; 33(1) : Characterization of the bipartite degron that regulates ubiquitin-independent degradation of thymidylate synthase [0459]
- Greussing et al., *J Vis Exp.* 2012 Nov 10; (69) : Monitoring of ubiquitin-proteasome activity in living cells using a Degron(dgn)-destabilized green fluorescent protein(GFP)-based reporter protein [0459]
- Panyam et al., *Adv Drug Deliv Rev.* 2012 Sep 13. pii: 50169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023 [0468]
- Lee et al., *Invest Ophthalmol Vis Sci* 35: 2543 2549, 1994 [0472]
- Borrás et al., *Gene Ther* 6: 515 524, 1999 [0472]
- Li and Davidson, *PNAS* 92: 7700 7704, 1995 [0472]
- Sakamoto et al., *H Gene Ther* 5:1088 1097, 1999 [0472]
- Ali et al., *Hum Gene Ther* 9: 81 86, 1998 [0472]
- Flannery et al., *PNAS* 94: 6916 6921, 1997 [0472]
- Bennett et al., *Invest Ophthalmol Vis Sci* 38: 2857 2863, 1997 [0472]
- Jomary et al., *Gene Ther* 4: 683 690, 1997 [0472]
- Rolling et al., *Hum Gene Ther* 10: 641 648, 1999 [0472]
- Ali et al., *Hum Mol Genet* 5: 591 594, 1996 [0472]
- Samulski et al., *J. Vir.* (1989) 63: 3822–3828 [0472]
- Mendelson et al., *Virology* (1988) 166: 154–165 [0472]
- Flotte et al., *PNAS* (1993) 90: 10613–10617 [0472]
- Miyoshi et al., *PNAS* 94: 10319 23, 1997 [0472]
- Takahashi et al., *J Virol* 73: 7812 7816, 1999 [0472]
- Bitter et al. (1987) *Methods in Enzymology*, 153: 516–544 [0474]
- Miyagishi et al., *Nature Biotechnology* 20, 497–500 (2002) [0477]
- Xia et al., *Nucleic Acids Res.* 2003 Sep 1; 31 (17) [0477]
- Chef et al. (1987) *Cell* 51: 7–19 [0480]
- Llewellyn, et al. (2010) *Nat. Med.* 16(10): 1161–1166 [0480]
- Oh et al. (2009) *Gene Ther* 16: 437 [0480]
- Sasaoka et al. (1992) *Mol. Brain Res.* 16: 274 [0480]
- Boundy et al. (1998) *J. Neurosci.* 18: 9989 [0480]
- Kaneda et al. (1991) *Neuron* 6: 583–594 [0480]
- Radovick et al. (1991) *Proc. Natl. Acad. Sci. USA* 88: 3402–3406 [0480]
- Oberdick et al. (1990) *Science* 248: 223–226 [0480]
- Bartge et al. (1988) *Proc. Natl. Acad. Sci. USA* 85: 3648–3652 [0480]
- Comb et al. (1988) *EMBO J.* 17: 3793–3805 [0480]
- Mayford et al. (1996) *Proc. Natl. Acad. Sci. USA* 93: 13250 [0480]
- Casanova et al. (2001) *Genesis* 31: 37 [0480]
- Liu et al. (2004) *Gene Therapy* 11: 52–60 [0480]
- Tozzo et al. (1997) *Endocrinol.* 138: 1604 [0481]
- Ross et al. (1990) *Proc. Natl. Acad. Sci. USA* 87: 9590 [0481]
- Pavjani et al. (2005) *Nat. Med.* 11: 797 [0481]
- Knight et al. (2003) *Proc. Natl. Acad. Sci. USA* 100: 14725 [0481]
- Kuriki et al. (2002) *Biol. Pharm. Bull.* 25: 1476 [0481]
- Sato et al. (2002) *J. Biol. Chem.* 277: 15703 [0481]
- Tabor et al. (1999) *J. Biol. Chem.* 274: 20603 [0481]
- Mason et al. (1998) *Endocrinol.* 139: 1013 [0481]
- Chen et al. (1999) *Biochem. Biophys. Res. Comm.* 262: 187 [0481]
- Kita et al. (2005) *Biochem. Biophys. Res. Comm.* 331: 484 [0481]
- Chakrabarti (2010) *Endocrinol.* 151: 2408 [0481]
- Platt et al. (1989) *Proc. Natl. Acad. Sci. USA* 86: 7490 [0481]
- Seo et al. (2003) *Molec. Endocrinol.* 17: 1522 [0481]
- Franz et al. (1997) *Cardiovasc. Res.* 35: 560–566 [0482]
- Robbins et al. (1995) *Ann. N. Y. Acad. Sci.* 752: 492–505 [0482]
- Linn et al. (1995) *Circ. Res.* 76: 584–591 [0482]
- Parmacek et al. (1994) *Mol. Cell. Biol.* 14: 1870–1885 [0482]
- Hunter et al. (1993) *Hypertension* 22: 608–617 [0482]
- Sartorelli et al. (1992) *Proc. Natl. Acad. Sci. USA* 89: 4047–4051 [0482]
- Akyürek et al. (2000) *Mol. Med.* 6: 983 [0483]

- Kim, et al. (1997) Mol. Cell. Biol. 17, 2266–2278 [0483]
- Li, et al., (1996) J. Cell Biol. 132, 849–859 [0483]
- Moessler, et al. (1996) Development 122, 2415–2425 [0483]
- Young et al. (2003) Ophthalmol. Vis. Sci. 44: 4076 [0484]
- Nicoud et al. (2007) J. Gene Med. 9: 1015 [0484]
- Yokoyama et al. (1992) Exp Eye Res. 55: 225 [0484]
- Grissa I et al., BMC Bioinformatics 2007; 8: 172 [0574]
- Grissa I et al., Nucleic Acids Res 2007 [0574]
- fruitfly.org/seq_tools/promoter.html [0575]

Schutzansprüche

1. Zusammensetzung, die Folgendes umfasst:
 - (i) ein Cas9-Polypeptid; und
 - (ii) ein auf DNA-gerichtetes RNA-Einzelmolekül, das Folgendes umfasst:
 - (a) ein auf DNA-gerichtetes Segment, das eine Nukleotidsequenz umfasst, die komplementär zu einer Sequenz in einer Ziel-DNA ist, und
 - (b) ein Protein-bindendes Segment, das mit einem Cas9-Polypeptid interagiert, wobei das Protein-bindende Segment zwei komplementäre Abschnitte von Nukleotiden umfasst, die so hybridisieren, dass ein doppelsträngiger RNA(dsRNA)-Duplex gebildet wird, wobei der dsRNA-Duplex komplementäre Nukleotide einer tracrRNA und einer CRISPR-RNA (crRNA) umfasst, und wobei die zwei komplementären Abschnitte von Nukleotiden kovalent durch zwischenliegende Nukleotide miteinander verbunden sind.
2. Zusammensetzung nach Anspruch 1, wobei der dsRNA-Duplex eine Länge von 8 Basenpaaren (bp) bis 30 bp aufweist.
3. Zusammensetzung nach einem der Ansprüche 1 oder 2, wobei die prozentuale Komplementarität zwischen den Nukleotiden, die hybridisieren, um den dsRNA-Duplex des Protein-bindenden Segments zu bilden, mindestens 70% beträgt.
4. Zusammensetzung, die Folgendes umfasst:
 - (i) ein chimäres Cas9-Polypeptid, wobei das chimäre Cas9-Polypeptid (a) eine Cas9-Polypeptidsequenz umfasst, die kovalent mit (b) einer Protein-Transduktionsdomäne verbunden ist; und
 - (ii) ein auf DNA gerichtetes RNA-Einzelmolekül, umfassend:
 - (a) ein auf DNA gerichtetes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in einer Ziel-DNA ist; und
 - (b) ein Protein-bindendes Segment, das mit einem Cas9-Polypeptid interagiert, wobei das Protein-bindende Segment zwei komplementäre Abschnitte von Nukleotiden umfasst, die so hybridisieren, dass ein doppelsträngiger RNA(dsRNA)-Duplex gebildet wird, wobei der dsRNA-Duplex komplementäre Nukleotide einer tracrRNA und einer CRISPR-RNA (crRNA) umfasst, und wobei die zwei komplementären Abschnitte von Nukleotiden kovalent durch zwischenliegende Nukleotide verbunden sind.
5. Zusammensetzung nach Anspruch 4, wobei die Protein-Transduktionsdomäne kovalent an den Amino-Terminus der Cas9-Polypeptidsequenz gebunden ist.
6. Zusammensetzung nach Anspruch 4, wobei die Protein-Transduktionsdomäne kovalent an den Carboxy-Terminus der Cas9-Polypeptidsequenz gebunden ist.
7. Zusammensetzung nach einem der Ansprüche 4–6, wobei die Protein-Transduktionsdomäne in der Lage ist, den Übergang vom Cytosol in das Innere einer Organelle einer Zelle zu ermöglichen.
8. Zusammensetzung, die Folgendes umfasst:
 - (i) ein chimäres Cas9-Polypeptid; wobei das chimäre Cas9-Polypeptid Folgendes umfasst:
 - (a) eine Cas9-Polypeptidsequenz, und
 - (b) eine heterologe Polypeptidsequenz, die eine Aktivität bereitstellt, welche ausgewählt ist aus: einer Ziel-DNA-modifizierenden Aktivität, einer Ziel-Polypeptid-modifizierenden Aktivität und einer Aktivität, welche die Transkription einer Ziel-DNA moduliert; und
 - (ii) ein auf DNA-gerichtetes RNA-Einzelmolekül, umfassend:
 - (a) ein auf DNA-gerichtetes Segment, das eine Nukleotidsequenz umfasst, welche komplementär zu einer Sequenz in einer Ziel-DNA ist, und
 - (b) ein Protein-bindendes Segment, das mit einem Cas9-Polypeptid interagiert, wobei das Protein-bindende Segment zwei komplementäre Abschnitte von Nukleotiden umfasst, die so hybridisieren, dass ein doppelsträngiger RNA(dsRNA)-Duplex gebildet wird, wobei der dsRNA-Duplex komplementäre Nukleotide einer tracrRNA und einer CRISPR-RNA (crRNA) umfasst, und wobei die beiden komplementären Abschnitte kovalent durch zwischenliegende Nukleotide miteinander verbunden sind.
9. Zusammensetzung nach Anspruch 8, wobei die heterologe Polypeptidsequenz eine Aktivität bereitstellt, welche die Transkription einer Ziel-DNA moduliert.

10. Zusammensetzung nach einem der Ansprüche 8 und 9, wobei die heterologe Polypeptid-Sequenz eine Aktivität bereitstellt, welche die Transkription einer Ziel-DNA erhöht.
11. Zusammensetzung nach einem der Ansprüche 8 und 9, wobei die heterologe Polypeptid-Sequenz eine Aktivität bereitstellt, welche die Transkription einer Ziel-DNA verringert.
12. Zusammensetzung nach Anspruch 8, wobei die heterologe Polypeptidsequenz eine Ziel-DNA-modifizierende Aktivität bereitstellt, die ausgewählt ist aus: einer Nuklease-Aktivität, einer Methyltransferase-Aktivität, einer Demethylase-Aktivität, einer DNA-Reparatur-Aktivität, einer DNA-Schädigungs-Aktivität, einer Deaminierungs-Aktivität, einer Dismutase-Aktivität, einer Alkylierungs-Aktivität, einer Depurinierungs-Aktivität, einer Oxidations-Aktivität, einer Pyrimidin-Dimer-bildenden Aktivität, einer Integrase-Aktivität, einer Transposase-Aktivität, einer Rekombinase-Aktivität, einer Polymerase-Aktivität, einer Ligase-Aktivität, einer Helicase-Aktivität, einer Photolyase-Aktivität und einer Glykosylase-Aktivität.
13. Zusammensetzung nach Anspruch 8, wobei die heterologe Polypeptidsequenz eine Ziel-Polypeptid-modifizierende Aktivität bereitstellt, die ausgewählt ist aus: einer Methyltransferase-Aktivität, einer Demethylase-Aktivität, einer Acetyltransferase-Aktivität, einer Deacetylase-Aktivität, einer Kinase-Aktivität, einer Phosphatase-Aktivität, einer Ubiquitin-Ligase-Aktivität, einer deubiquitinierenden Aktivität, einer Adenylierungs-Aktivität, einer Deadenylierungs-Aktivität, einer SUMOylierenden Aktivität, einer deSUMOylierenden Aktivität, einer Ribosylierungs-Aktivität, einer Deribosylierungs-Aktivität, einer Myristoylierungs-Aktivität und einer Demyristoylierungs-Aktivität.
14. Zusammensetzung nach einem der Ansprüche 8–13, wobei die Cas9-Polypeptidsequenz eine oder mehrere Aminosäure-Änderungen relativ zum entsprechenden Wildtyp-Cas9-Polypeptid aufweist, wobei die Aminosäure-Änderungen die Nuklease-Aktivität der Cas9-Polypeptidsequenz relativ zum entsprechenden Wildtyp-Cas9-Polypeptid verringern.
15. Zusammensetzung nach Anspruch 14, wobei die Cas9-Polypeptidsequenz im Wesentlichen keine Nuklease-Aktivität umfasst.
16. Zusammensetzung nach einem der Ansprüche 8–15, wobei das chimäre Cas9-Polypeptid kovalent an eine Protein-Transduktionsdomäne gebunden ist.
17. Zusammensetzung nach Anspruch 16, wobei die Protein-Transduktionsdomäne kovalent an den Amino-Terminus des chimären Cas9-Polypeptids gebunden ist.
18. Zusammensetzung nach Anspruch 16, wobei die Protein-Transduktionsdomäne kovalent an den Carboxy-Terminus des chimären Cas9-Polypeptids gebunden ist.
19. Zusammensetzung nach einem der Ansprüche 16–18, wobei die Protein-Transduktionsdomäne in der Lage ist, den Übergang vom Cytosol in das Innere einer Organelle einer Zelle zu ermöglichen.
20. Zusammensetzung nach einem der Ansprüche 1–19, wobei die Zusammensetzung ein pharmazeutisches Vehikel umfasst.
21. Zusammensetzung nach Anspruch 20, wobei die Zusammensetzung ein pharmazeutisches Hilfsmittel umfasst.
22. Zusammensetzung nach einem der Ansprüche 1–19 zur Verwendung in einem therapeutischen Verfahren.
23. Verwendung der Zusammensetzung nach einem der Ansprüche 1–19 in einem therapeutischen Verfahren.

Es folgen 128 Seiten Zeichnungen

Anhängende Zeichnungen

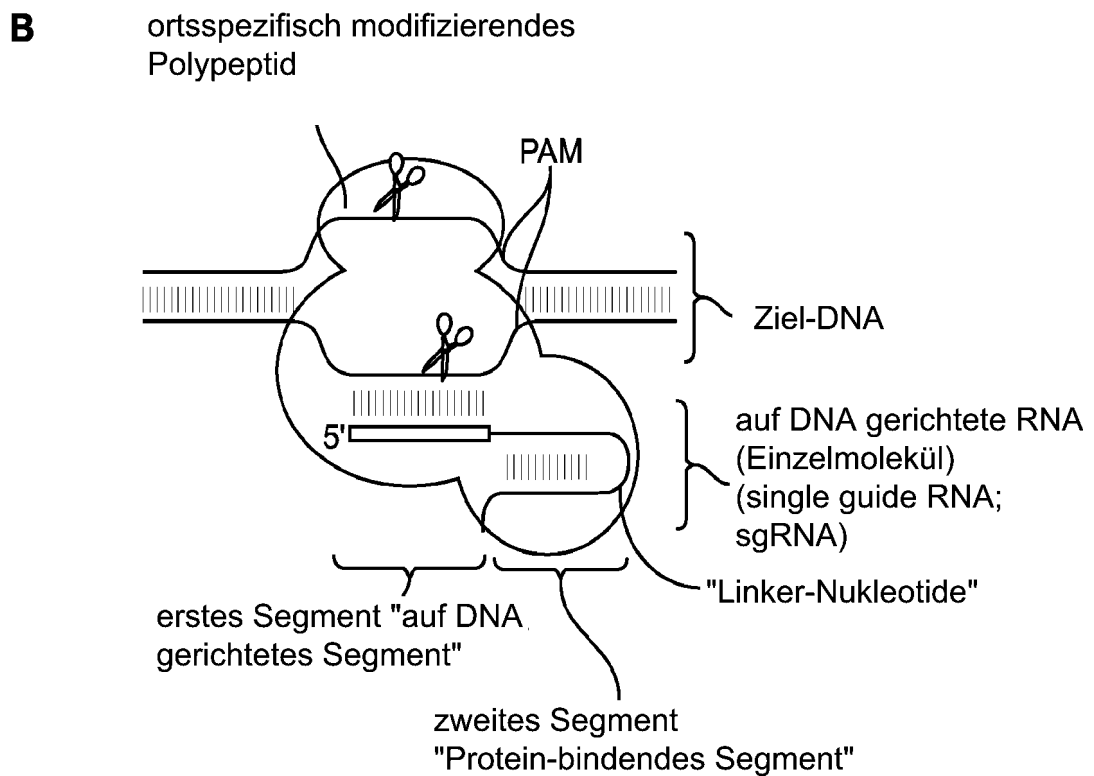
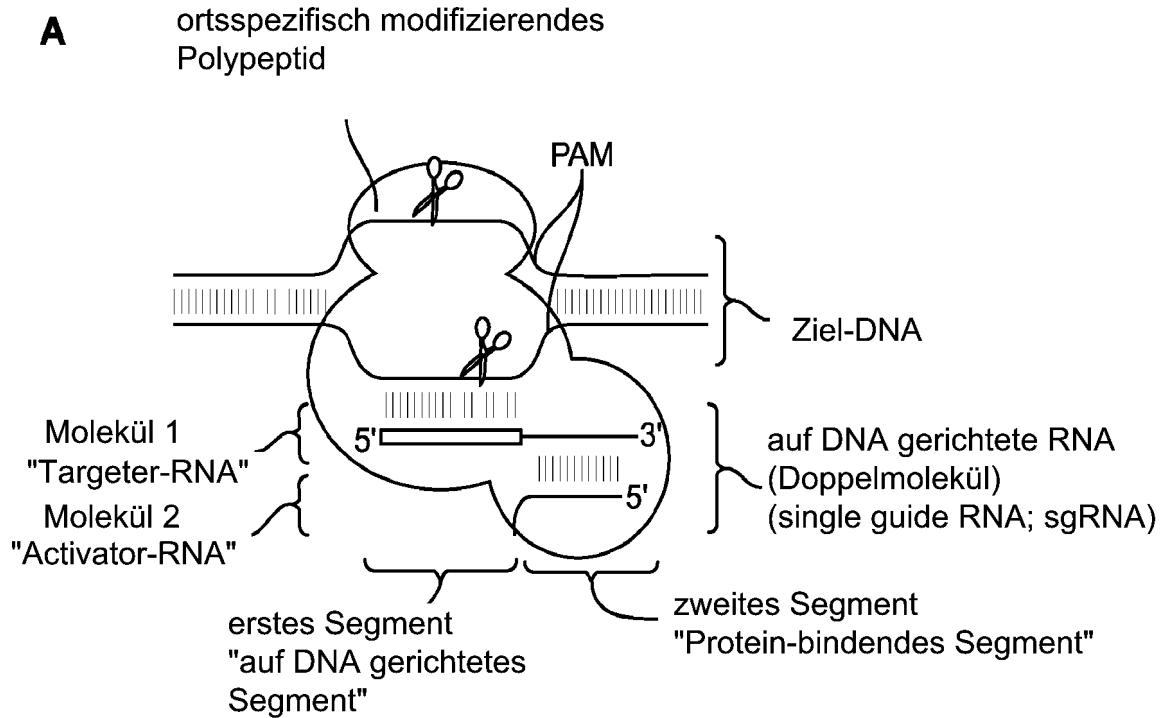


FIG. 1

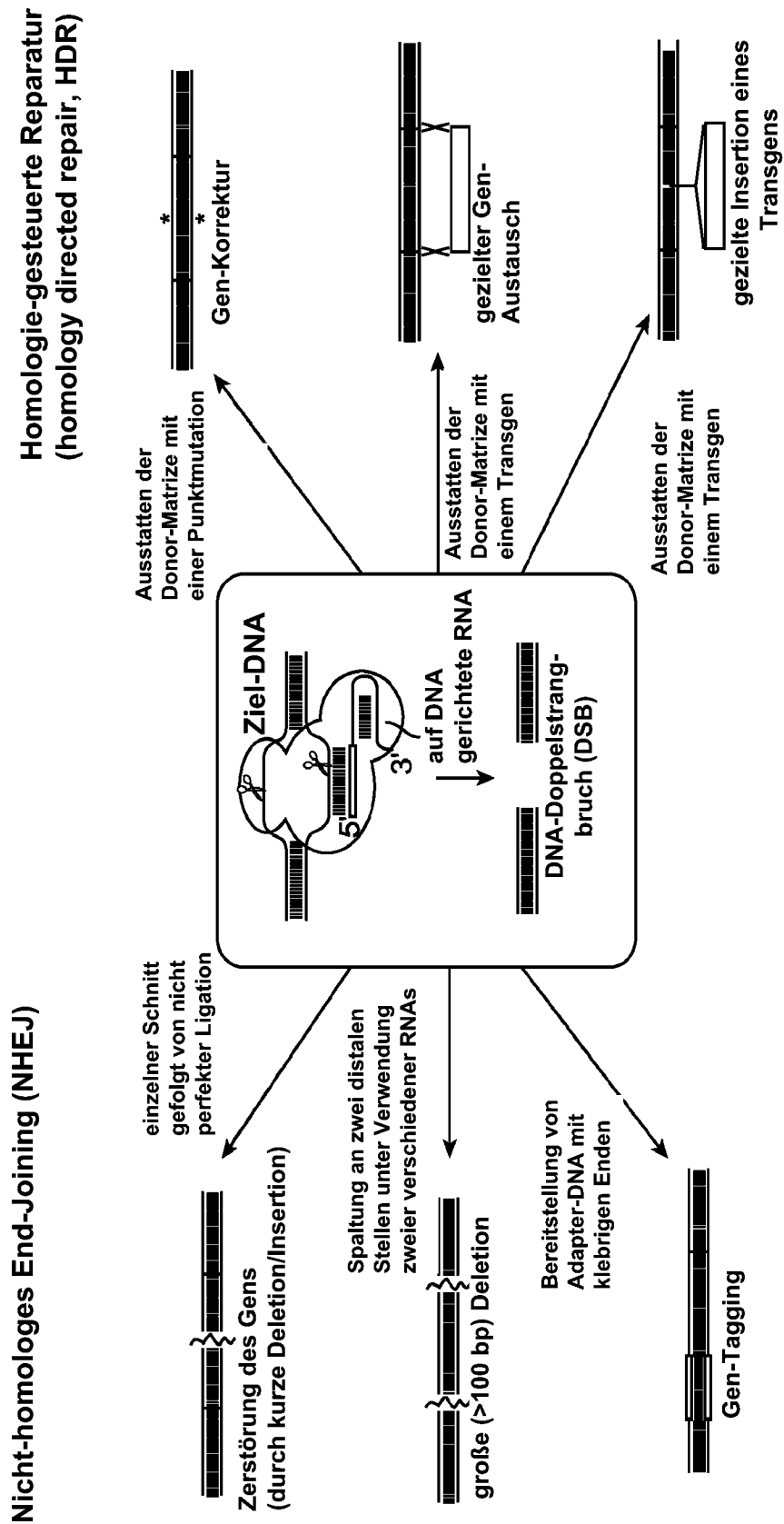


FIG. 2

Cas9/Csn1 Streptococcus pyogenes

Motive

*
 1 MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKLKGLGNTDRHGIKKNLIGALL
 FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
 SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLADSTDKVDLRLIYL
 ALAHMIKFRGHFLIEGDLNPDNSVDKLFQILVQTYNQLFEENPINASRVDA
 KAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
 AKLQLSKD TYDDDLNLLAQIGDQYADFLAAKNLS DATLLSDILRVNSEITK
 APLSASMIKRYDEHHQDLTLLKALVRQQLP EKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLAKLNREDLLRKQRTFDNGSIPYQIHLGEL
 HAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE
 TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYN
 ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC
 FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI
 LDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 2 PAIKKGILQTVKVVDELVKVMGRHKPENIVIE MARENQTTQKGQKNSRERM
 KRIE EG I K E L G S D I L K E Y P V E N T Q L Q N E K L Y L Y L Q N G R D M Y V D Q E L D I N R L
 *
 3 SDYDV D H I V P Q S F L K D D S I D N K V L T R S D K N R G K S D N V P S E E V V K K M K N Y W
 R Q L L N A K L I T Q R K F D N L T K A E R G G L S E L D K V G F I K R Q L V E T R Q I T K H V A Q I L D
 4 S R M N T K Y D E N D K L I R E V R V I T L K S K L V S D F R K D F Q F Y K V R E I N N Y H H A H D A Y
 L N A V V G T A L I K K Y P K L E S E F V Y G D Y K V Y D V R K M I A K S E Q E I G K A T A K Y F F Y S
 N I M N F F K T E I T L A N G E I R K R P L I E T N G E T G E I V W D K G R D F A T V R K V L S M P Q V
 N I V K K T E V Q T G G F S K E S I L P K R N S D K L I A R K K D W D P K K Y G G F D S P T V A Y S V L
 V V A K V E K G K S K K L K S V K E L L G I T I M E R S S F E K D P I D F L E A K G Y K E V R K D L I I K L
 P K Y S L F E L E N G R K R M L A S A G E L Q K G N E L A L P S K Y V N F L Y L A S H Y E K L K G S P
 E D N E Q K Q L F V E Q H K H Y L D E I I E Q I S E F S K R V I L A D A N L D K V L S A Y N K H R D K P I
 R E Q A E N I I H L F T L T N L G A P A A F K Y F D T T I D R K R Y T S T K E V L D A T L I H Q S I T G L Y
 E T R I D L S Q L G G D

FIG. 3A

Cas9/Csn1 Streptococcus pyogenes

Domänen

1 MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKLKGLGNTDRHGIKKNLIGAL
 LFDSGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL
 ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLADSTDKVDLRLI
 YLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASRV
 DAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLA
 EDAKLQLSKD TYDDLDNLLAQIGDQYADLFLAAKNLSDATLLSDILRVNSEI
 TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYID
 GGASQEEFYKFIKPILEKMDGTEELLAKLNREDLLRKQRTFDNGSIPYQIHL
 GELHAILRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRK
 SEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTV
 YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI
 ECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLF
 EDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA

2 GSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSR
 ERMKRIEEGIKELGSDILKEYPVENTQLQNEKLYLYYLQNGRDMYVDQEL
 DINRLSDYVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKK
 MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKVGFIKRQLVETRQIT
 KHVAQILDSRMNTKYDENDKLIREVRVITLKSCLVSDFRKDFQFYKVIN
 NYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIG
 KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV
 RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF
 DSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKDPIDFLEAKGY
 KEVRKDLIIKLPKYSLEFELNGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL
 SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD
 ATLIHQSIITGLYETRIDLSQLGGD

FIG. 3B

A.

Sequenzidentitäten relativ zu: **S. pyogenes Cas9/Csn1**

Spezies	Referenzsequenz	Sequenzidentitäten - Muskel-Alignment		
		volle Länge	Domäne 1	Domäne 2
		% Identität	% Identität	% Identität
Streptococcus pyogenes M1 GAS	NP_269215	100.0	100.0	100.0
Streptococcus pyogenes MGAS5005	YP_282132.1	99.9	99.4	100.0
Listeria innocua Clip11262	NP_472073	54.3	60.0	64.9
Pasteurella multocida subsp. multocida str. Pm70	NP_246064.1	19.7	29.0	25.9
Streptococcus thermophilus LMD-9 Csn1-A	YP_820832	59.2	75.6	72.4
Streptococcus thermophilus LMD-9 Csn1-B	YP_820161.1	20.6	27.3	26.8
Neisseria meningitidis Z2491	YP_002342100.1	20.2	33.6	28.1
Streptococcus mutans UA159	NP_721764	64.9	78.1	74.1
Streptococcus gordonii str. Challis substr. CH1	YP_001450662.1	19.8	28.2	27.0
Campylobacter jejuni subsp. jejuni NCTC 11168	YP_002344900.1	19.6	30.3	26.3
Treponema denticola ATCC 35405	NP_970941	32.5	47.3	38.8

B.

Sequenzidentitäten relativ zu: **N. meningitidis Cas9/Csn1**

Spezies	Referenzsequenz	Sequenzidentitäten - Muskel-Alignment		
		volle Länge	Domäne 1	Domäne 2
		% Identität	% Identität	% Identität
Streptococcus pyogenes M1 GAS	NP_269215	20.2	33.6	28.1
Streptococcus pyogenes MGAS5005	YP_282132.1	20.3	34.5	28.1
Listeria innocua Clip11262	NP_472073	18.8	33.6	25.9
Pasteurella multocida subsp. multocida str. Pm70	NP_246064.1	64.3	72.1	69.0
Streptococcus thermophilus LMD-9 Csn1-A	YP_820832	19.6	35.3	25.9
Streptococcus thermophilus LMD-9 Csn1-B	YP_820161.1	25.8	35.7	35.1
Neisseria meningitidis Z2491	YP_002342100.1	100.0	100.0	100.0
Streptococcus mutans UA159	NP_721764	19.2	36.1	25.5
Streptococcus gordonii str. Challis substr. CH1	YP_001450662.1	25.3	37.5	35.8
Campylobacter jejuni subsp. jejuni NCTC 11168	YP_002344900.1	34.7	45.0	41.0
Treponema denticola ATCC 35405	NP_970941	18.8	31.5	25.5

FIG. 4

	Motiv 1	Motiv 2	Motiv 4
	*		
<i>S. pyogenes</i>	...IGLDIGTNSVGWAVI...	...IVIEMARE...	...HHAHDAYL...
<i>L. pneumophila</i>	...IGIDLGGKFTGVCLS...	...MMQRLAYE...	...SHAIDATL...
<i>G. proteobacterium</i>	...IAIDLGAFTGVVALY...	...IIIEHIARK...	...SHVVDVAVC...
<i>L. innocua</i>	...IGLDIGTNSVGWAVL...	...IVVEMARE...	...HHAHDAYL...
<i>L. gasseri</i>	...VGLDVGTNSCGWVAM...	...IAIEFTRD...	...HHAIDAYL...
<i>E. rectale</i>	...LALDIGIASVGWAIL...	...IVIEMPRD...	...HHAVDAML...
<i>S. lugdunensis</i>	...LGLDIGITSVGYGLI...	...IIIEELARE...	...HHAEDALI...
<i>M. synoviae</i>	...IGFDLGVASVGWSIV...	...VVIEMARE...	...HHAVDASI...
<i>M. mobile</i>	...LGLDLGIASVGWCLT...	...IVVEVTRS...	...HHAEDAYF...
<i>W. succinogenes</i>	...LGVDLGISSLGWAIV...	...VHFELARE...	...HHAVDAIL...
<i>F. columnare</i>	...LGLDLGTNSIGWAIR...	...IHIEMARE...	...HHTIDAIT...
<i>F. succinogenes</i>	...LGLDLGTNSIGWAVV...	...IHLELGRD...	...HHAMDIV...
<i>B. fragilis</i>	...LGLDLGTNSIGWALV...	...IRVELARE...	...HHAMDALT...
<i>A. cellulolyticus</i>	...LGVDVGERSIGLAIV...	...IVVELARG...	...HHAVDVAV...
<i>B. dentium</i>	...IGIDVGLMSVGLAAI...	...VQIEHVRE...	...HHAVDAAV...
	Motiv 3		
	*		
<i>S. pyogenes</i>	...DVDHIVPQSFLKD-----DSIDNKVLTRSDKN...		
<i>L. pneumophila</i>	...EIDHIYPRSLSKKHFGVIFNSEVNLIYCSSLGN...		
<i>G. proteobacterium</i>	...EIDHIIPRSLTGRTKKTVFNSEANLIYCSSLGN...		
<i>L. innocua</i>	...DIDHIVPQSFITD-----NSIDNLVLTSSAGN...		
<i>L. gasseri</i>	...DIDHILPQSFIKD-----DSLENRVLVKKAVN...		
<i>E. rectale</i>	...EIDHIIPRSISFD-----DARSNKVLVYRSEN...		
<i>S. lugdunensis</i>	...EVDHIIPRSVSFD-----NSYHNKVLVKQSEN...		
<i>M. synoviae</i>	...EIDHVIPIYSKSD-----DSWFKLLVKKSTN...		
<i>M. mobile</i>	...DIDHIVPRISFD-----DSFSNLVIVNKLND...		
<i>W. succinogenes</i>	...EIDHILPRSRSD-----DSFANKVLCLARAN...		
<i>F. columnare</i>	...DIEHTIPRSISQD-----NSQMKTLCCLKFN...		
<i>F. succinogenes</i>	...EIEHVIPQSLYFD-----DSFSNKVICEAEVN...		
<i>B. fragilis</i>	...DIEHIIPQARLFD-----DSFSNKTLEARSVN...		
<i>A. cellulolyticus</i>	...ELDHIVPRTDGG-----SNRHENLAITCGACN...		
<i>B. dentium</i>	...EMDHIVPRKGVGS-----TNTRVNLAACAACN...		

FIG. 5

FIG. 6

A

	1				36
<i>L. innocua</i>	(1)	GUUUUAGAGCUAUGUUU	UUUUGAAUGCUAA	CAAAAC	
<i>S. pyogenes</i>	(1)	GUUUUAGAGCUAUGC	UGUUUUUGAAUGGU	CCCAAAAC	
<i>S. mutans</i>	(1)	GUUUUAGAGCUGUGUUUUU	C	GAAUGGUUCCAAAC	
<i>S. thermophilus1</i>	(1)	GUUUUAGAGCUGUGUUUUU	C	GAAUGGUUCCAAAC	

B

	1				37
<i>C. jejuni</i>	(1)	AUUUUACC	AUAAAGAAA	UUUAAAAGGGA	CUAAAC
<i>S. pyogenes</i>	(1)	GUUUUAGA	GCUAUGCUGUUUU	GAUGGUCC	CAAAAC
<i>F. novicida</i>	(1)	GUUUCAGUUGCUGAAU	UAUUUGGUAACUA	CUGUUAG	
<i>M. mobile</i>	(1)	GUUUUGGU	GUAGUAUCAUUU	UAUGUAUU	CUAAAC
<i>N. meningitidis</i>	(1)	GUUGUAGC	UCCUUUUCUCAUU	UCGCAGUGCUA	CAAU
<i>P. multocida</i>	(1)	GUUGUAGU	UCCUUUCUCAUU	UCGCAGUGCUA	CAAU
<i>S. thermophilus2</i>	(1)	GUUUUUGU	ACUCUCAAGA	UUUUAAGUA	ACUGUA
					CAAC

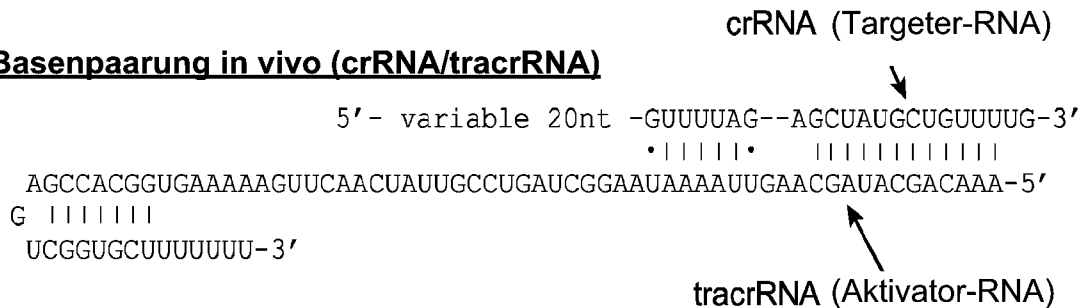
FIG. 7

Stamm	Anzahl von CRISPRs	CASS4 CRISPR Identifikatoren ^a	CRISPR REPEAT:tracrRNA Basenpaarung ^b
<i>Streptococcus pyogenes</i> SF370	2	NC_002737_1	5' GUUUUAG--AGCUAUGCUGUUUUGAAUGGUCCAAAAC 3' • 3' AAAUUGAACGAUACGACAAAACUUACCAAGGUUGUU 5'
<i>Streptococcus mutans</i> UA159	1	NC_004350_1	GUUUUAG--AGCUGUGUUGUUUCGAAUGGUUCCAAAAC • • AAAUUGAACGACACAACAAGCUUACUAAGGUUGUG
<i>Streptococcus thermophilus</i> LMD-9	3	NC_008532_5	GUUUUAG--AGCUGUGUUGUUUCGAAUGGUUCCAAAAC • AAAUUGAGCGACACAACAAGCUUACCAAGUUUGG
		NC_008532_2	GUUUUUGUACUCU-CAAGAUUUAAGUAACUGUACAAC • AAACAUCGAAGACGUUCUAAAUCAUUGACACAUUC
<i>Listeria innocua</i> Clip11262	1	NC_003212_2	GUUUUAG--AGCUAUGUUAUUUUGAAUGCUAACAAAAC • • AAAUUGAACGAUACAUAUAAAACUUUAUGAUUGUUUA
<i>Treponema denticola</i> ATCC35405	1	NC_002967_1	GUUUGAG--AGUUGUGUAUUUAAGAUGGAUCUCAAC • • AACUUGAGCAACACAUAUAAAUUCUACCUAGAAUUUA
<i>Neisseria meningitidis</i> Z2491	2	NC_003116_10	GUUGUAGCUCUUUCUUAUUCGAGUGCUACAAC • UAACAUCGUUGCCAAGAGUAAAGCGUCACGUGUUA
<i>Streptococcus gordonii</i> str. Challis substr. CH1	1	NC_009785_2	GUUUUUGUACUCU-CAAGAUUUAAGUAACUGUACAAC • • AAACAUCGAAGACGUUCUAAAUCAUUGACACAUUC
<i>Bifidobacterium bifidum</i> S17	1	NC_014616_1	GUUUCA-AUGCCUGUCAGAUCAAUGACUUUGACCAC • • AGUUAAAUAACGACAGUCCAGUUAACUGGAACUAGUA
<i>Lactobacillus salivarius</i> UCC118	1	NC_007929_1	GUUUCAGAAGUAUGUUAUAUCAAUAAAGGUUAAGACC • • • AAGUUGAGUCUUACA AUUAGUUAACUCCAGUUUUGG CUAACAGUAGUUUACCAAUAAUUCAGCAACUGAAAC • • • UUGUGUUAUGUAUGGUUUAUUAAGAUUGUUGACUUUG CUAACAGUAGUUUACCAAUAAUUA-GCAACUGAAAC • UUUAAUUAUUAUGGUUUAUUAUUAACG-AGACAUUA
<i>Francisella tularensis</i> subsp. novicida U112	2	NC_008601_1 ^c	CCAAUAAUCCCUCAUCUAAAAUCCA-ACCACUGAAAC AUUUAUUCUUAGUAGAUUUAAAGCUAUGG-GACUUUA
<i>Legionella pneumophila</i> str. Paris	1	NC_006368_1	CCAAUAAUCCCUCAUCUAAAAUCCA-ACCACUGAAAC AUUUAUUCUUAGUAGAUUUAAAGCUAUGG-GACUUUA

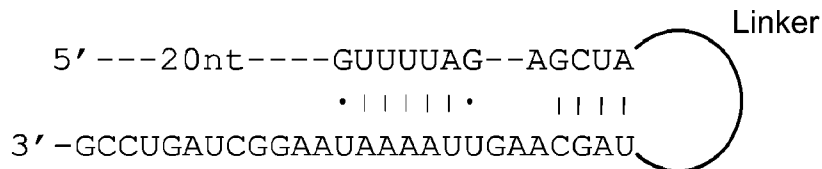
FIG. 8

Streptococcus pyogenes

Basenpaarung in vivo (crRNA/tracrRNA)

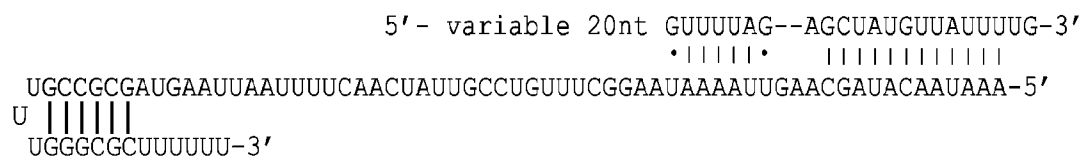


Beispiel für ein auf DNA gerichtetes RNA-Einzelmolekül



Listeria innocua

Basenpaarung in vivo (crRNA/tracrRNA)



Beispiel für ein auf DNA gerichtetes RNA-Einzelmolekül

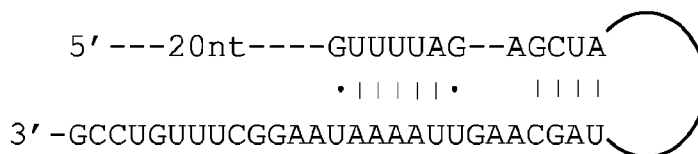


FIG. 9

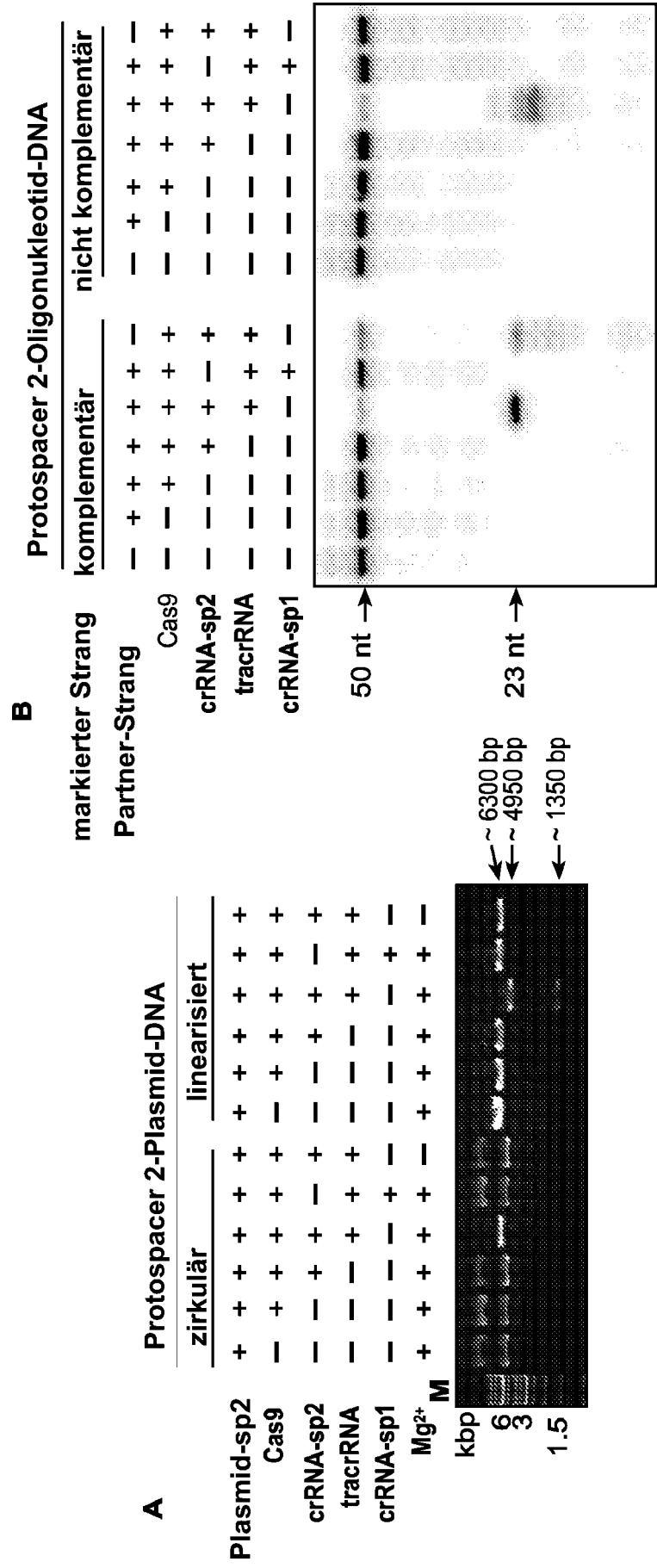
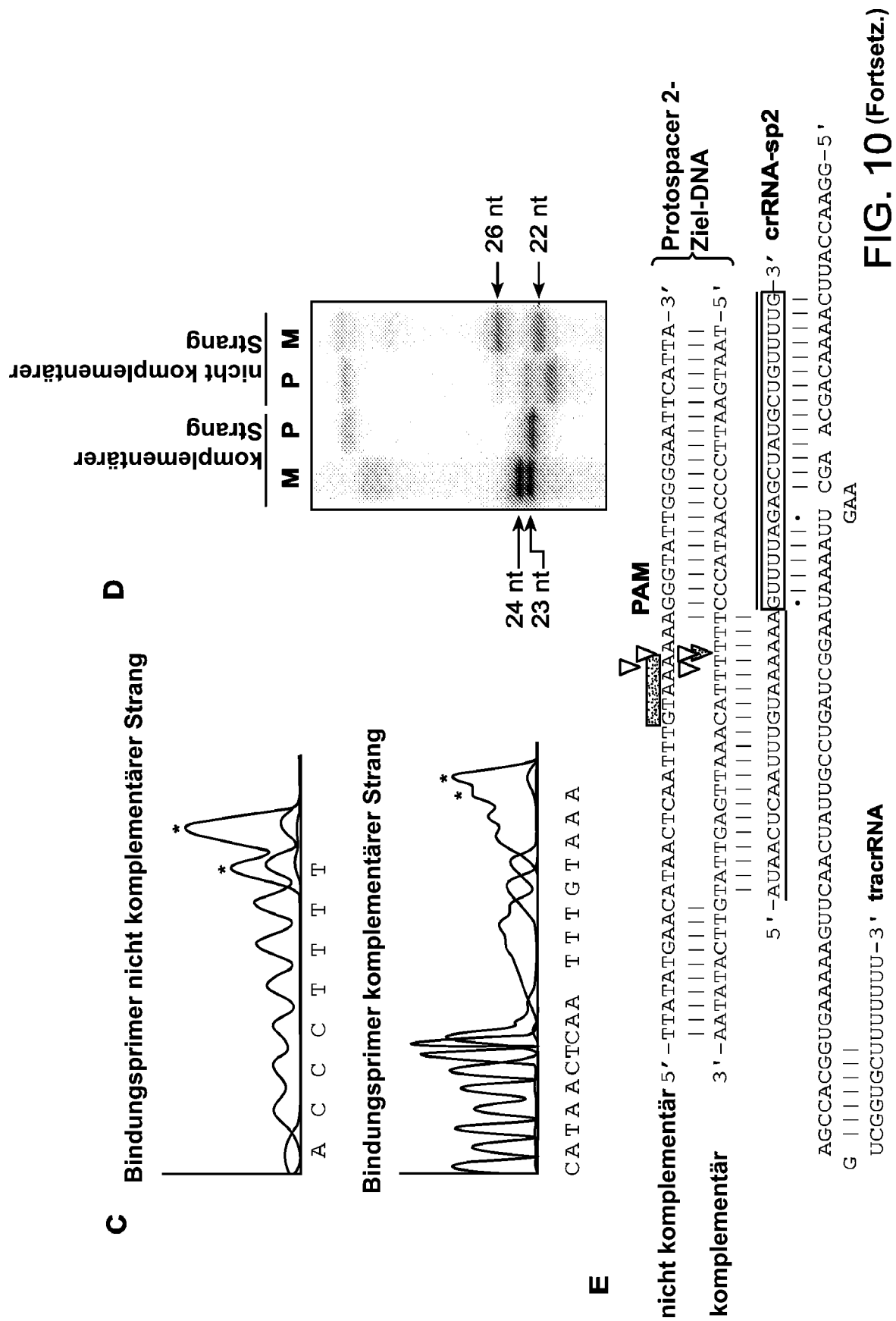


FIG. 10



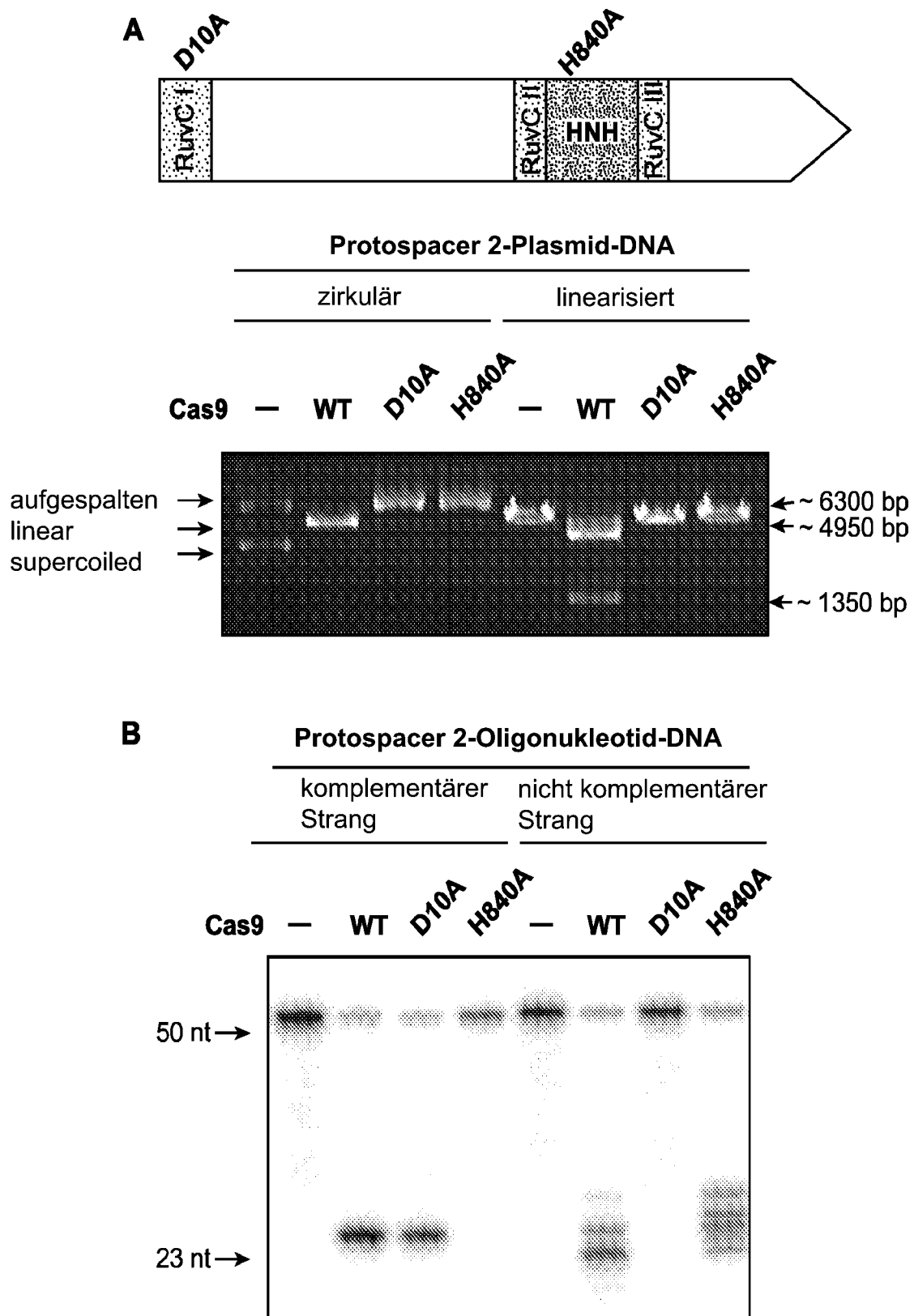


FIG. 11

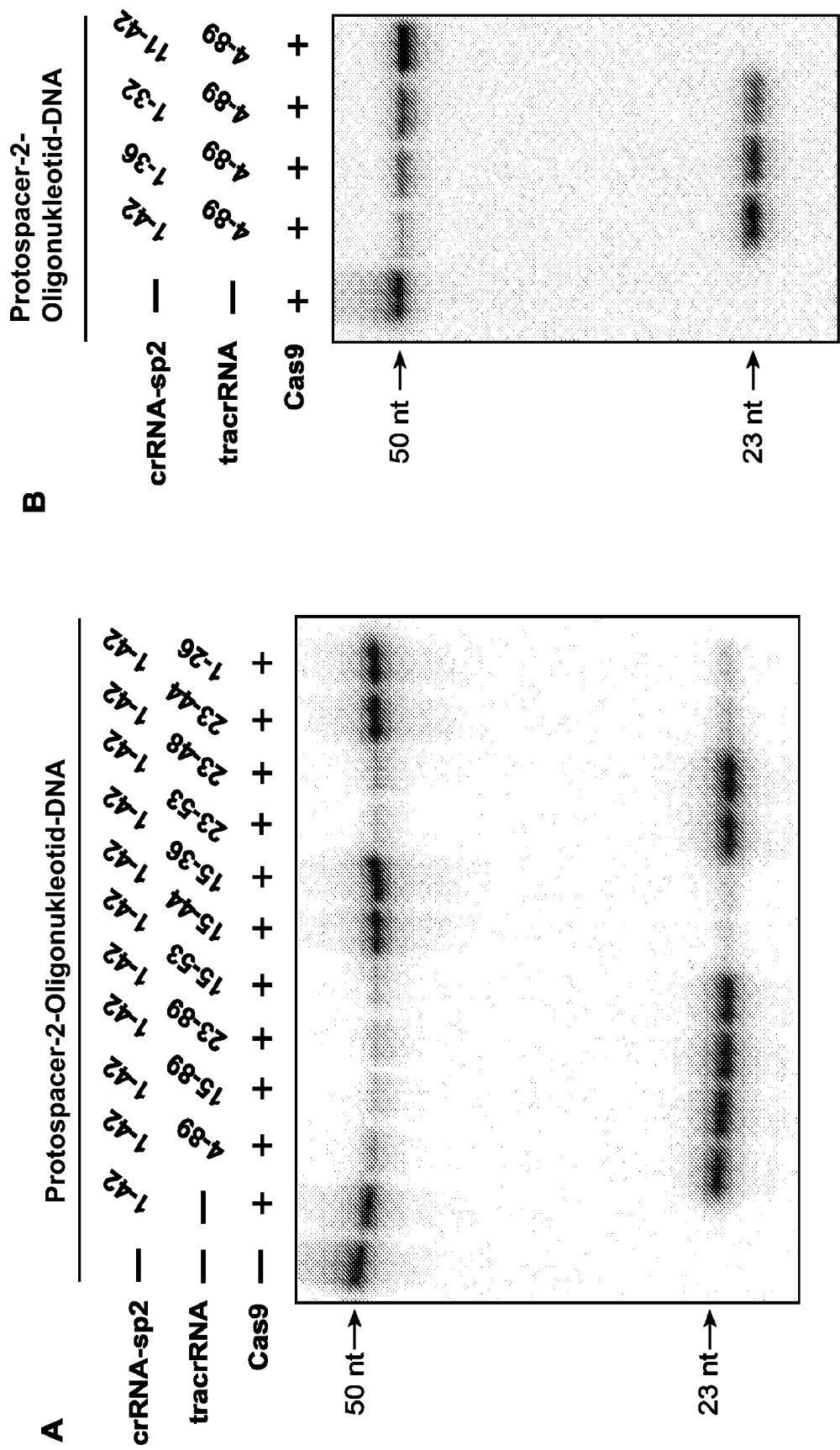


FIG. 12

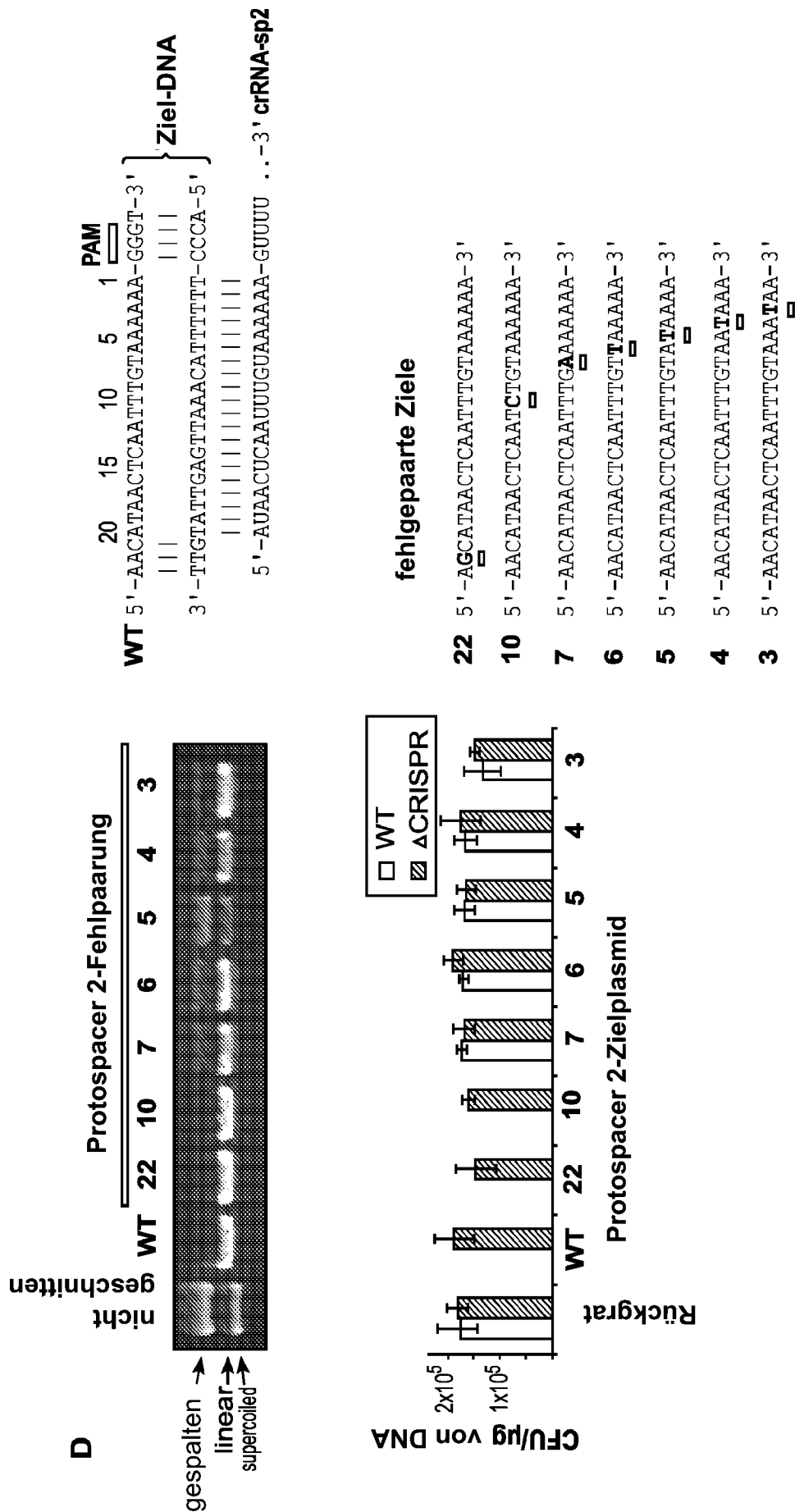
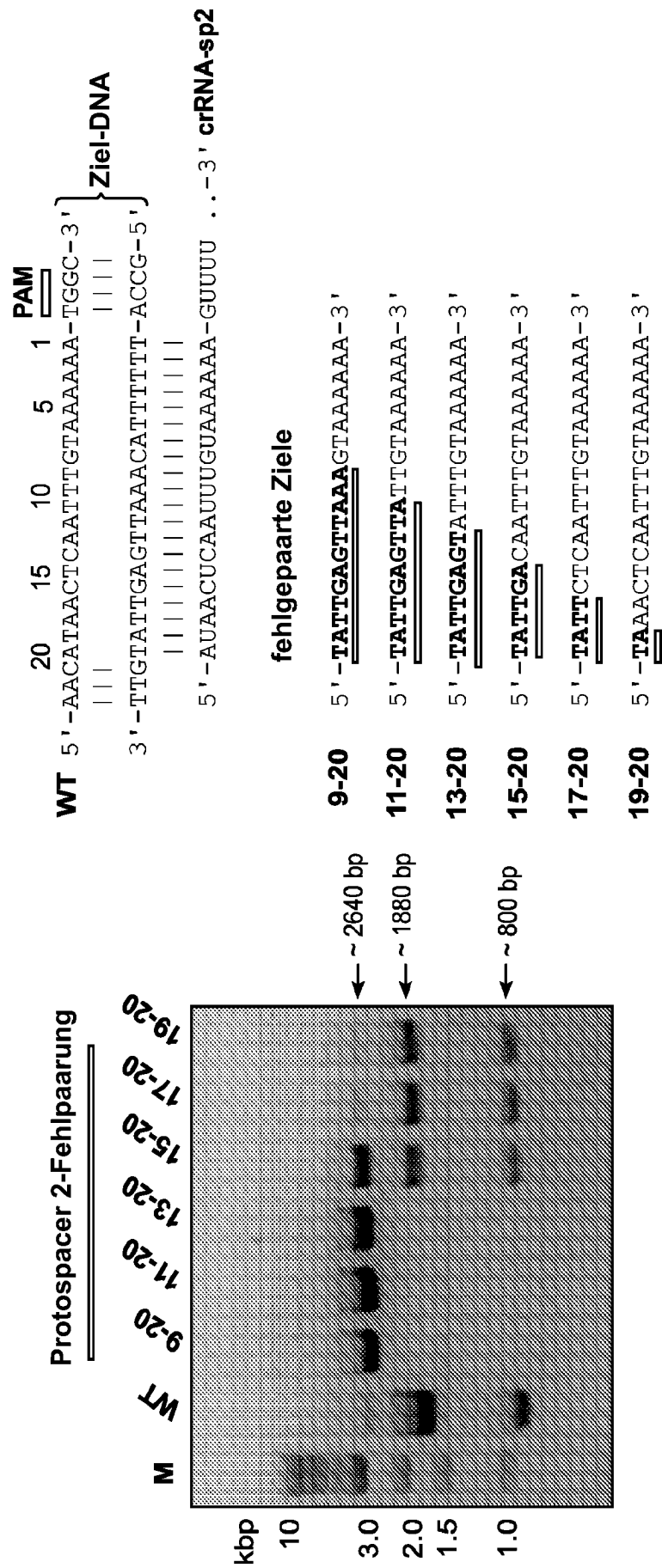


FIG. 12 (Fortsetz. 2)

E**FIG. 12** (Fortsetz. 3)

A

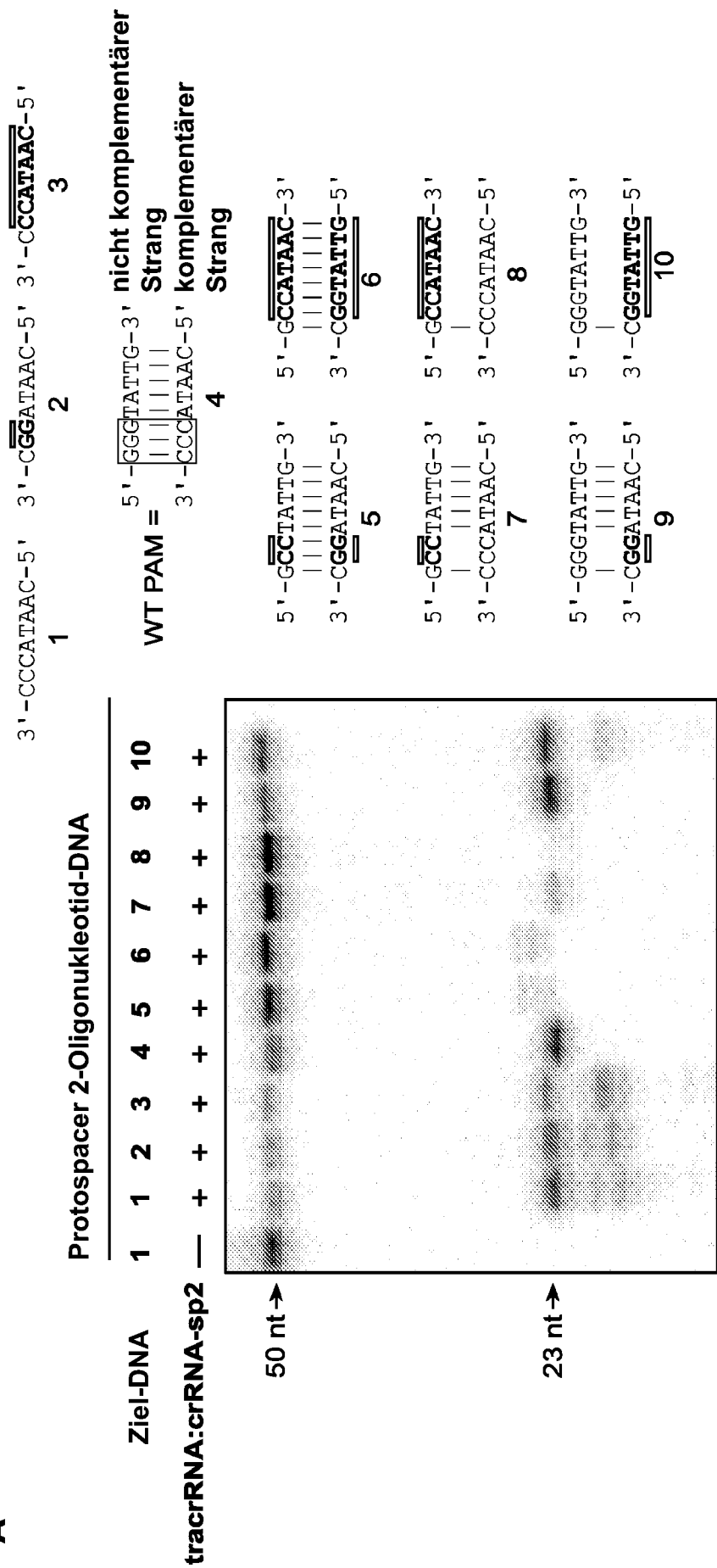


FIG. 13

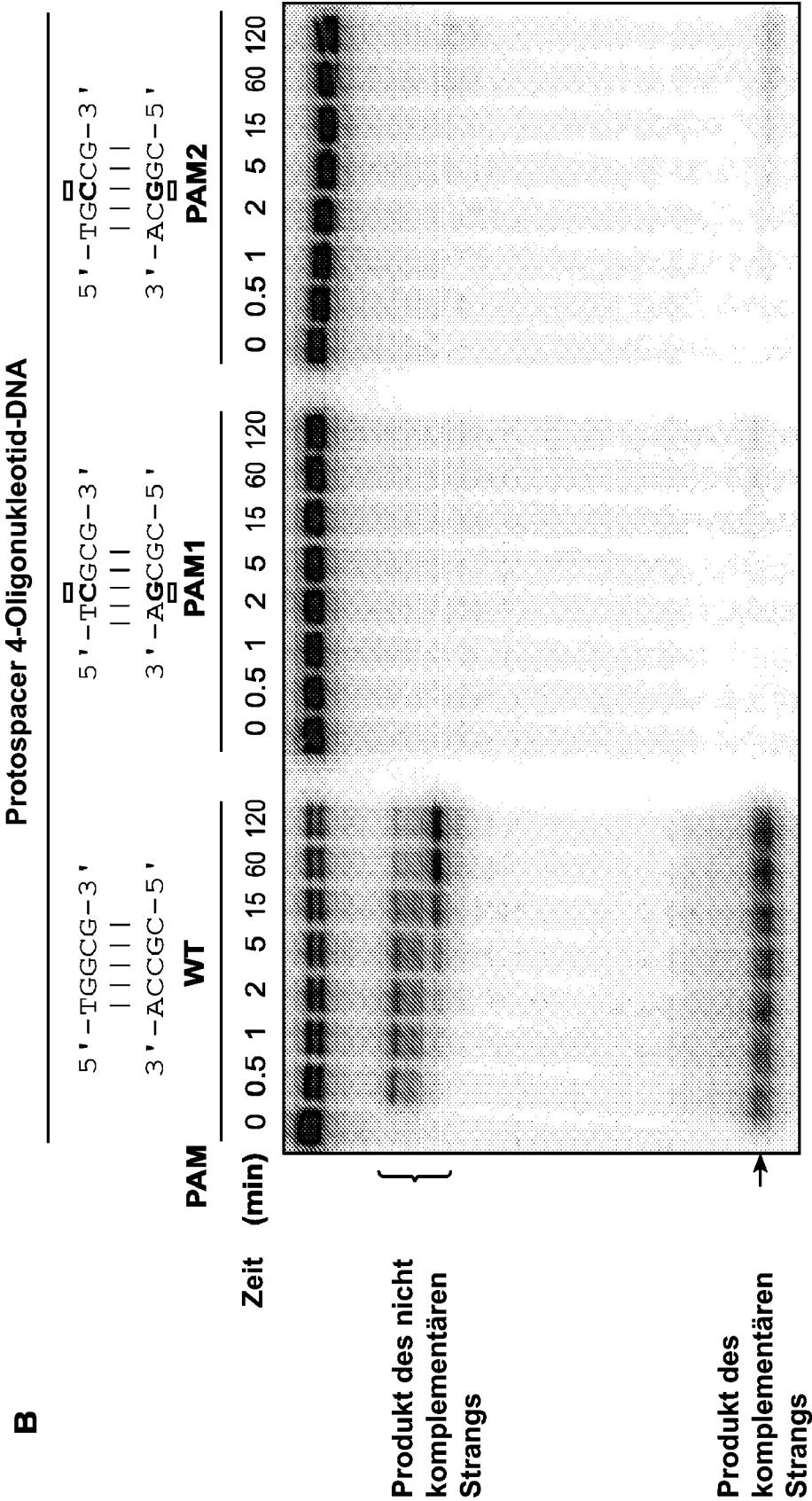


FIG. 13 (Fortsetz. 1)

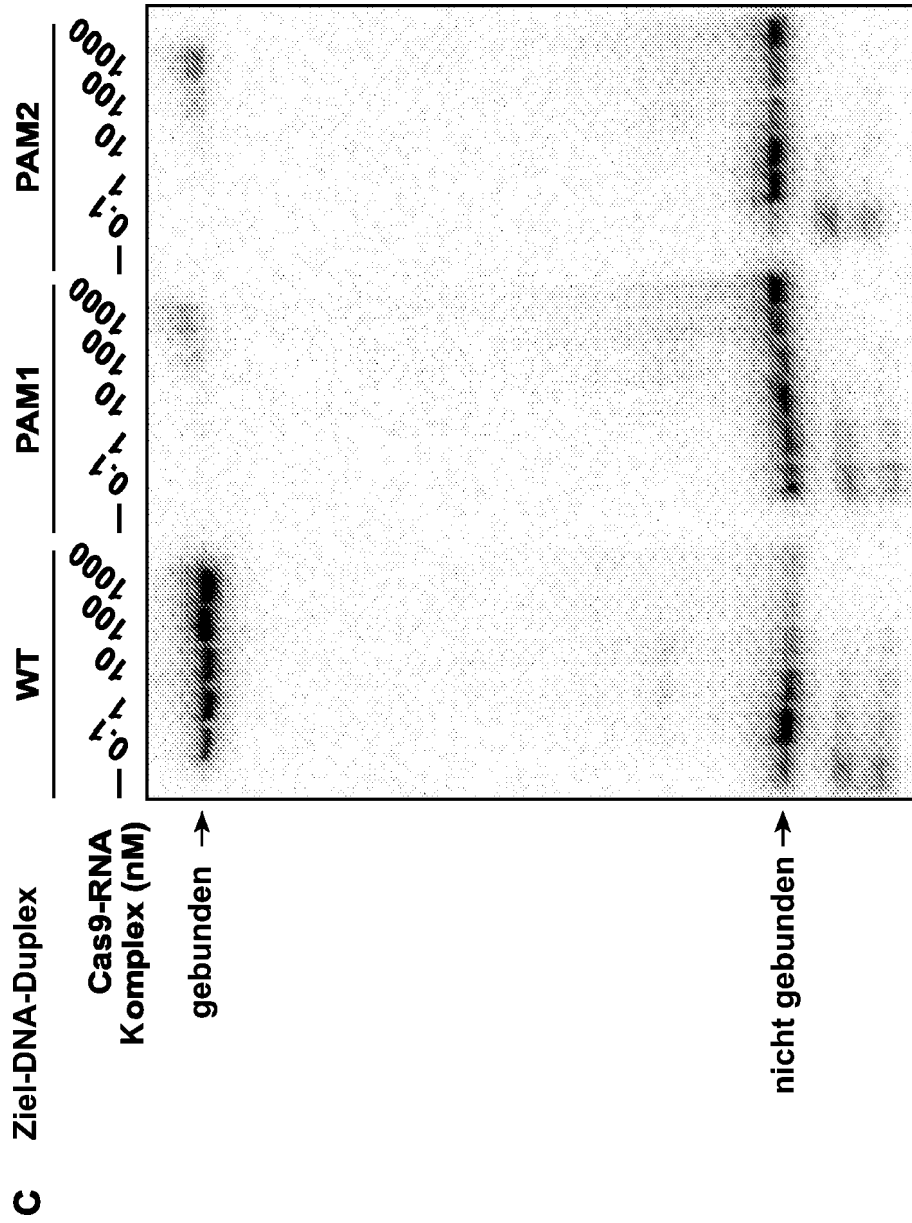


FIG. 13 (Fortsetz. 2)

A

Protospacer 4 Plasmid-DNA

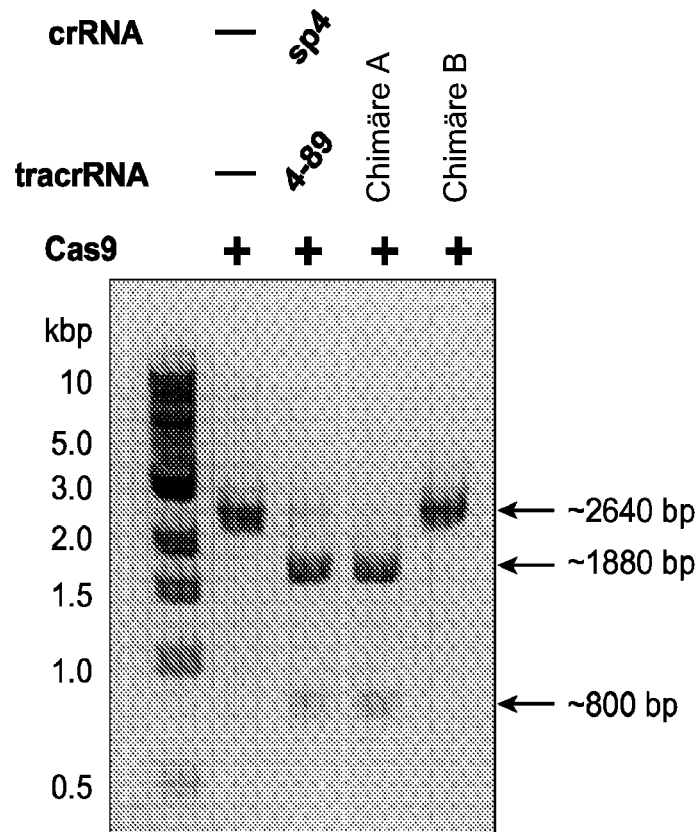
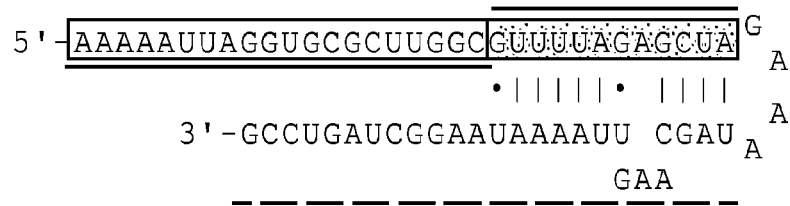
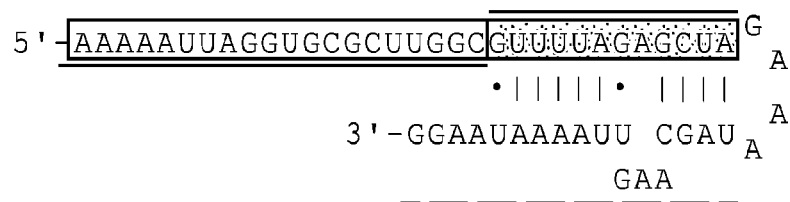
Chimäre **A**Chimäre **B**

FIG. 14

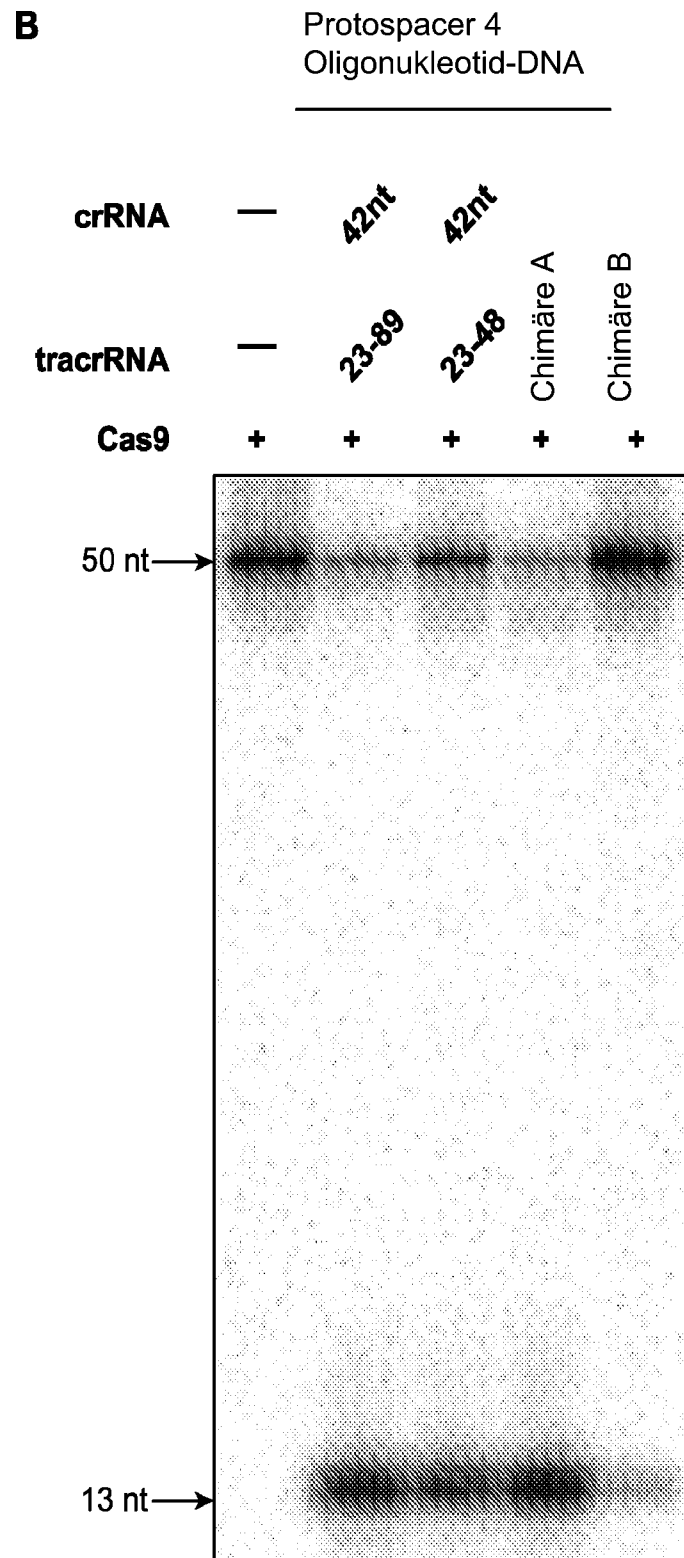


FIG. 14 (Fortsetz. 1)

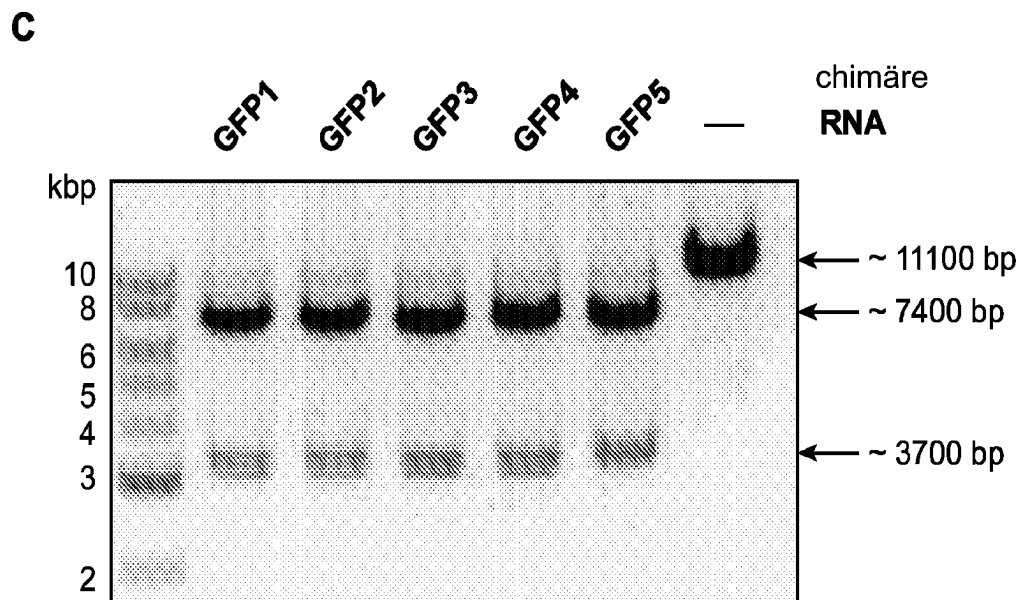


FIG. 14 (Fortsetz. 2)

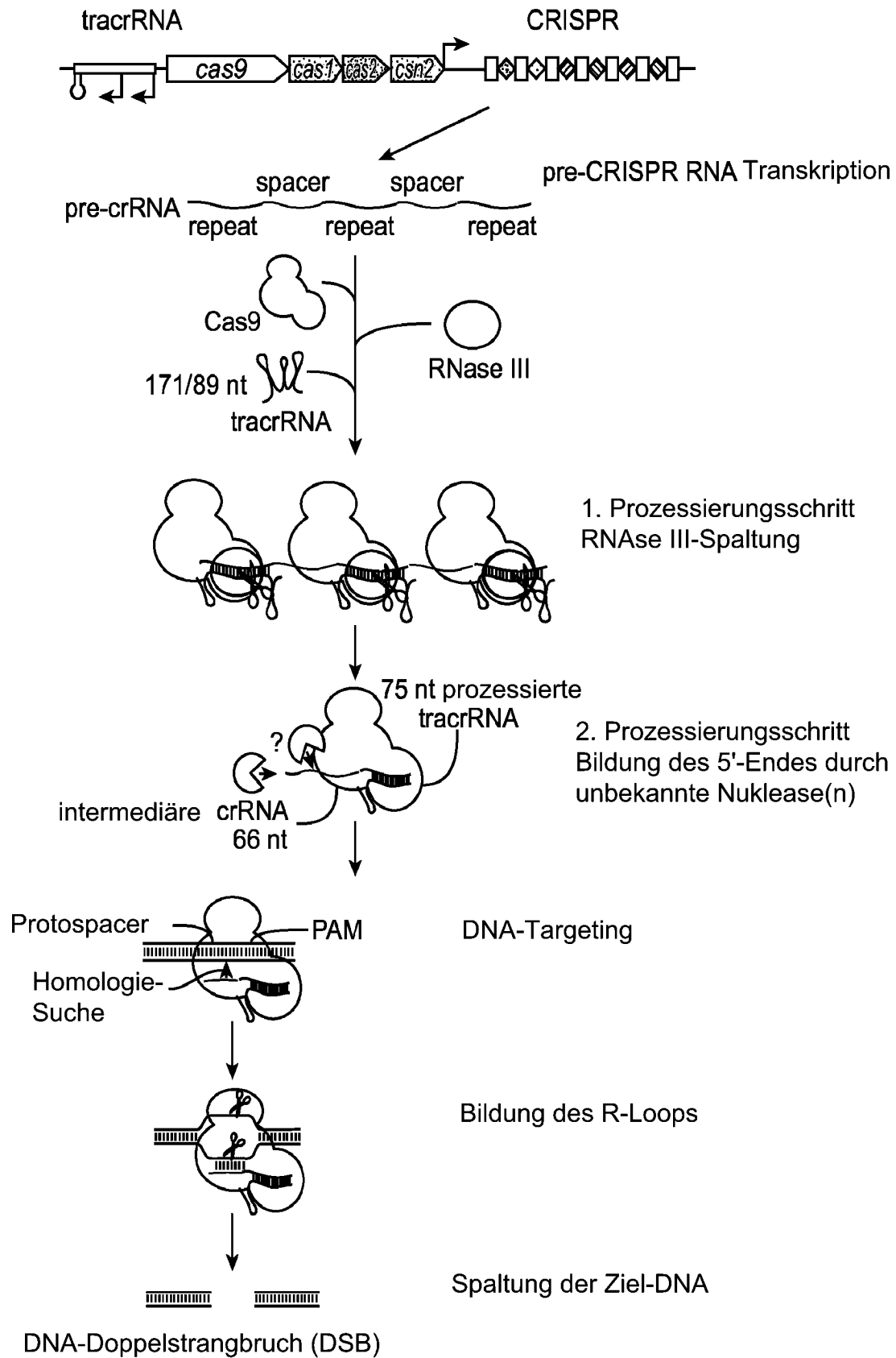


FIG. 15

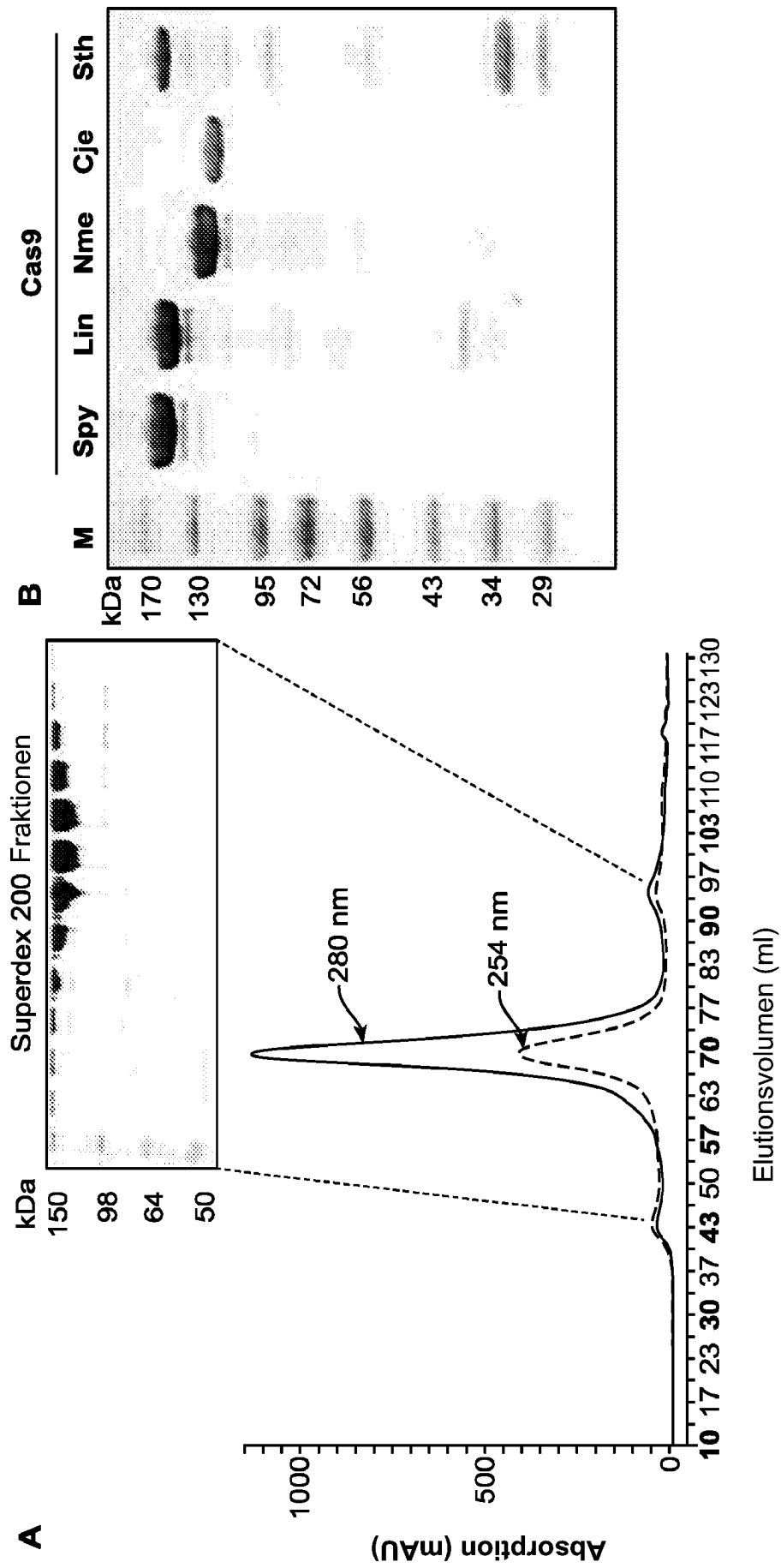


FIG. 16

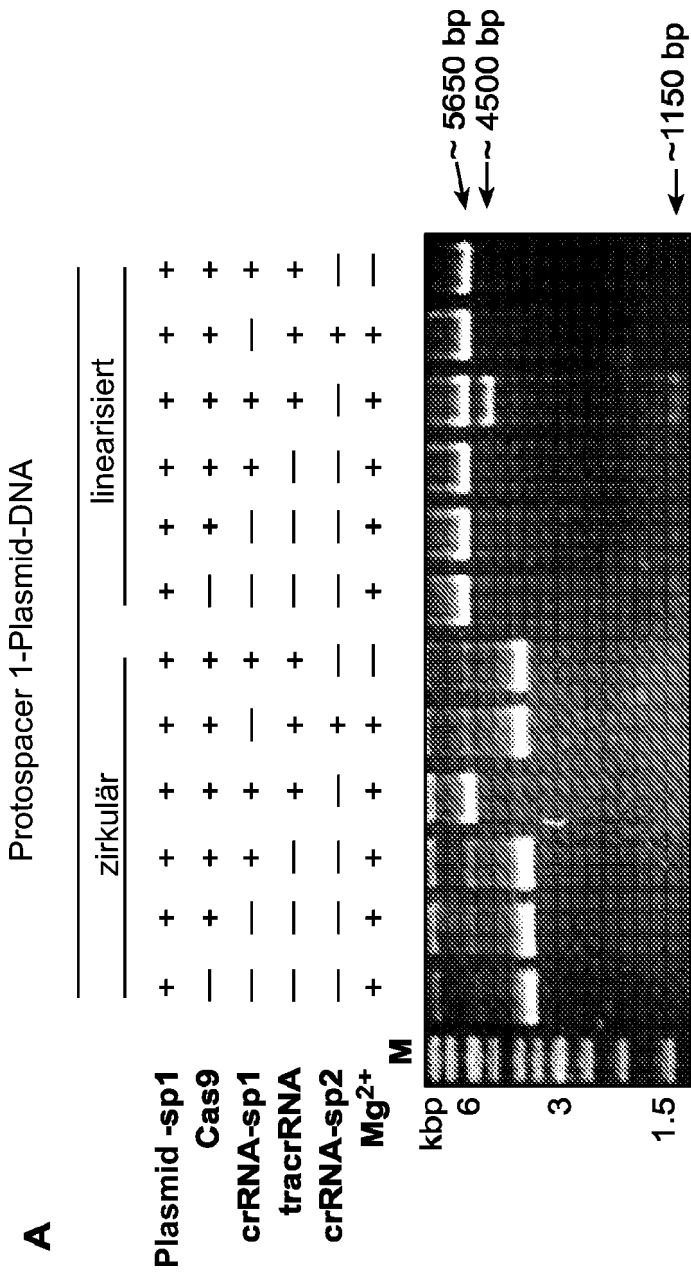
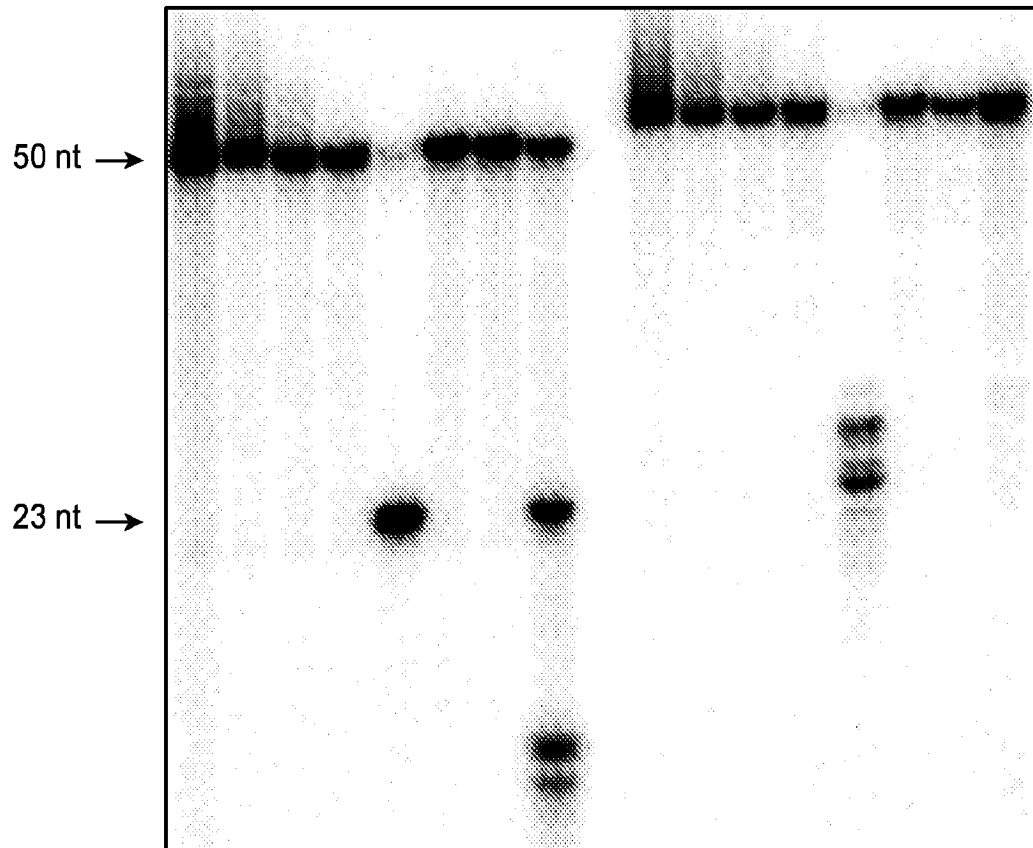


FIG. 17

B**Protospacer 1-Oligonukleotid-DNA**

markierter Strang	komplementär								nicht-komplementär							
Partner-Strang	—	+	+	+	+	+	+	—	—	+	+	+	+	+	+	—
Cas9	—	—	+	+	+	+	+	+	—	—	+	+	+	+	+	+
crRNA-sp1 (42 nt)	—	—	—	+	+	—	—	+	—	—	—	+	+	—	—	+
tracrRNA	—	—	—	—	+	+	+	+	—	—	—	—	+	+	+	+
crRNA-sp2	—	—	—	—	—	+	—	—	—	—	—	—	—	+	—	—
crRNA-sp4	—	—	—	—	—	—	+	—	—	—	—	—	—	—	+	—



nicht-komplementär

PAM

5' -TGCGCTGGTTGATTCTTCTTTCGCTTTTGGGTATTGGGAATTCATTA-3'

|||||

|||||

3' -ACGCGACCAACTAAAGAAGAACGCGAAAAACCCATAACCCCTTAAGTAAT-5'

komplementär

|||||

5' -GAUUUCUUUCUUUGGCUUUUUUUUUUAGAGCUAUGCUGUUUUU-3' crRNA-sp1

• ||||| •

AGCCACGGUGAAAAAGUUCACUAUUGCCUGAUCGGAUAAAAUU CGAUACGACAAAACUUACCAAGGUUG-5'

G |||||

GAA

UCGGUGCUUUUUUUU3'

tracrRNA

FIG. 17 (Fortsetz. 1)

C

markierter Strang	Protospacer 4-Oligonukleotid-DNA													
	komplementär							nicht-komplementär						
Partner-Strang	—	+	+	+	+	+	+	—	—	+	+	+	+	+
Cas9	—	—	+	+	+	+	+	—	—	+	+	+	+	+
crRNA-sp4 (42 nt)	—	—	—	+	+	—	—	+	—	—	—	+	+	—
tracrRNA	—	—	—	—	+	+	+	+	—	—	—	—	+	+
crRNA-sp2	—	—	—	—	—	+	—	—	—	—	—	—	+	—
crRNA-sp1	—	—	—	—	—	—	+	—	—	—	—	—	—	+

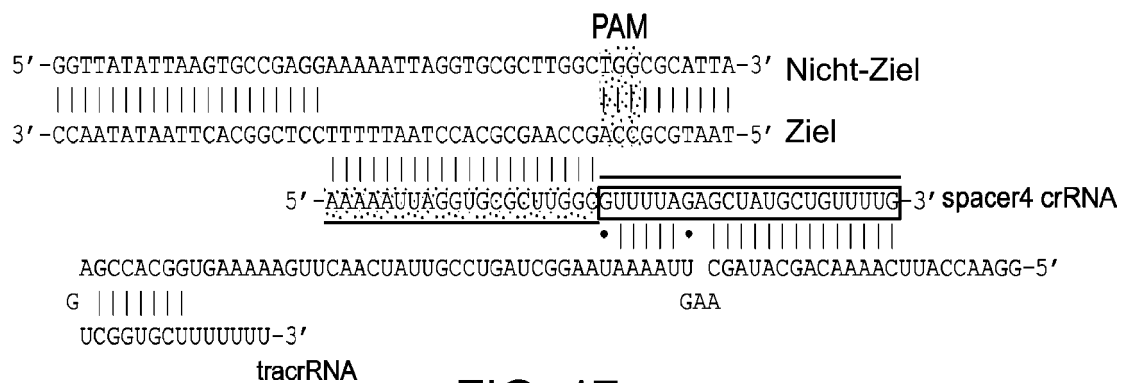
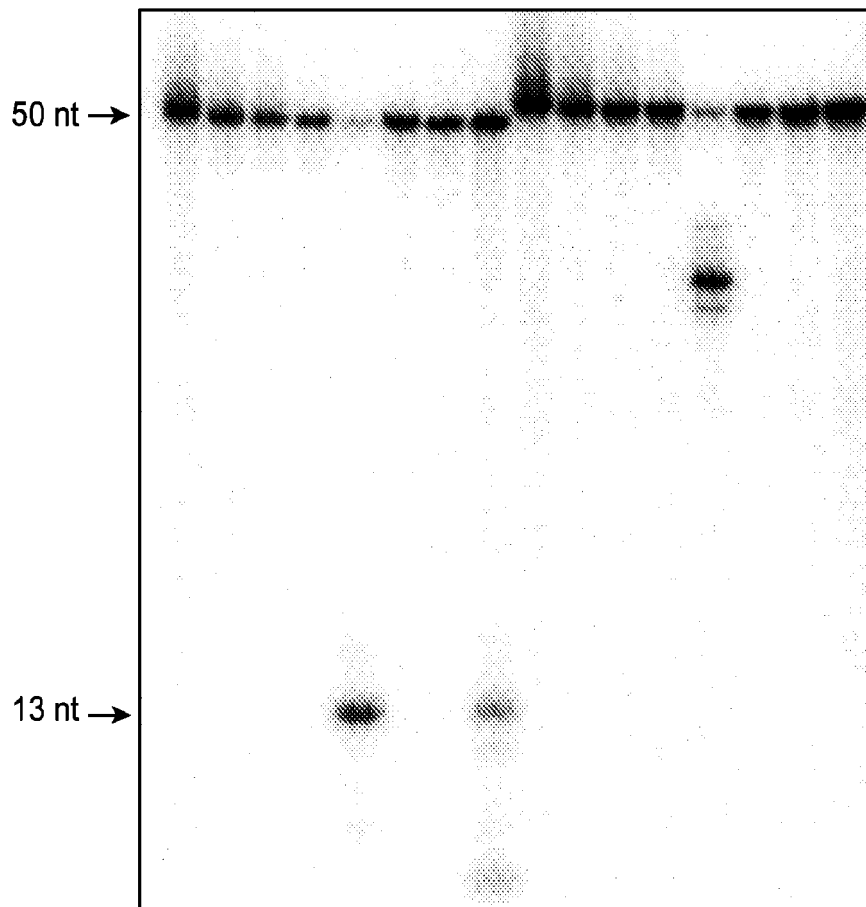


FIG. 17 (Fortsetz. 2)

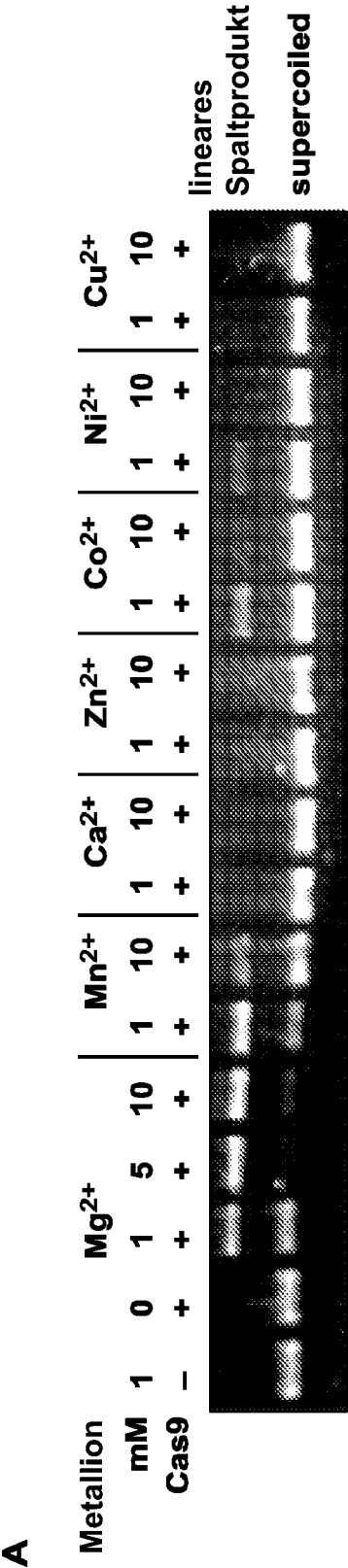


FIG. 18

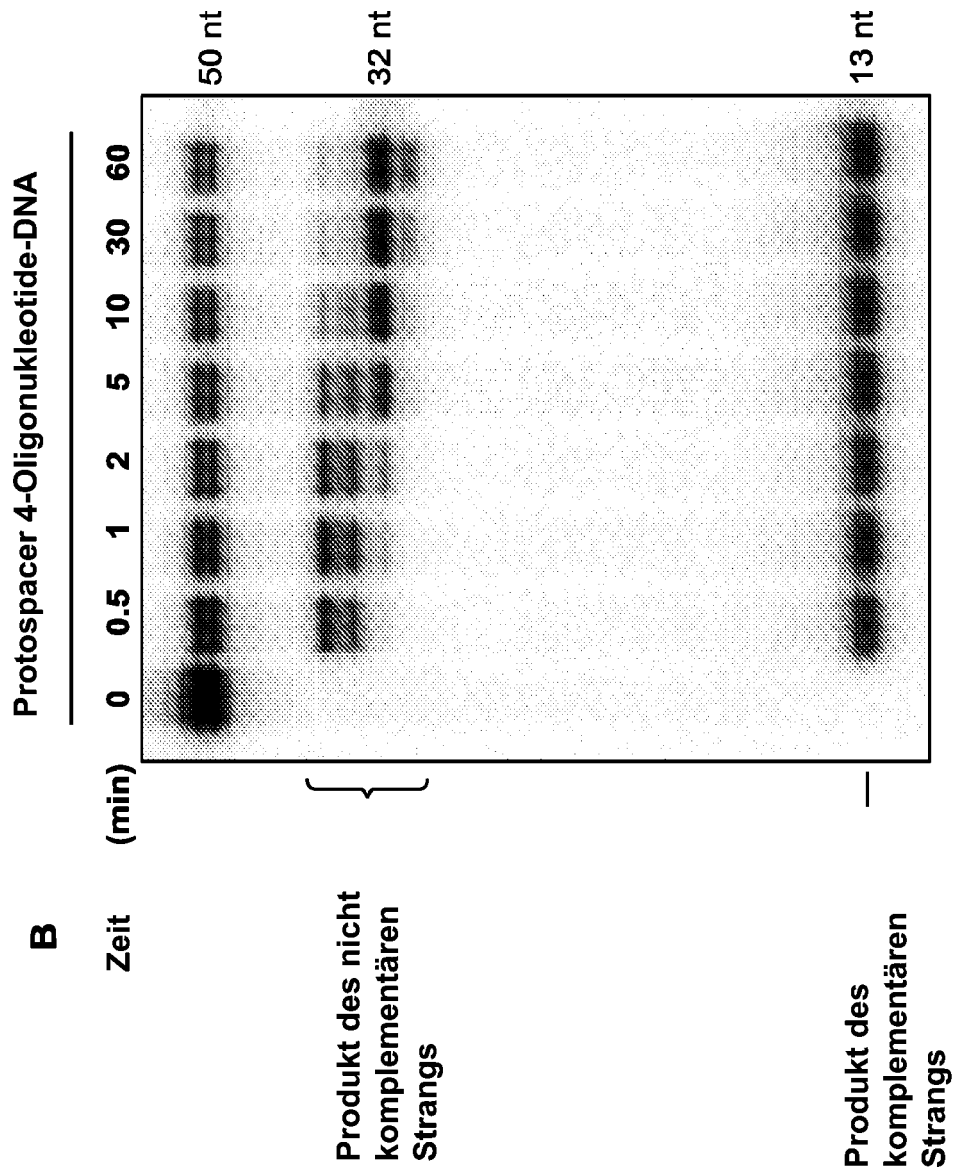


FIG. 18 (Fortsetz.)

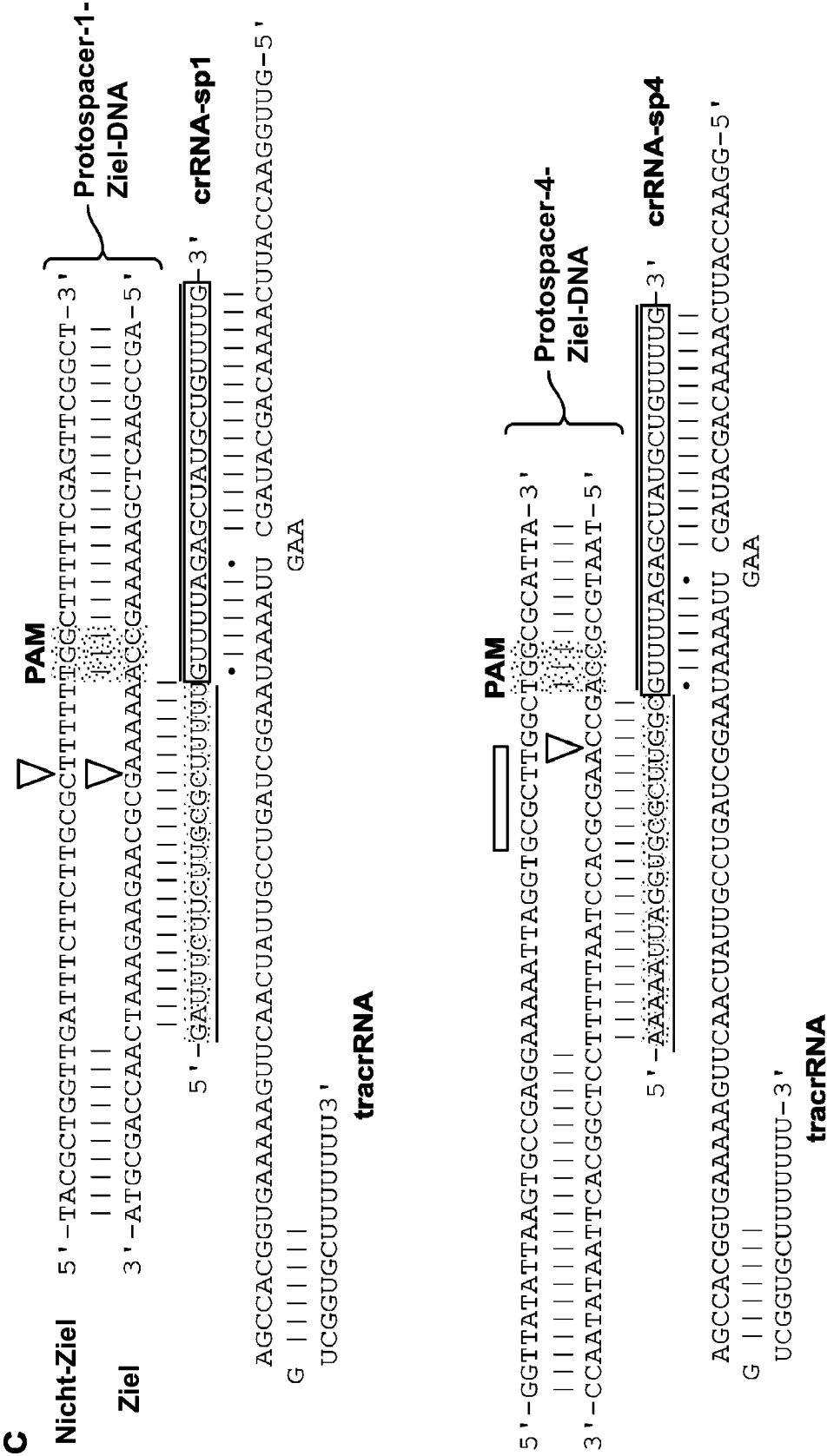
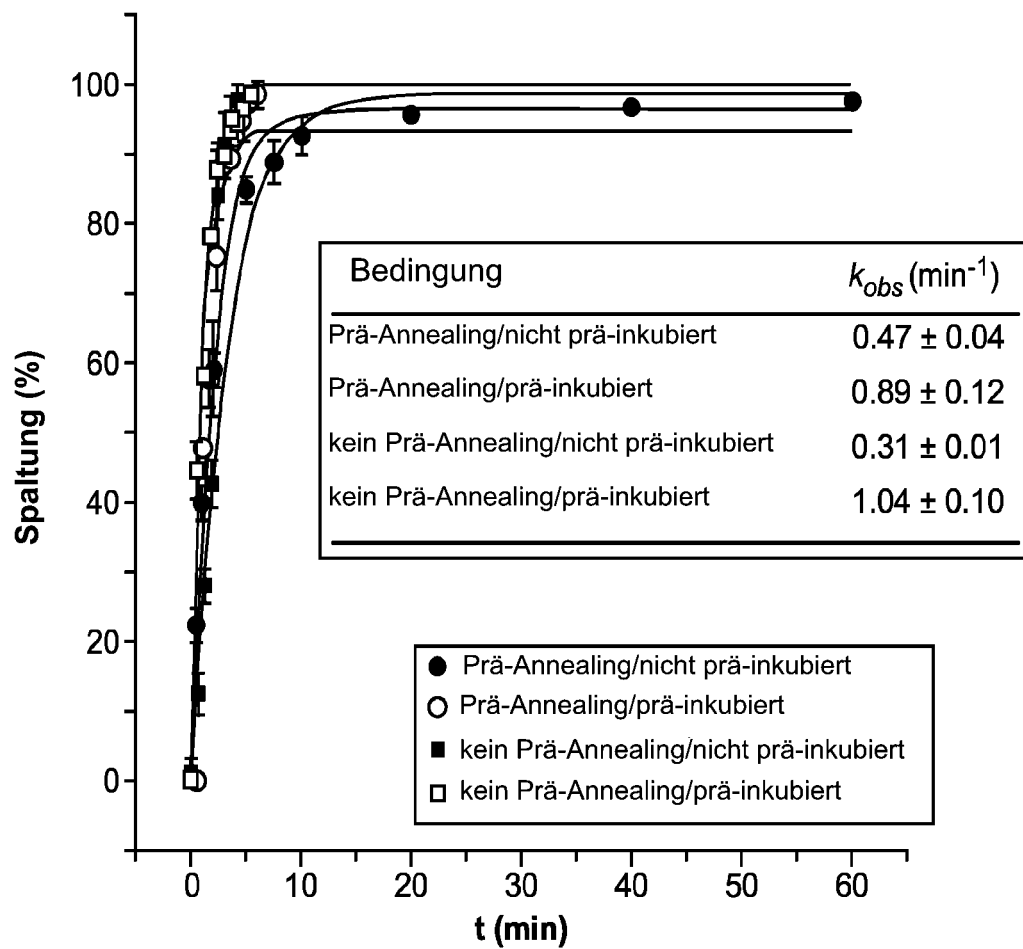
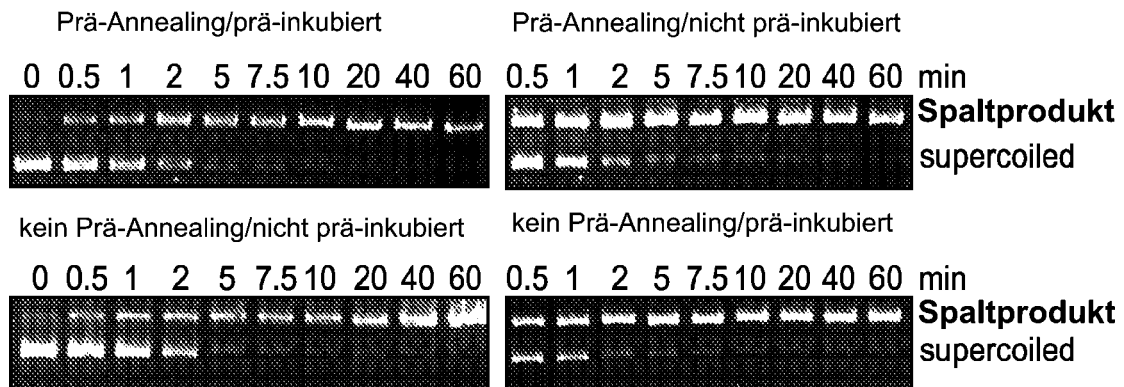
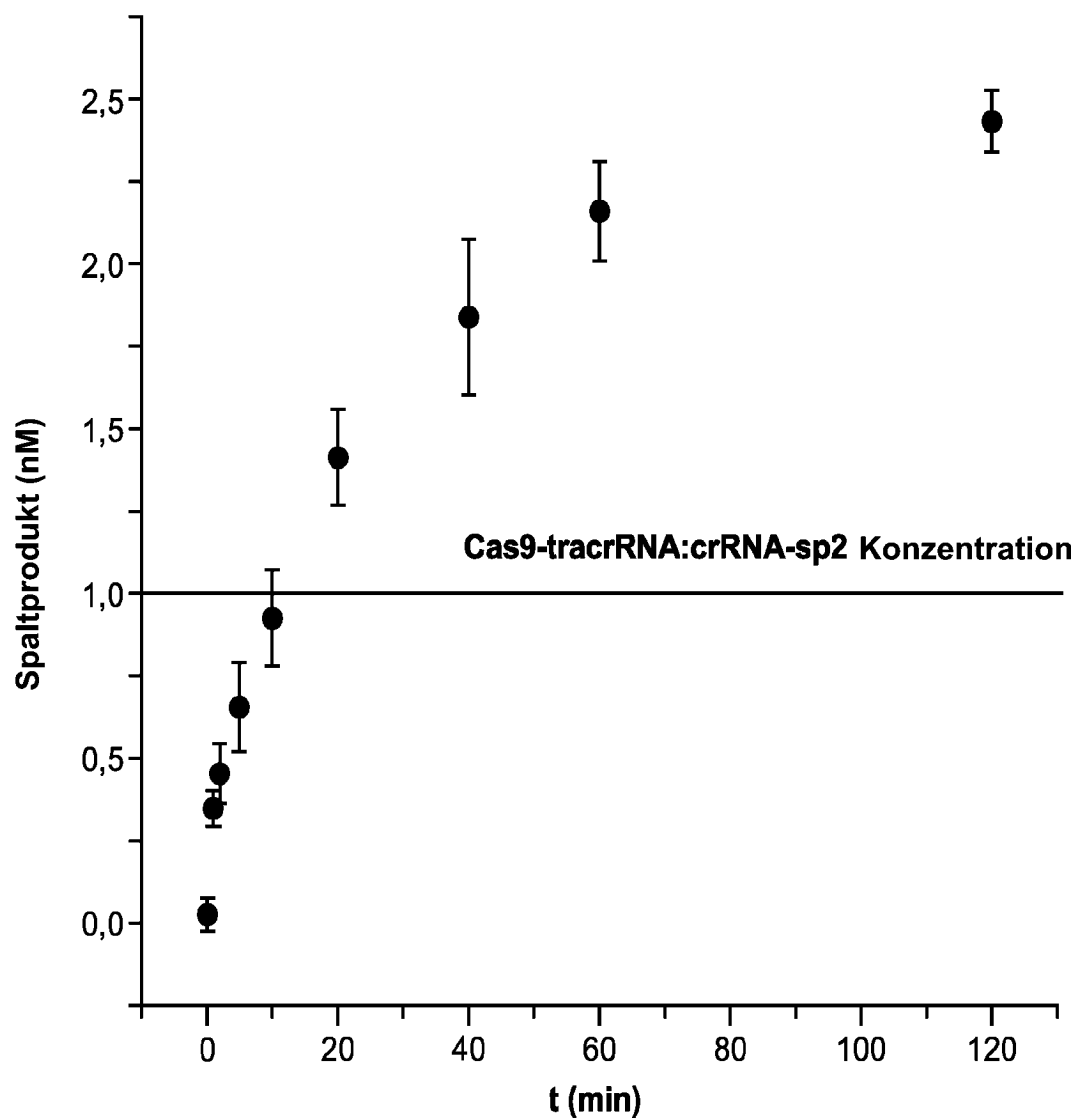
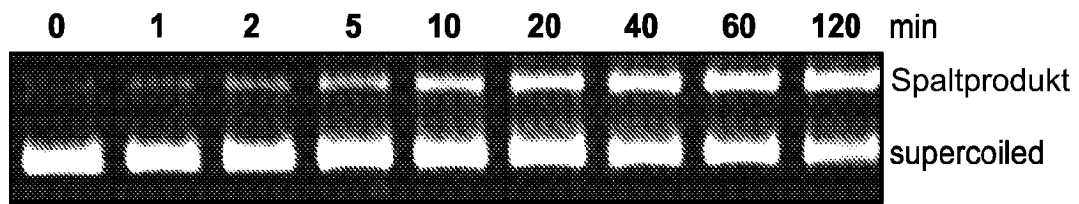


FIG. 19 (Fortsetz.)

A**FIG. 20**

B**FIG. 20** (Fortsetz.)

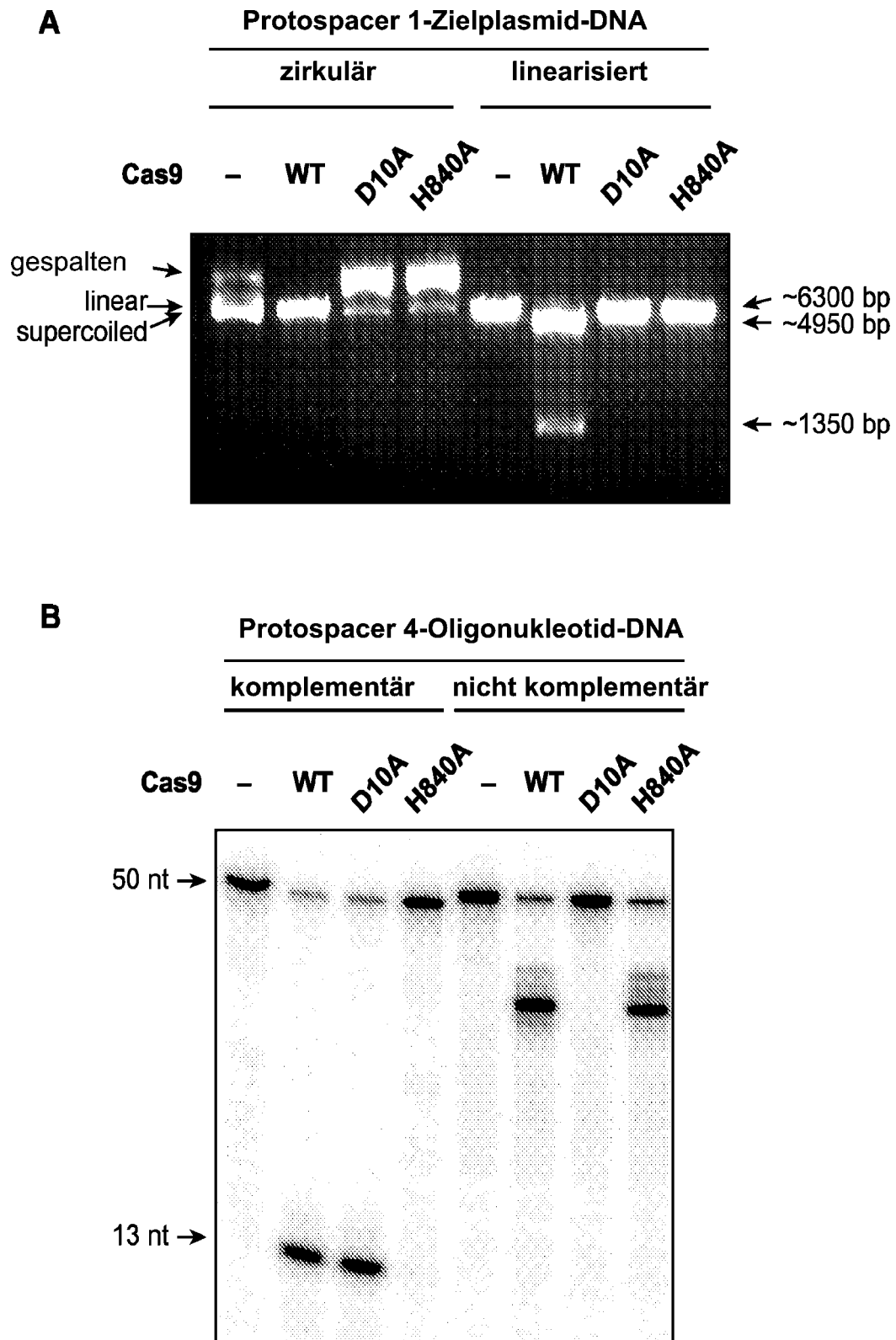


FIG. 21

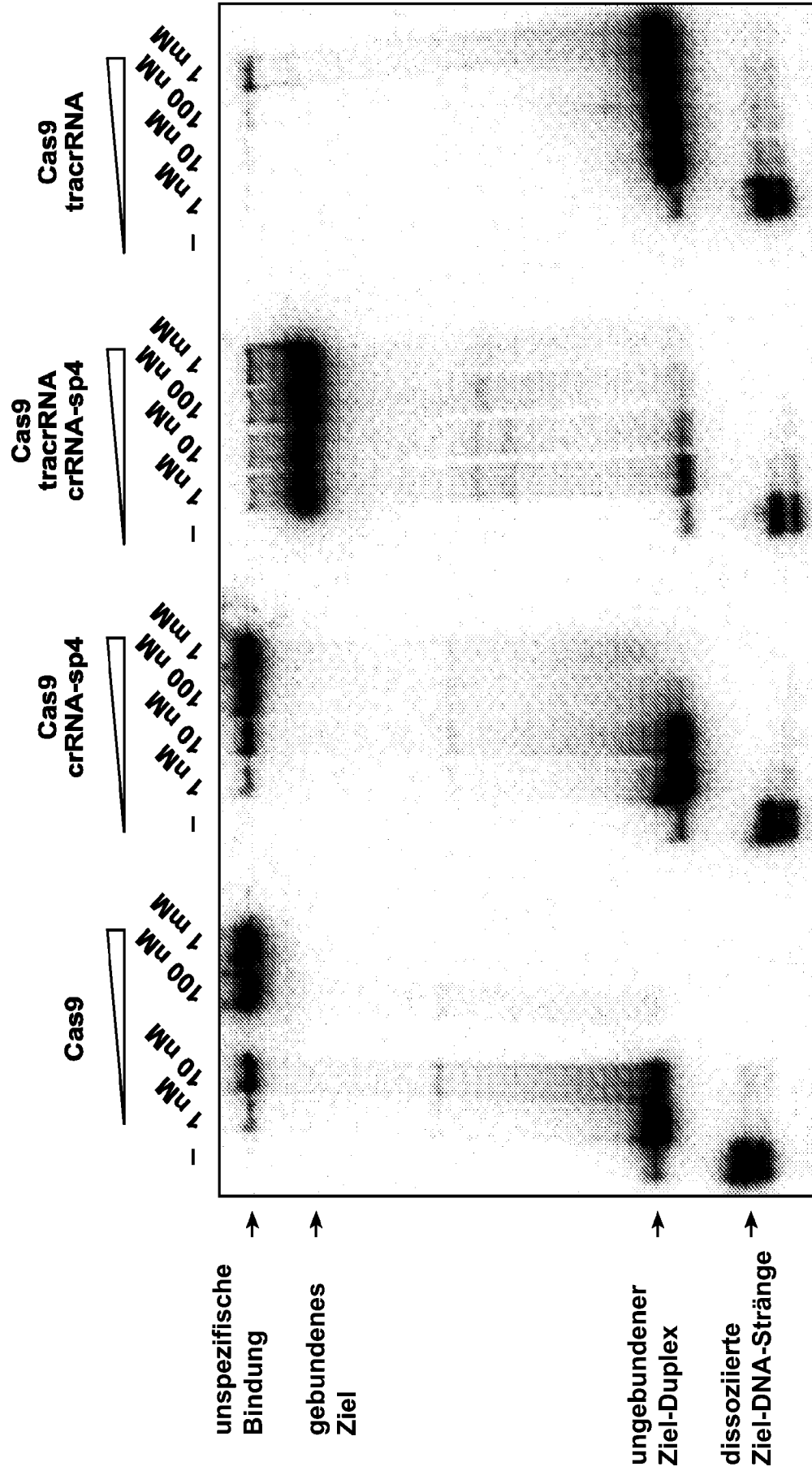
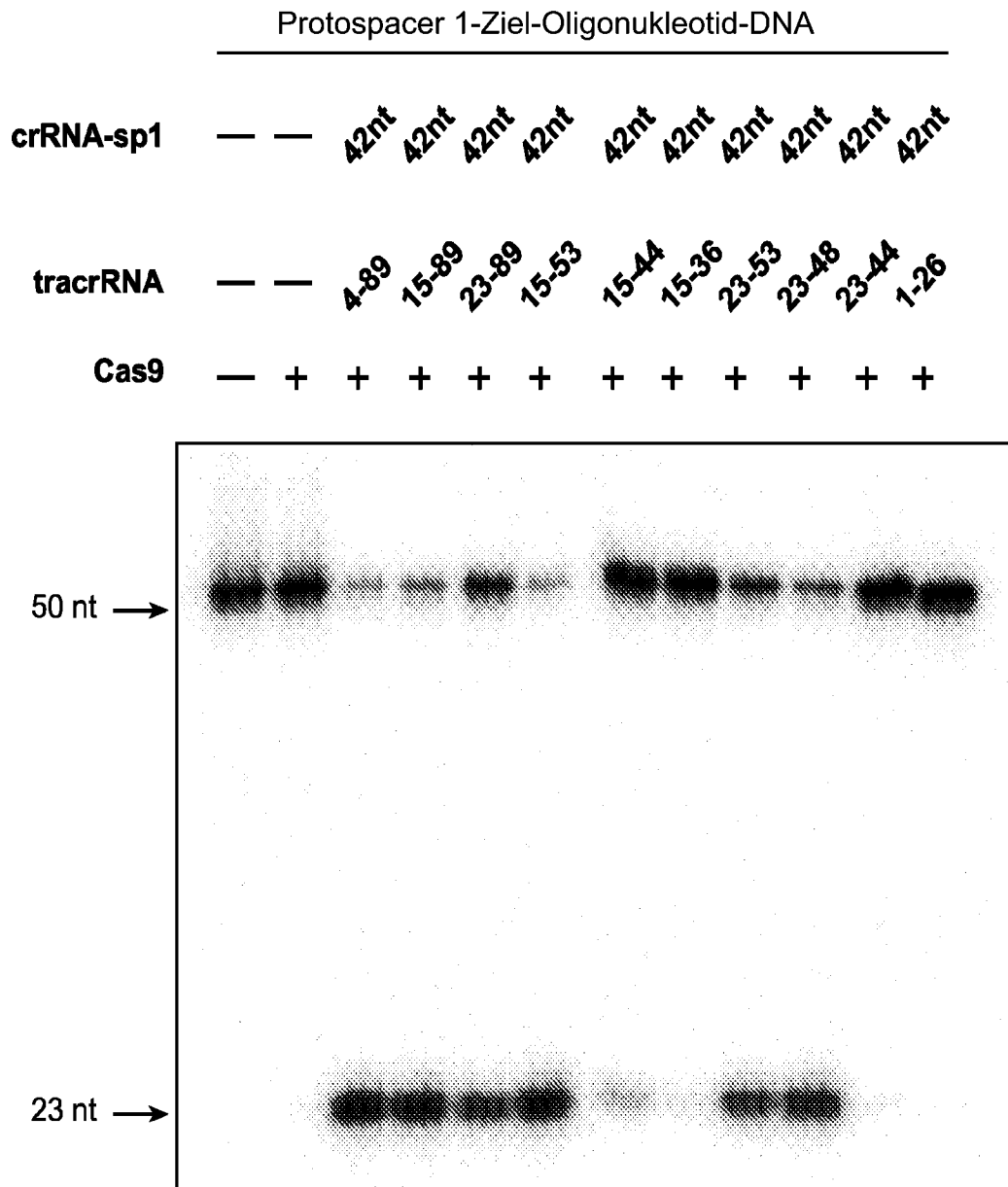


FIG. 22

A**FIG. 23**

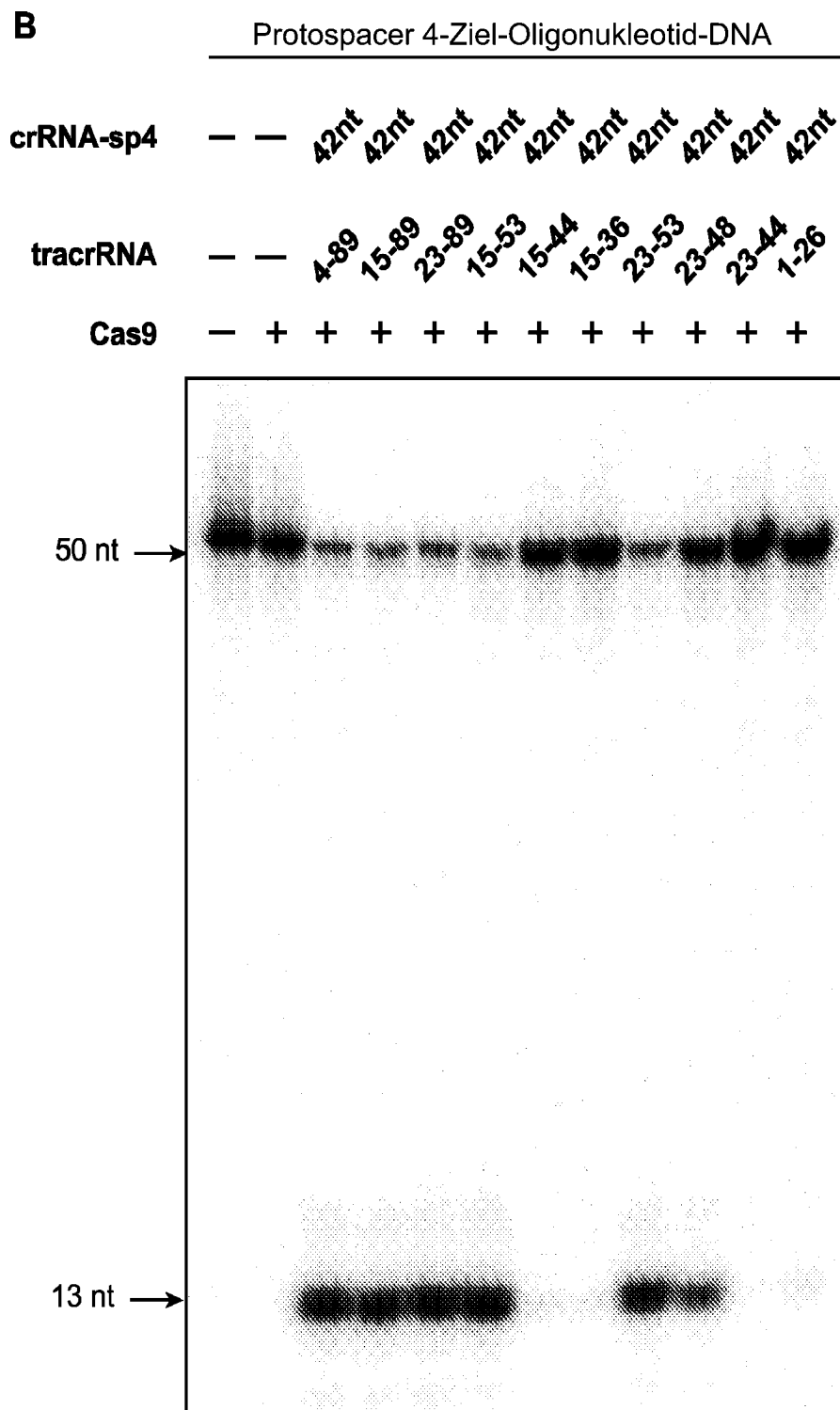


FIG. 23 (Fortsetz.)

A

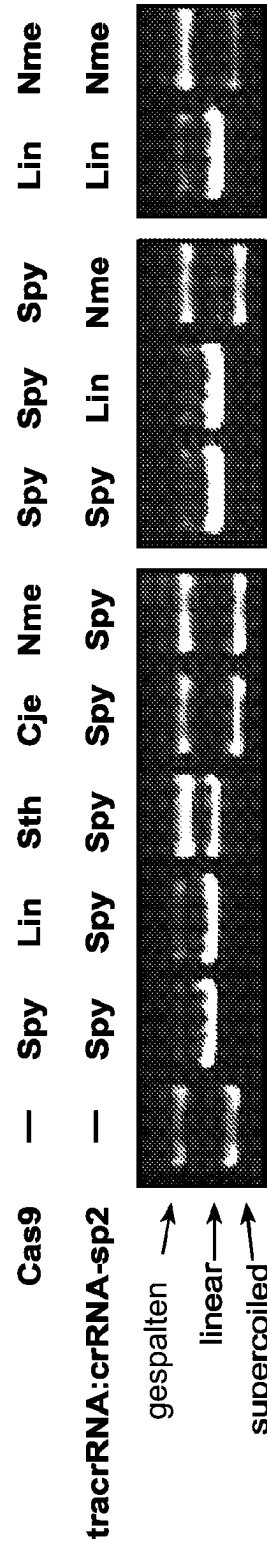


FIG. 24

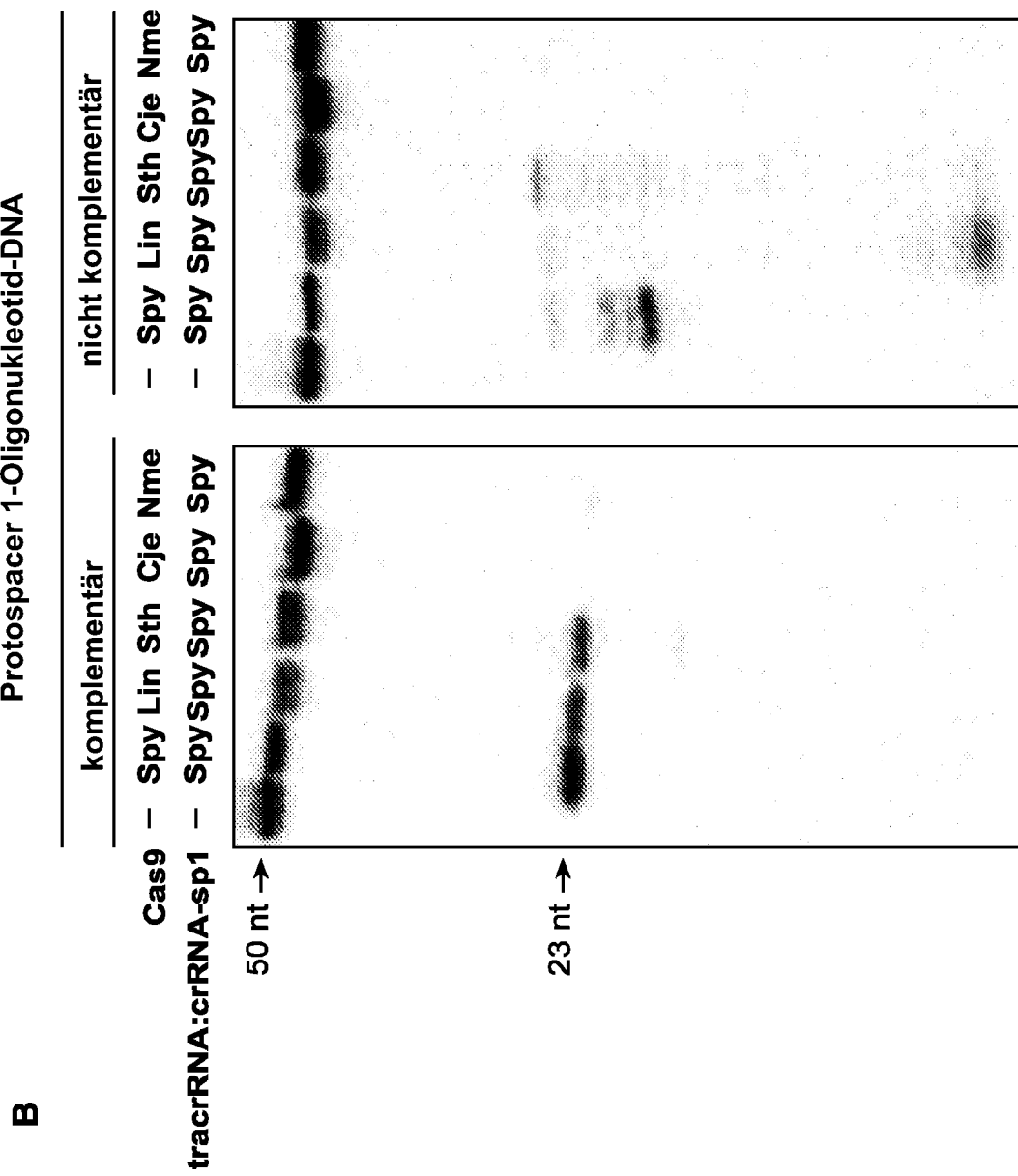


FIG. 24 (Fortsetz. 1)

C

	<i>S. pyogenes</i>	<i>L. innocua</i>	<i>S. thermophilus</i>	<i>C. jejuni</i>	<i>N. meningitidis</i>
<i>S. pyogenes</i>	x	54	58	16	16
<i>L. innocua</i>	54	x	52	15	14
<i>S. thermophilus</i>	58	52	x	16	15
<i>C. jejuni</i>	16	15	16	x	32
<i>N. meningitidis</i>	16	14	15	32	x

FIG. 24 (Fortsetz. 2)

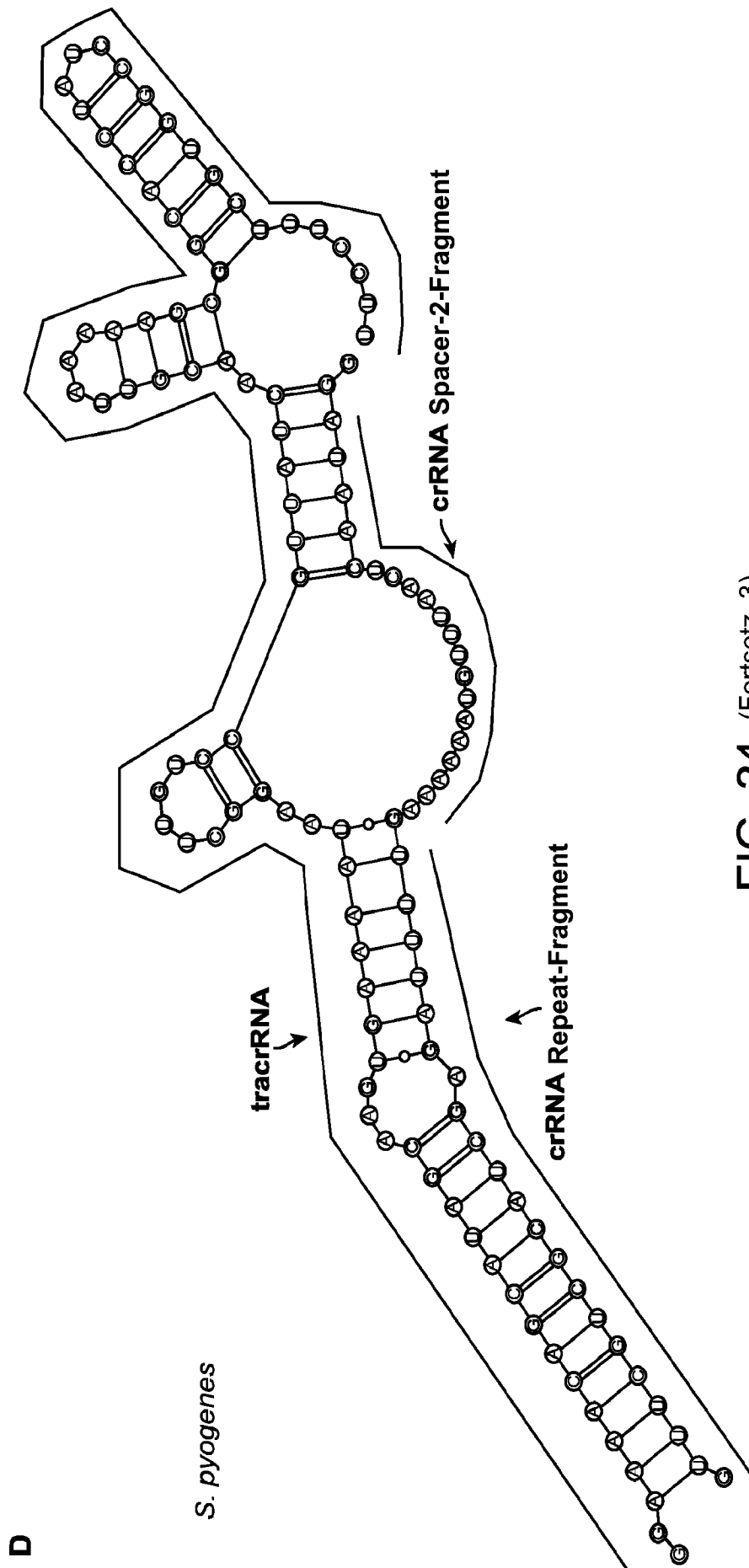


FIG. 24 (Fortsetz. 3)

D

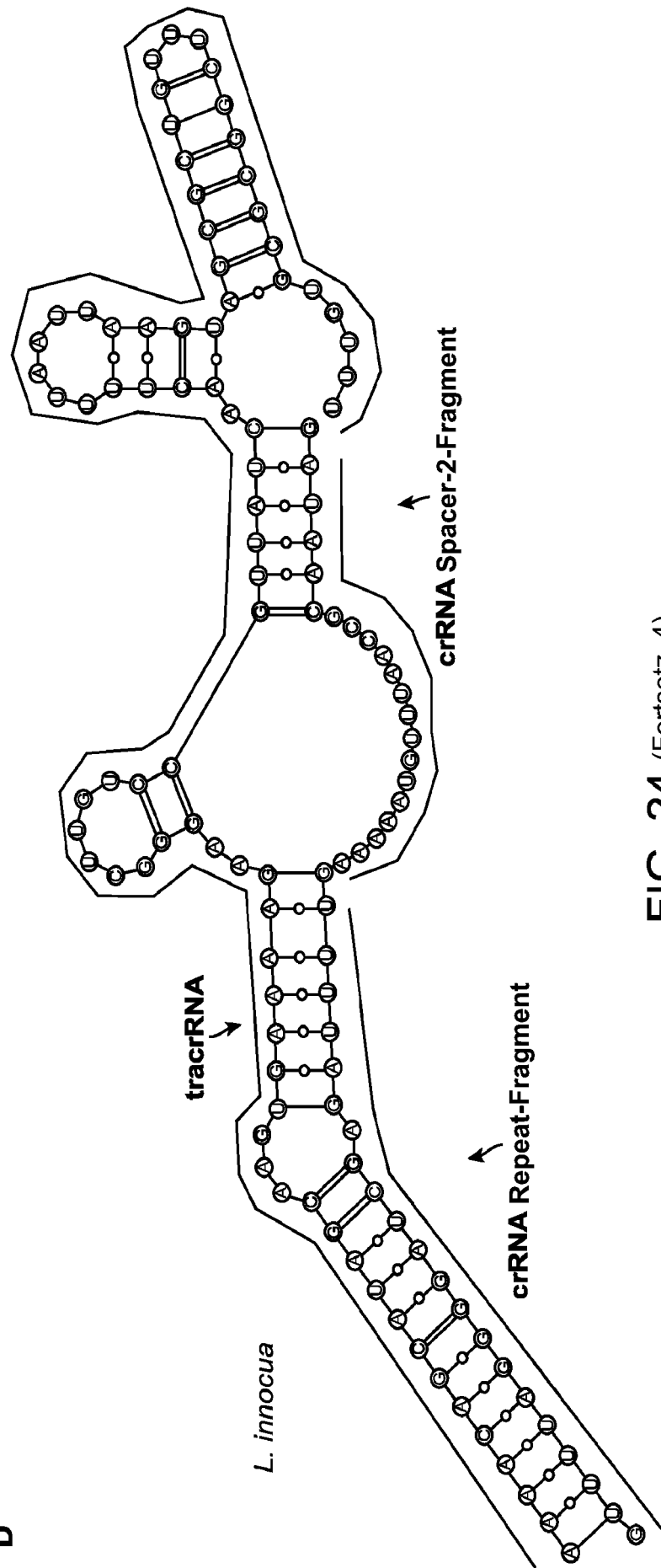


FIG. 24 (Fortsetz. 4)

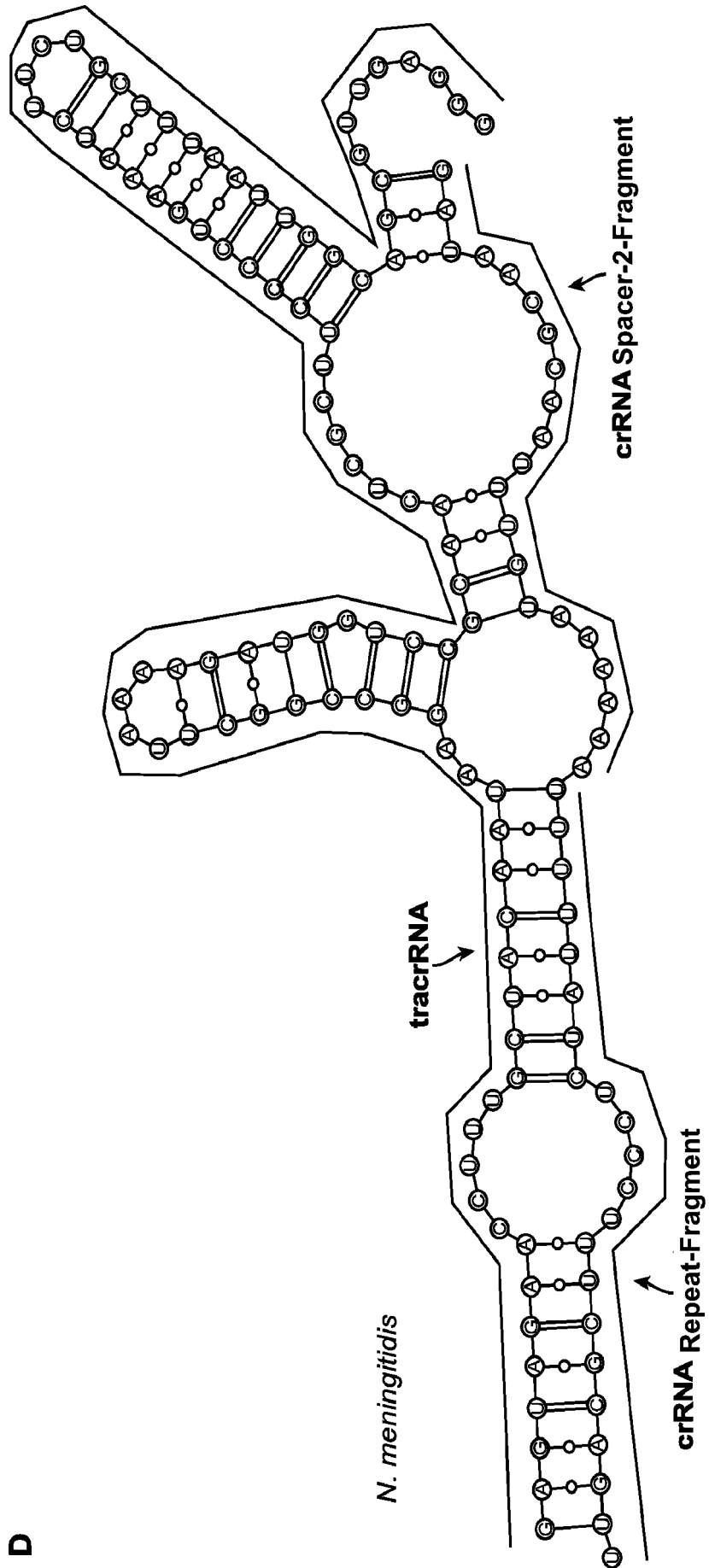


FIG. 24 (Fortsetz. 7)

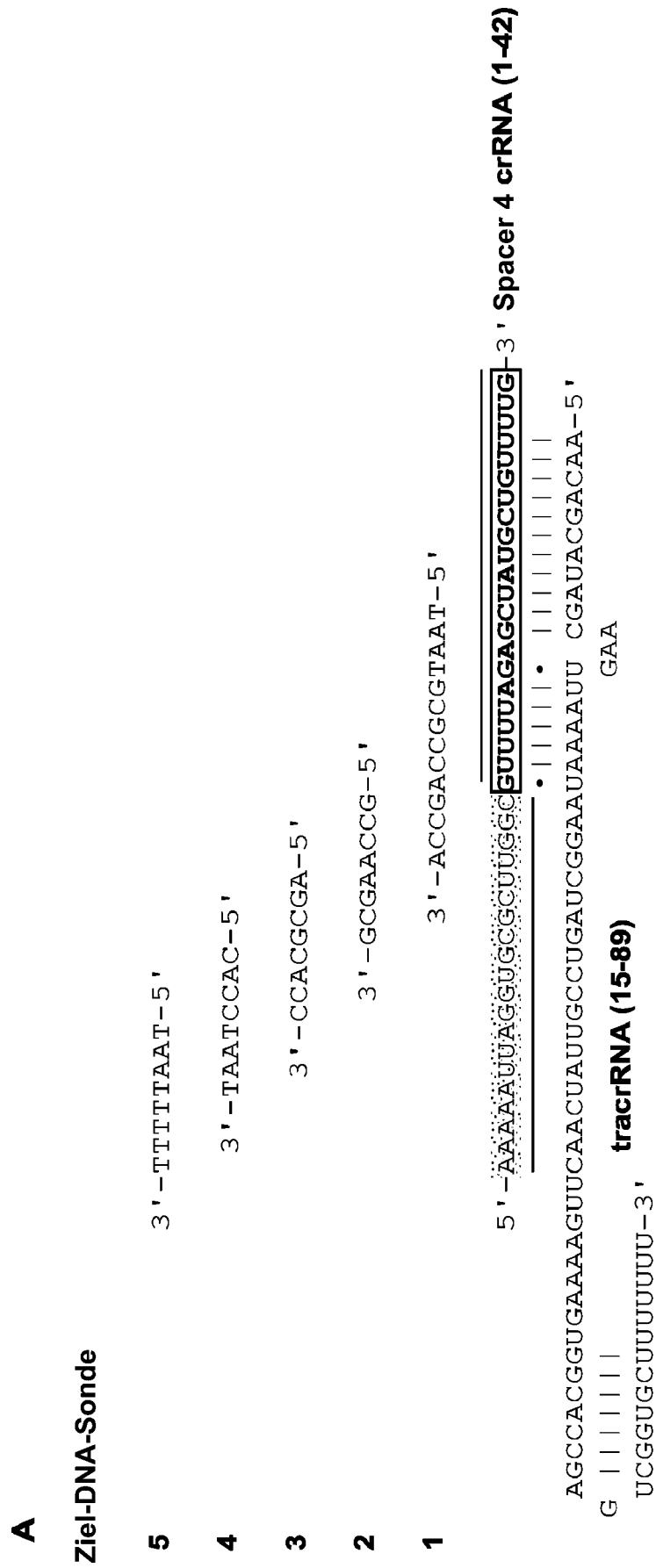


FIG. 25

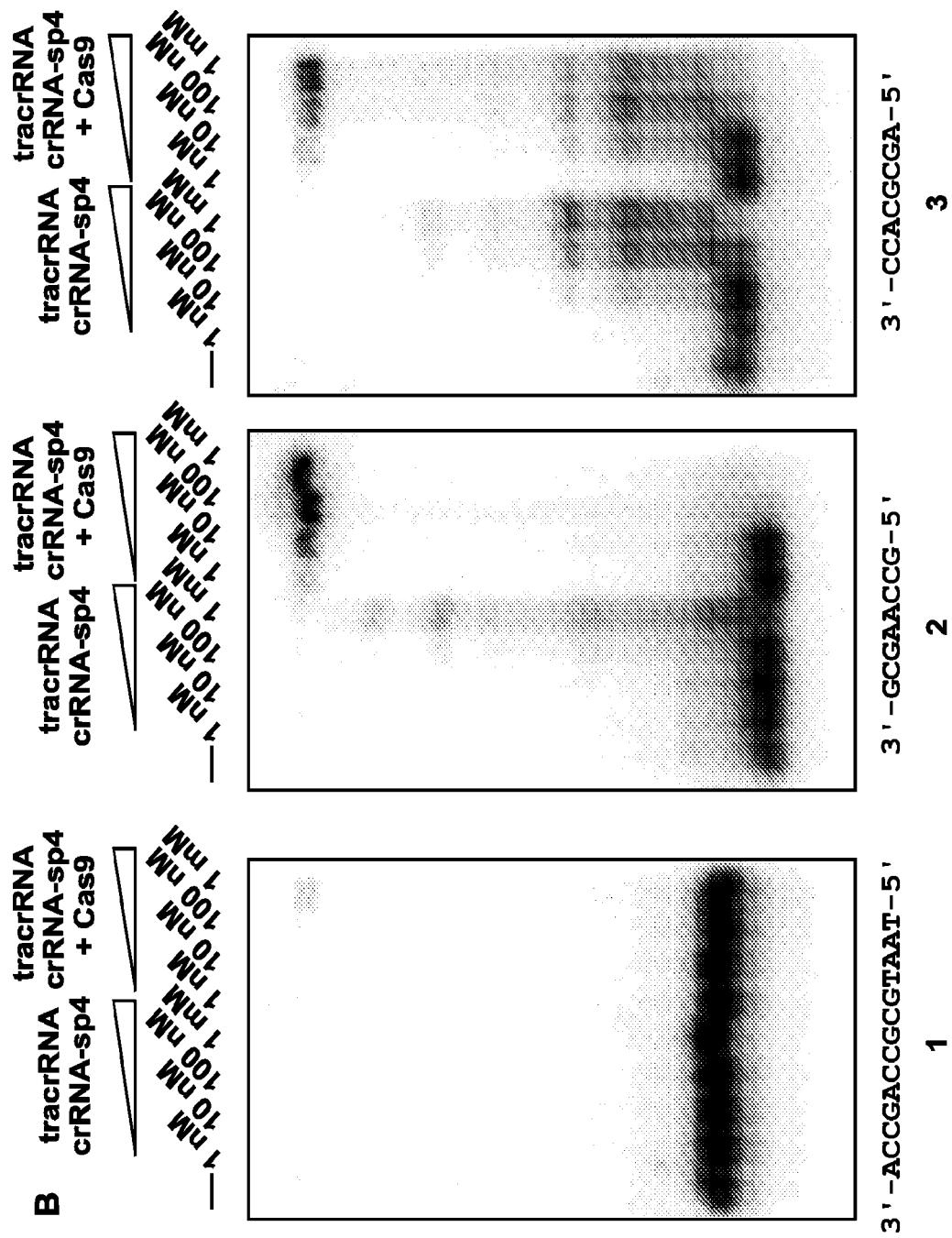


FIG. 25 (Fortsetz. 1)

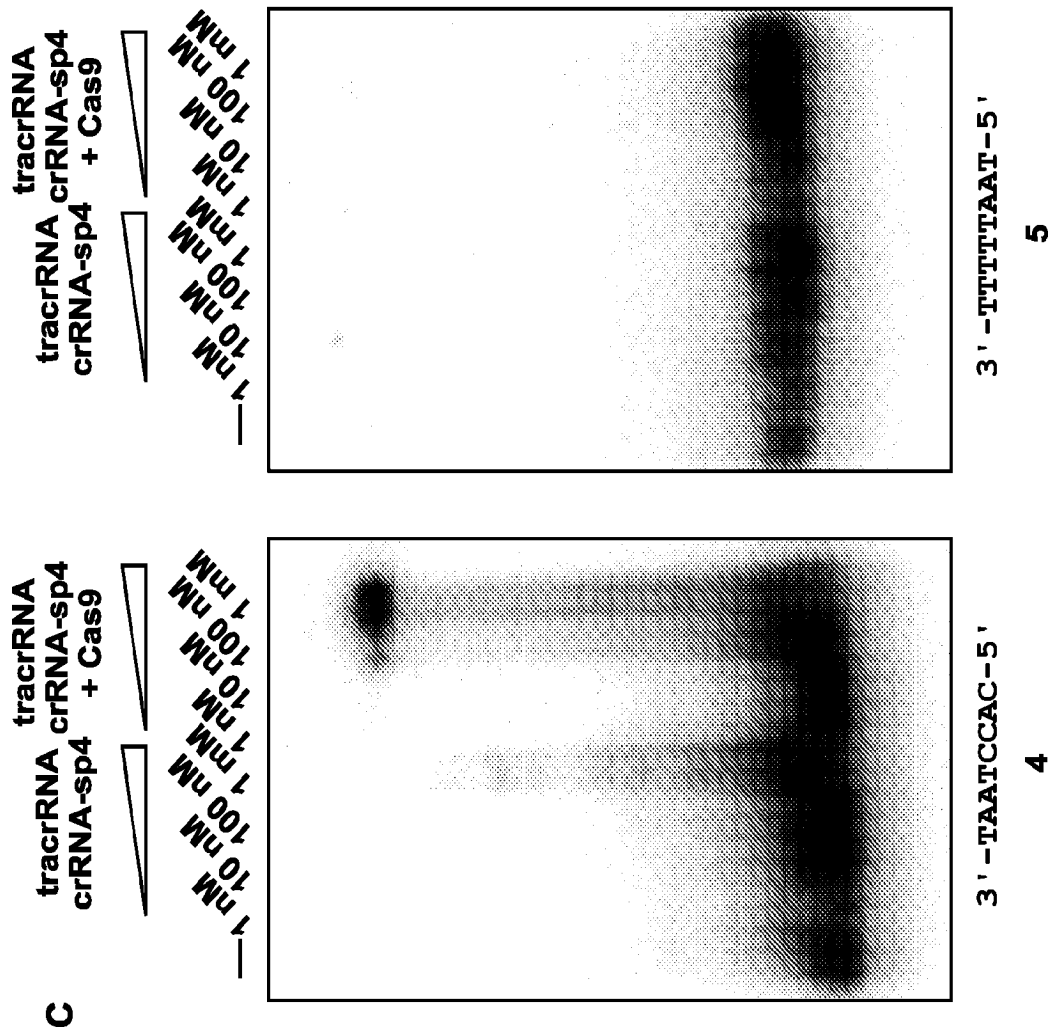


FIG. 25 (Fortsetz. 2)

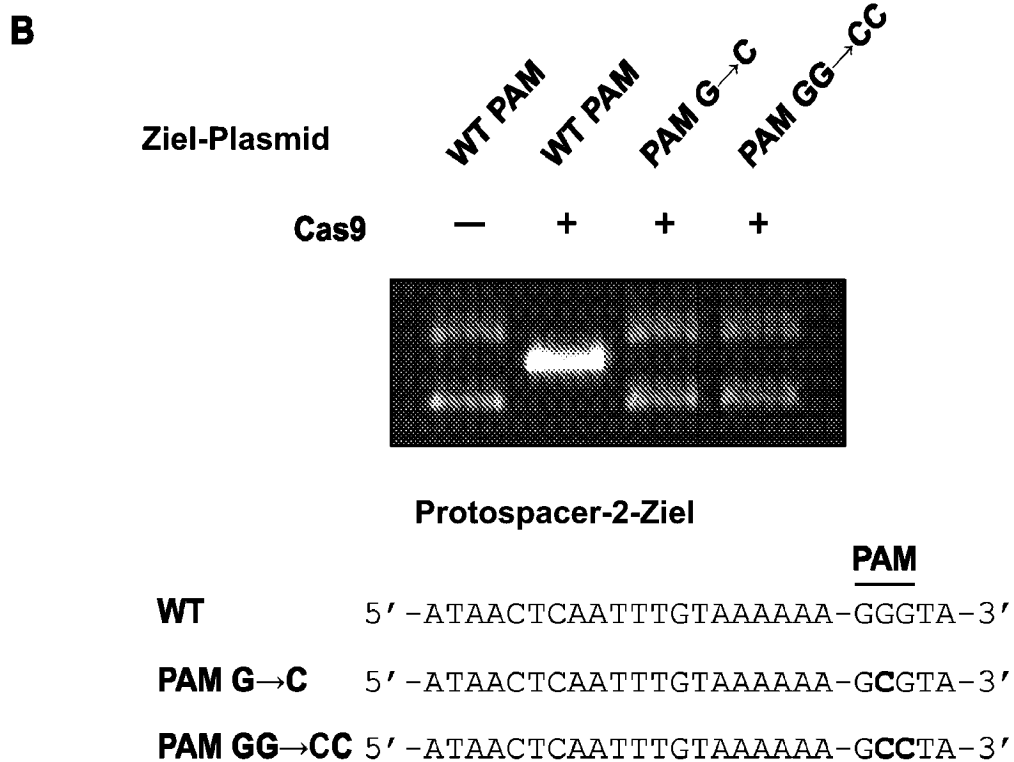
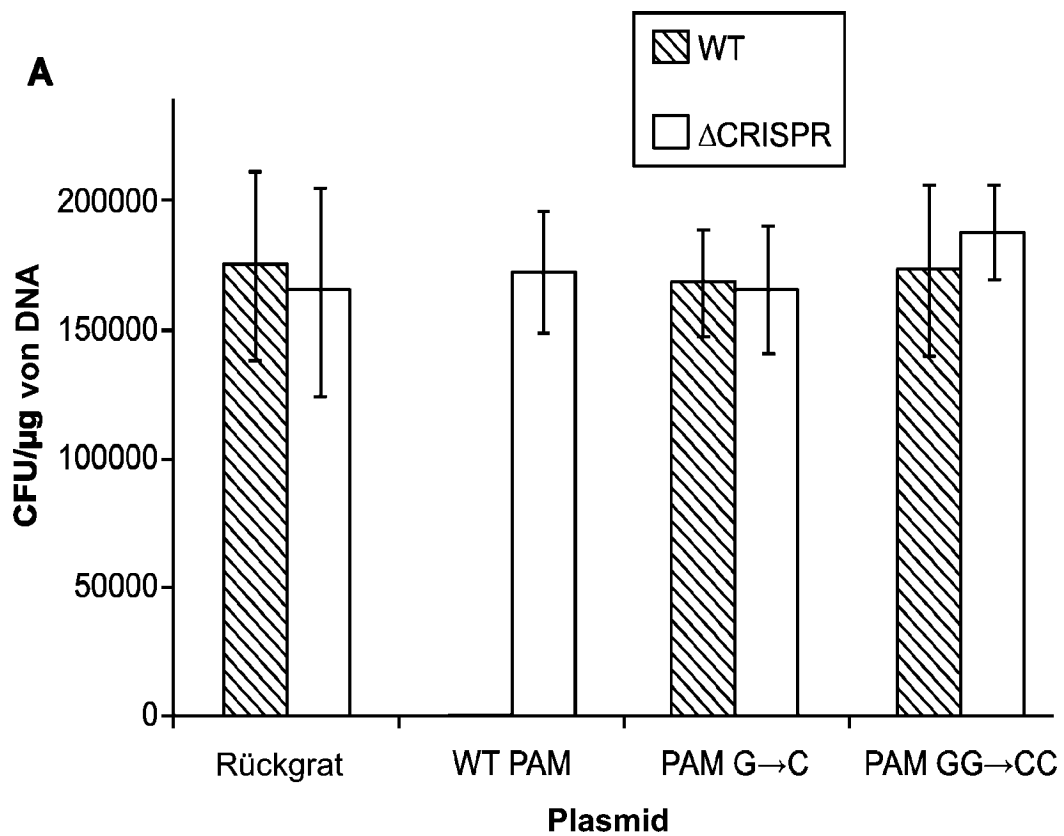
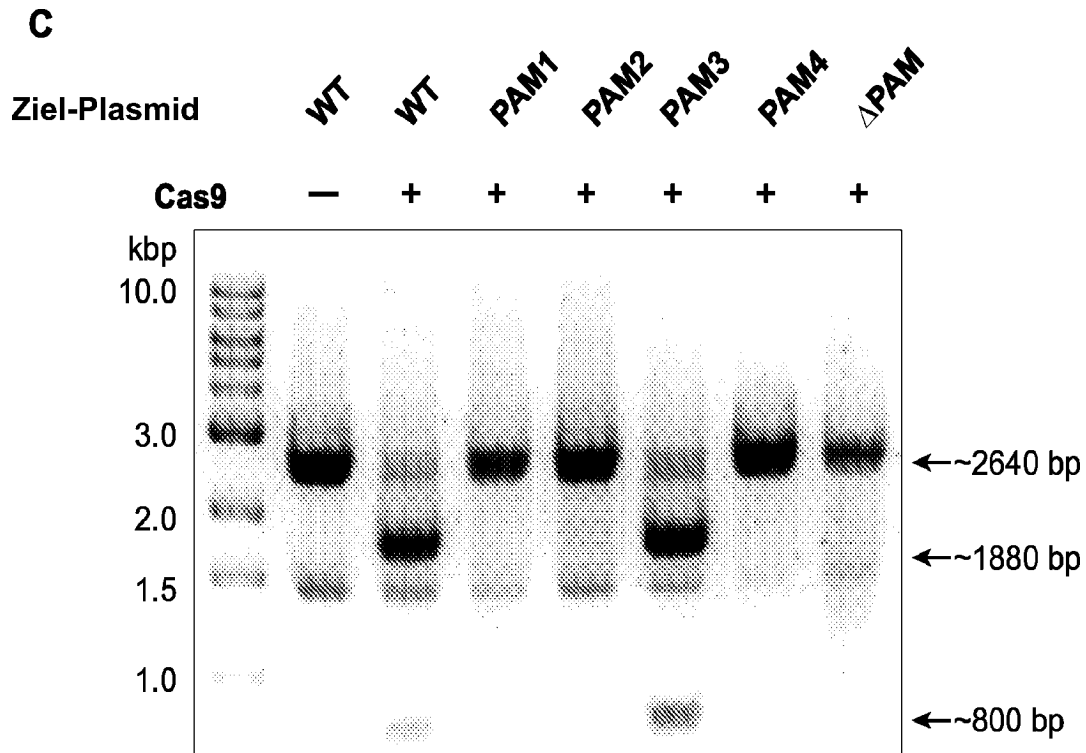


FIG. 26



Protospacer 4-Ziel

	<u>PAM</u>
WT	5' -AAAAATTAGGTGCGCTTGGC-TGGCG-3'
PAM1	5' -AAAAATTAGGTGCGCTTGGC-T CG CG-3'
PAM2	5' -AAAAATTAGGTGCGCTTGGC-TG CC G-3'
PAM3	5' -AAAAATTAGGTGCGCTTGGC-TGG CC -3'
PAM4	5' -AAAAATTAGGTGCGCTTGGC-T CCCC -3'
DPAM	5' -AAAAATTAGGTGCGCTTGGC-T-3'

FIG. 26 (Fortsetz.)

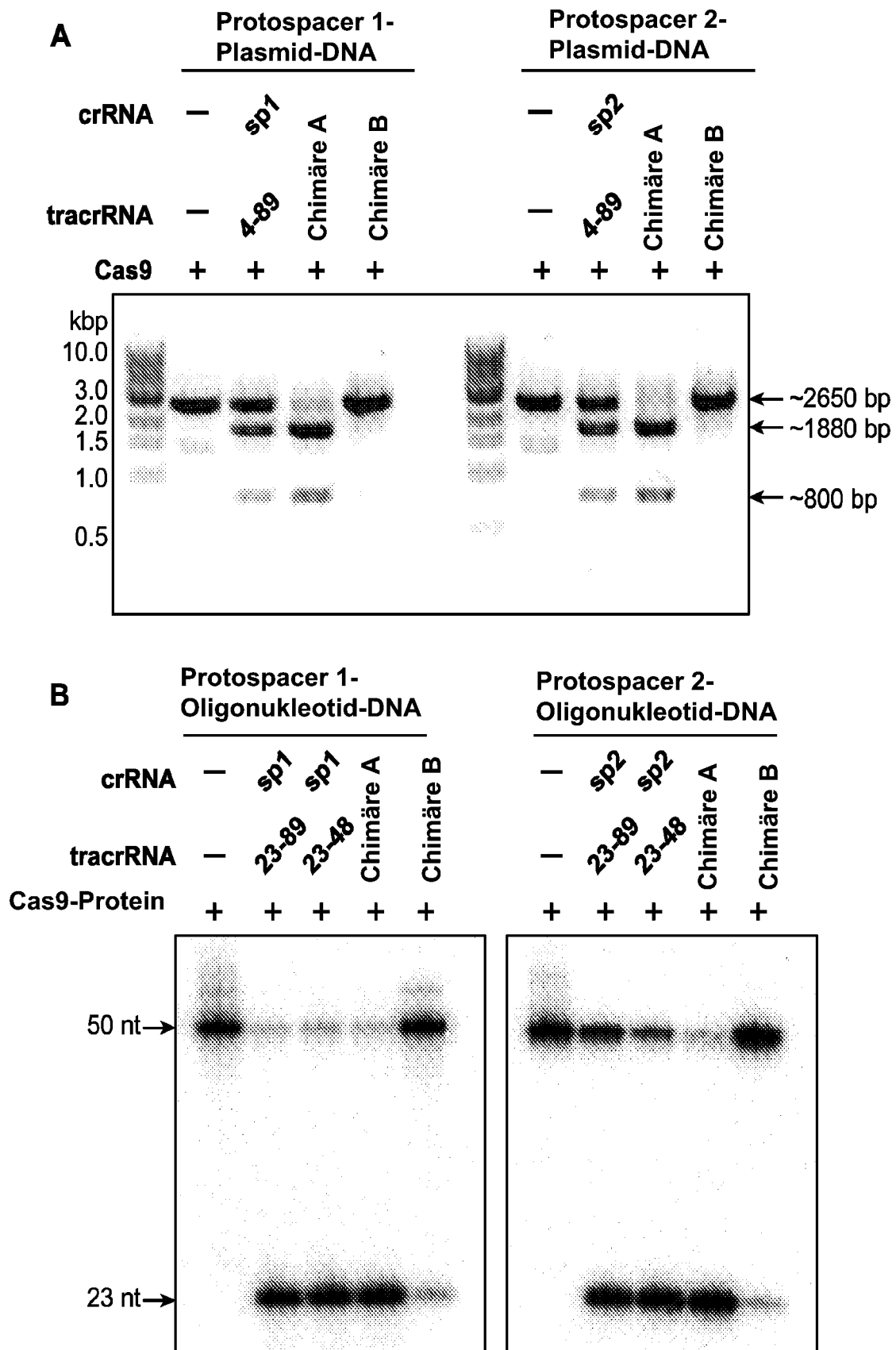
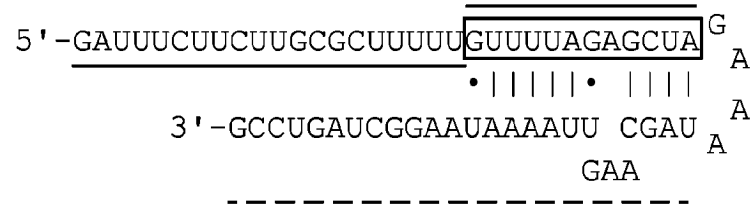
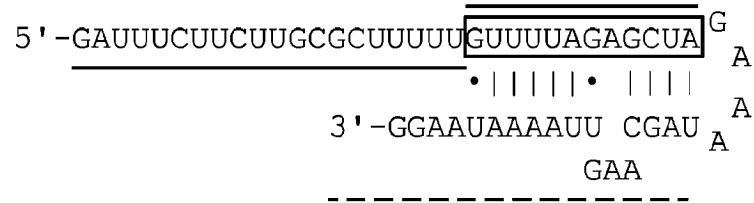
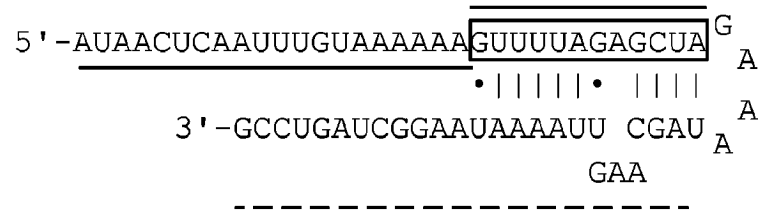
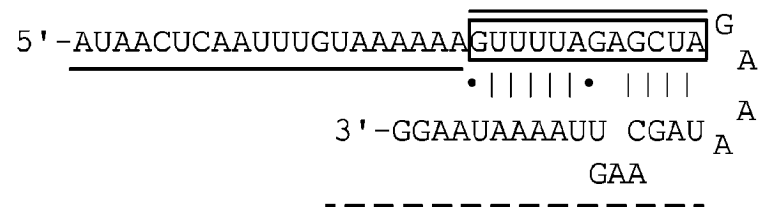


FIG. 27

C Auf Protospacer 1 gerichtete chimäre RNAsChimäre **A**Chimäre **B**

Auf Protospacer 2 gerichtete chimäre RNAs

Chimäre **A**Chimäre **B****FIG. 27** (Fortsetz.)

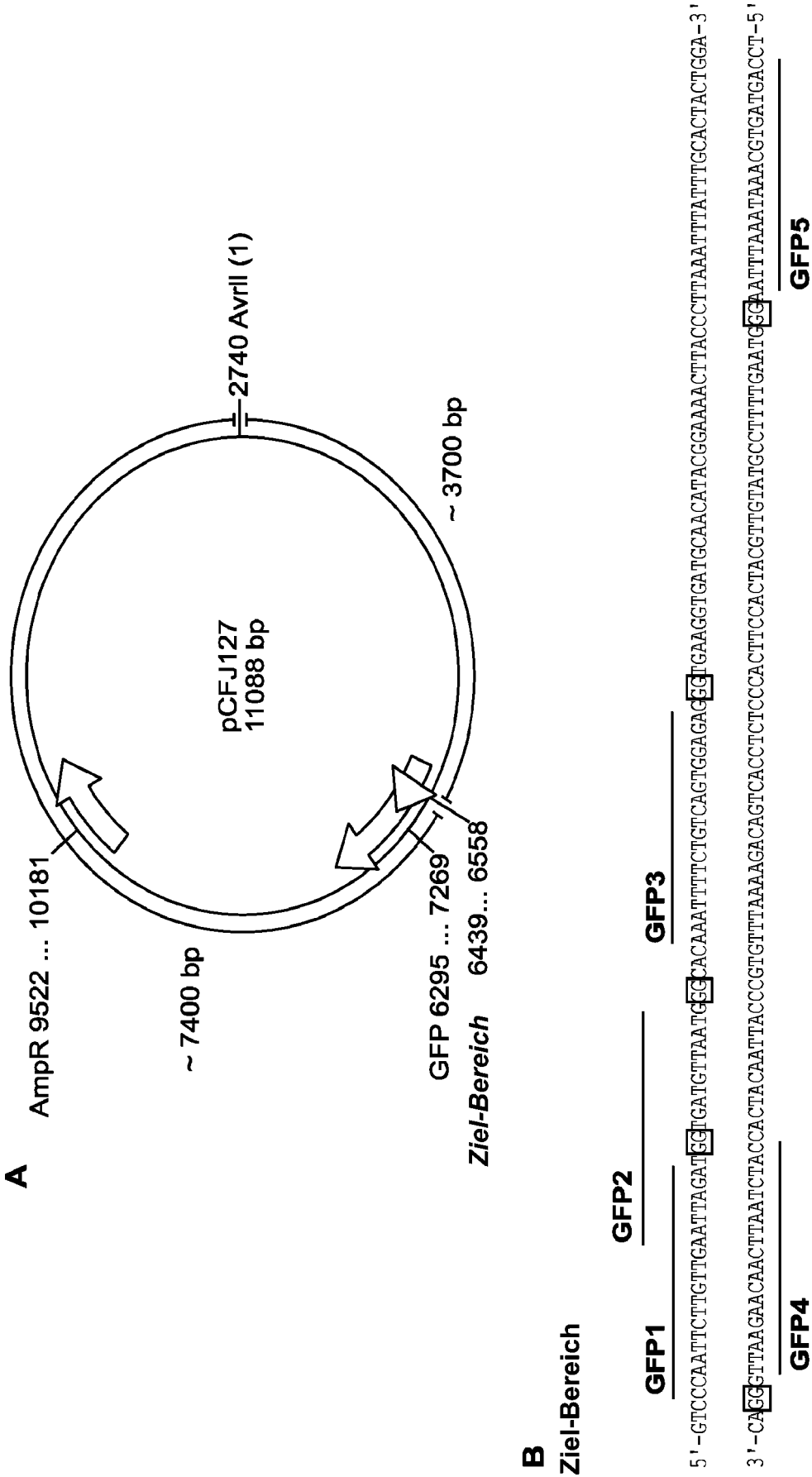


FIG. 28

C	Ziel-Sequenz	PAM
GFP1	CCAAATTC TTGTTGAATTAGA-TGGTGA	5'-CCAAUUCU UUGAAUUAGAGUUUUAGAGCUA ^G A • • A 3'- CUGAUCGGGAAUAAAAUU CGAU ^A A GAA
GFP2	AATTAGATGGTGATGTTAAT-GGGCAC	5'-AAUUAGAU UGAUGUUAAUGUUUUAGAGCUA ^G A • • A 3'-GCCUGAUCGGGAAUAAAAUU CGAU ^A A GAA
GFP3	AAATTTTCTGTCAGTGGAGA-GGGTGA	5'-AAAUUUUCUGUCAGUGGAGAGUUUUAGAGCUA ^G A • • A 3'-GCCUGAUCGGGAAUAAAAUU CGAU ^A A GAA
GFP4	CATCTAATTCAACAAGAATT-GGGACA	5'-CAUCUAAUUCAAACAAGAAUUUGUUUUAGAGCUA ^G A • • A 3'-GCCUGAUCGGGAAUAAAAUU CGAU ^A A GAA
GFP5	CAGTAGTGCAAAATAAATTTA-AGGGTA	5'-CAGUAGUGCAAAUAAAUUUUAGUUUUAGAGCUA ^G A • • A 3'-GCCUGAUCGGGAAUAAAAUU CGAU ^A A GAA

FIG. 28 (Fortsetz. 1)

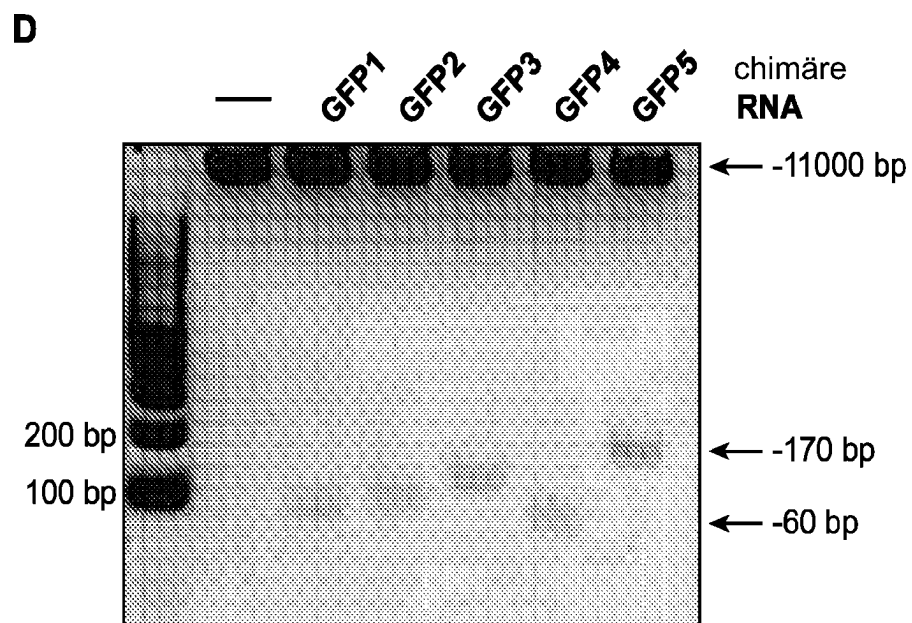


FIG. 28 (Fortsetz. 2)

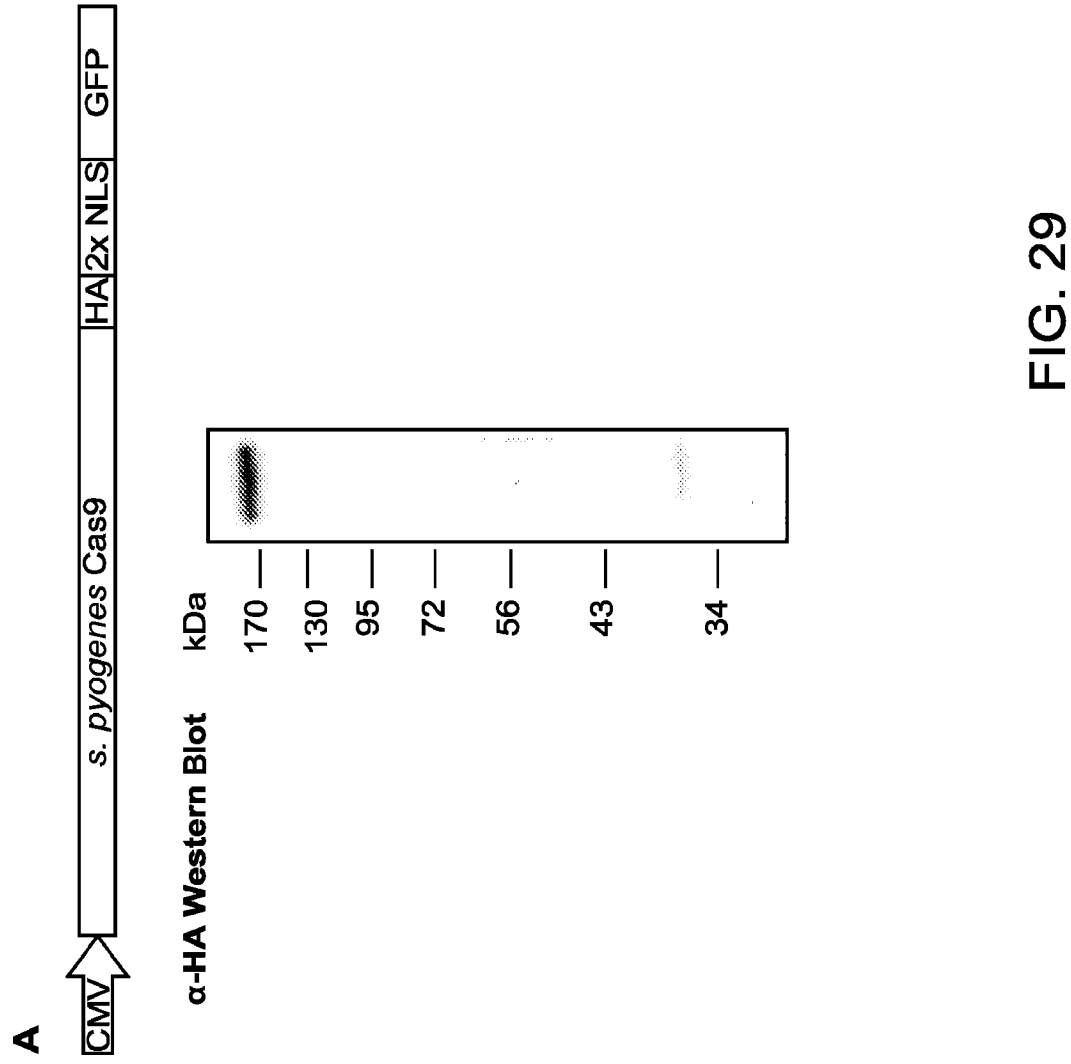
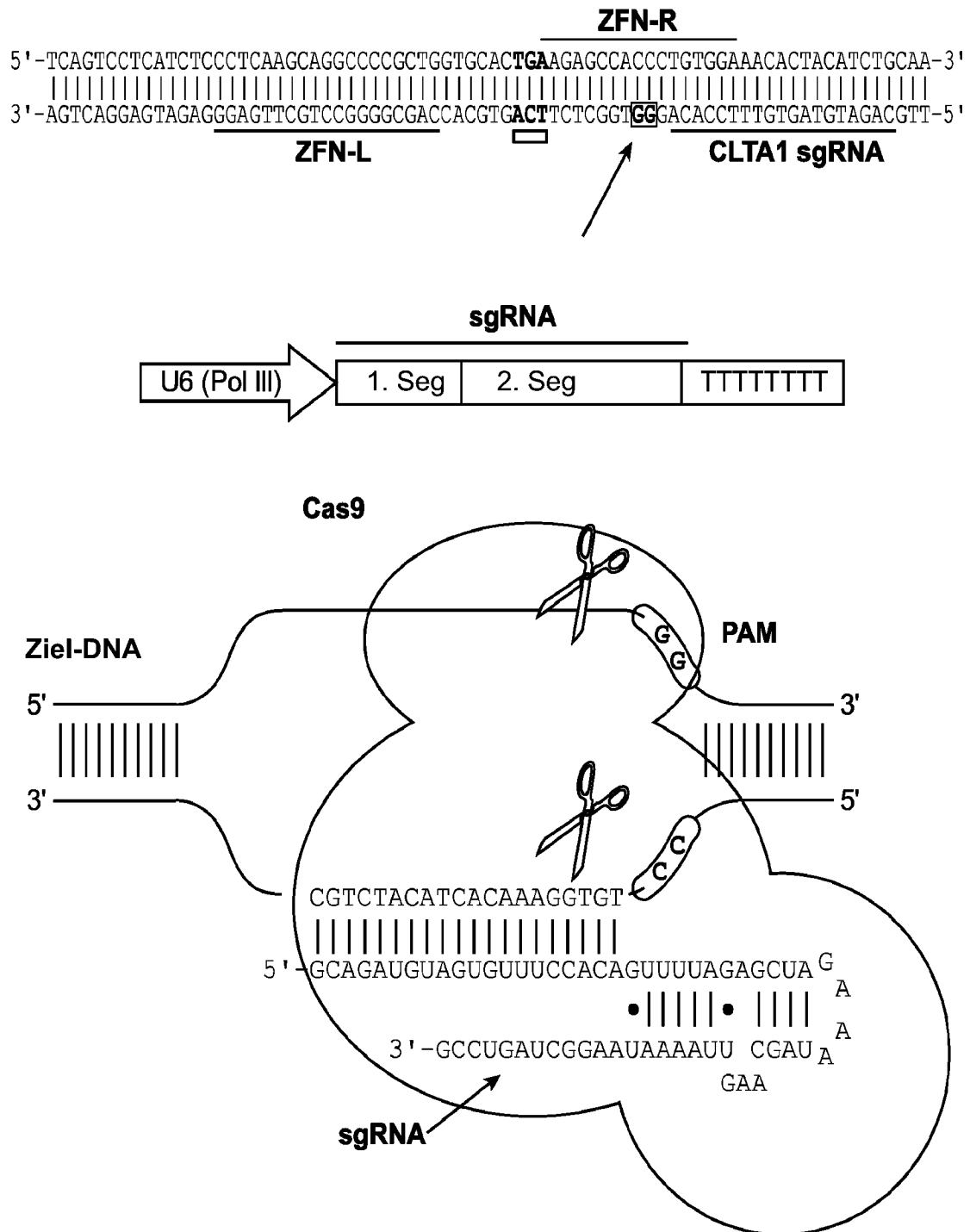


FIG. 29

C**FIG. 29** (Fortsetz. 1)

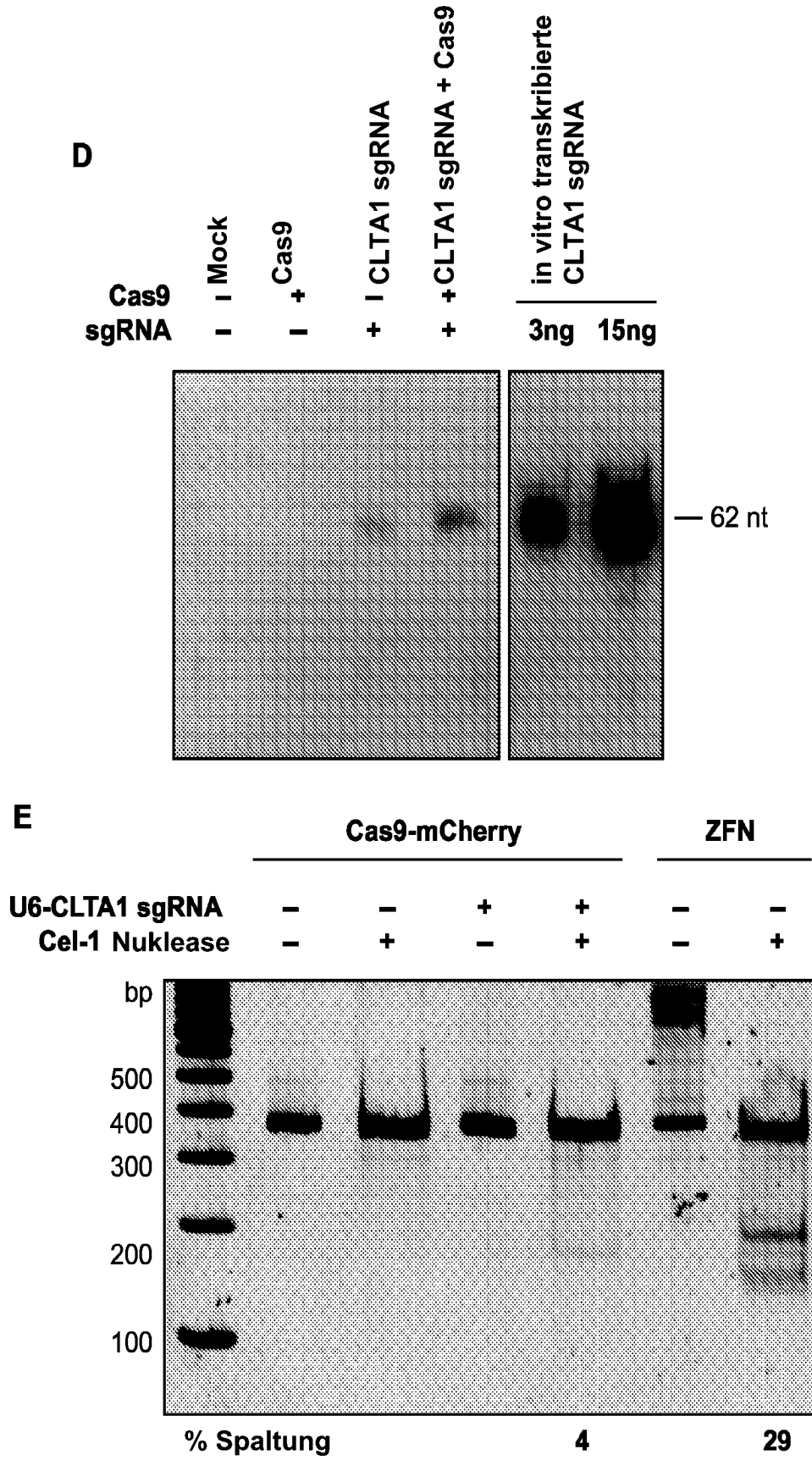
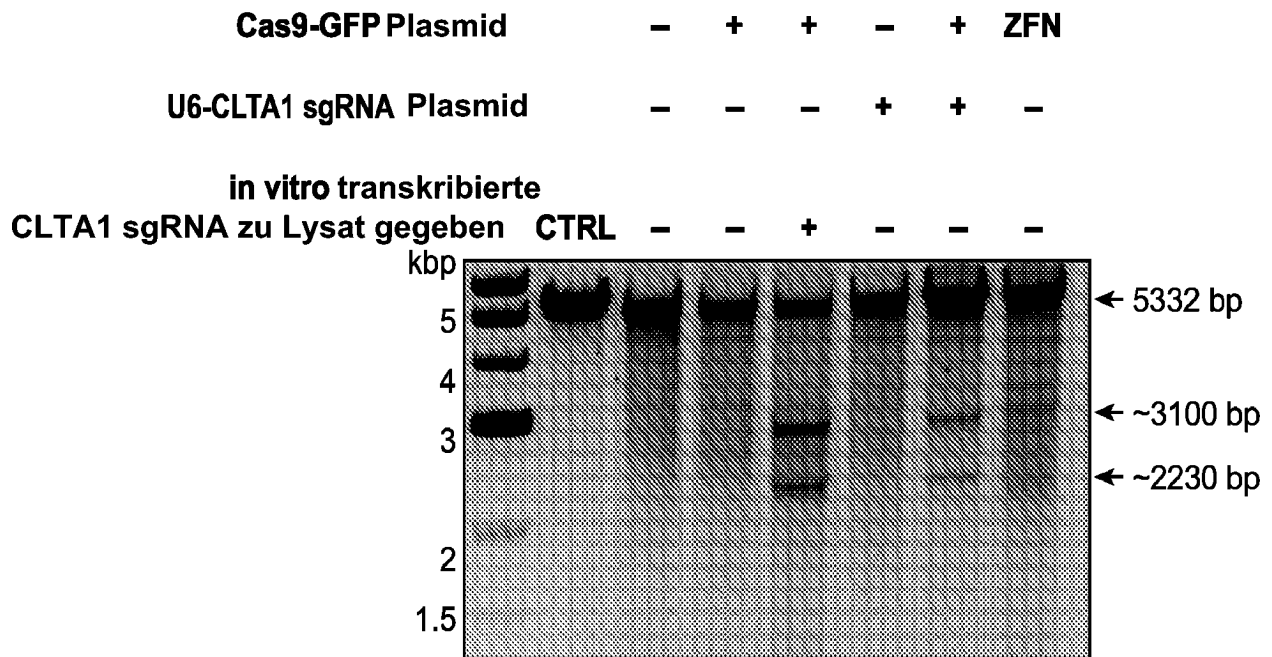
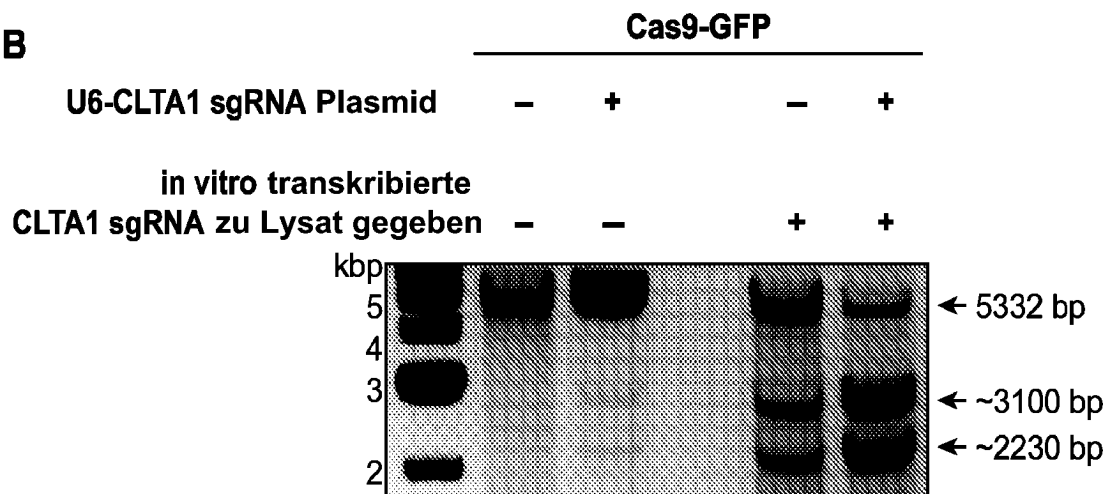
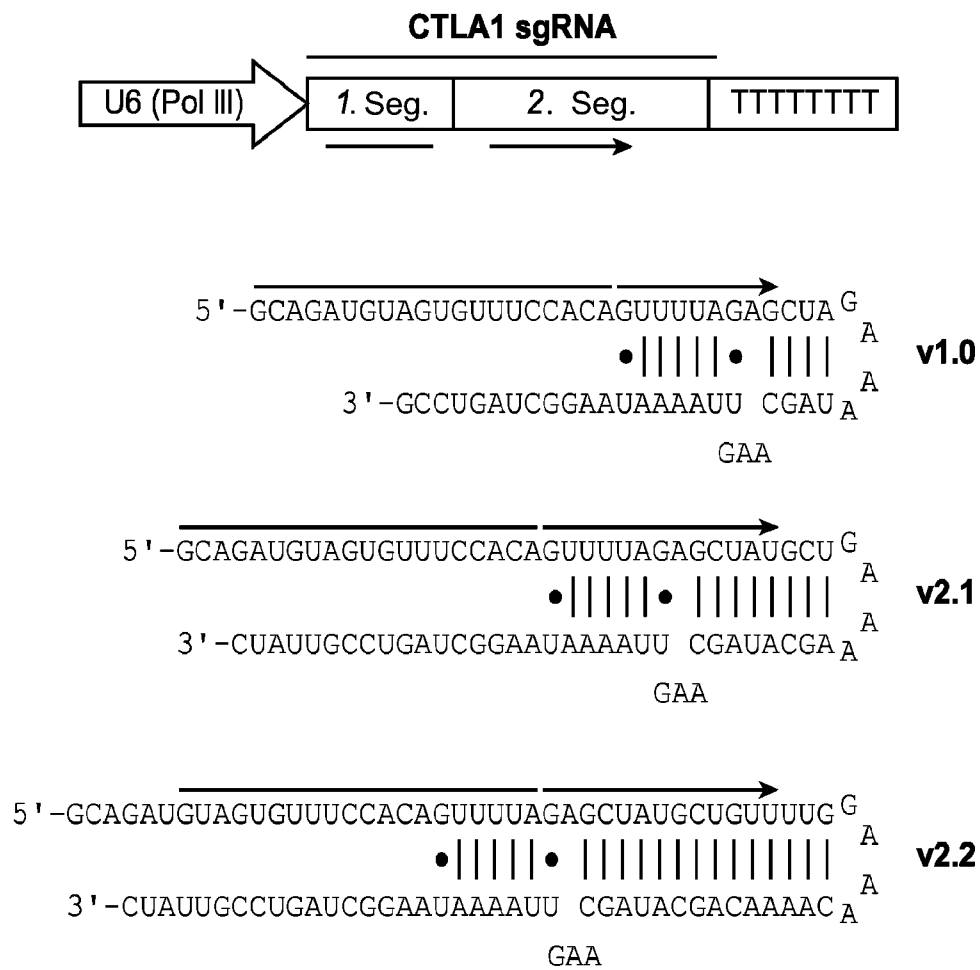


FIG. 29 (Fortsetz. 2)

A**B****FIG. 30**

A**FIG. 31**

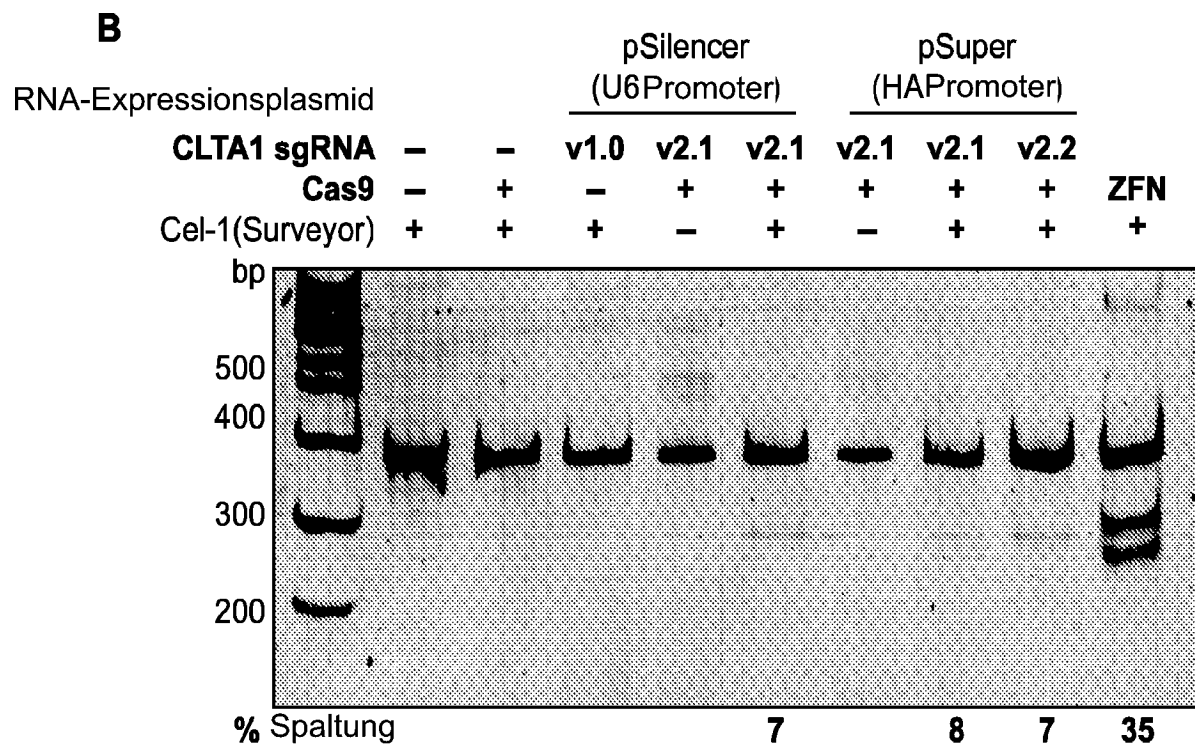


FIG. 31 (Fortsetz.)

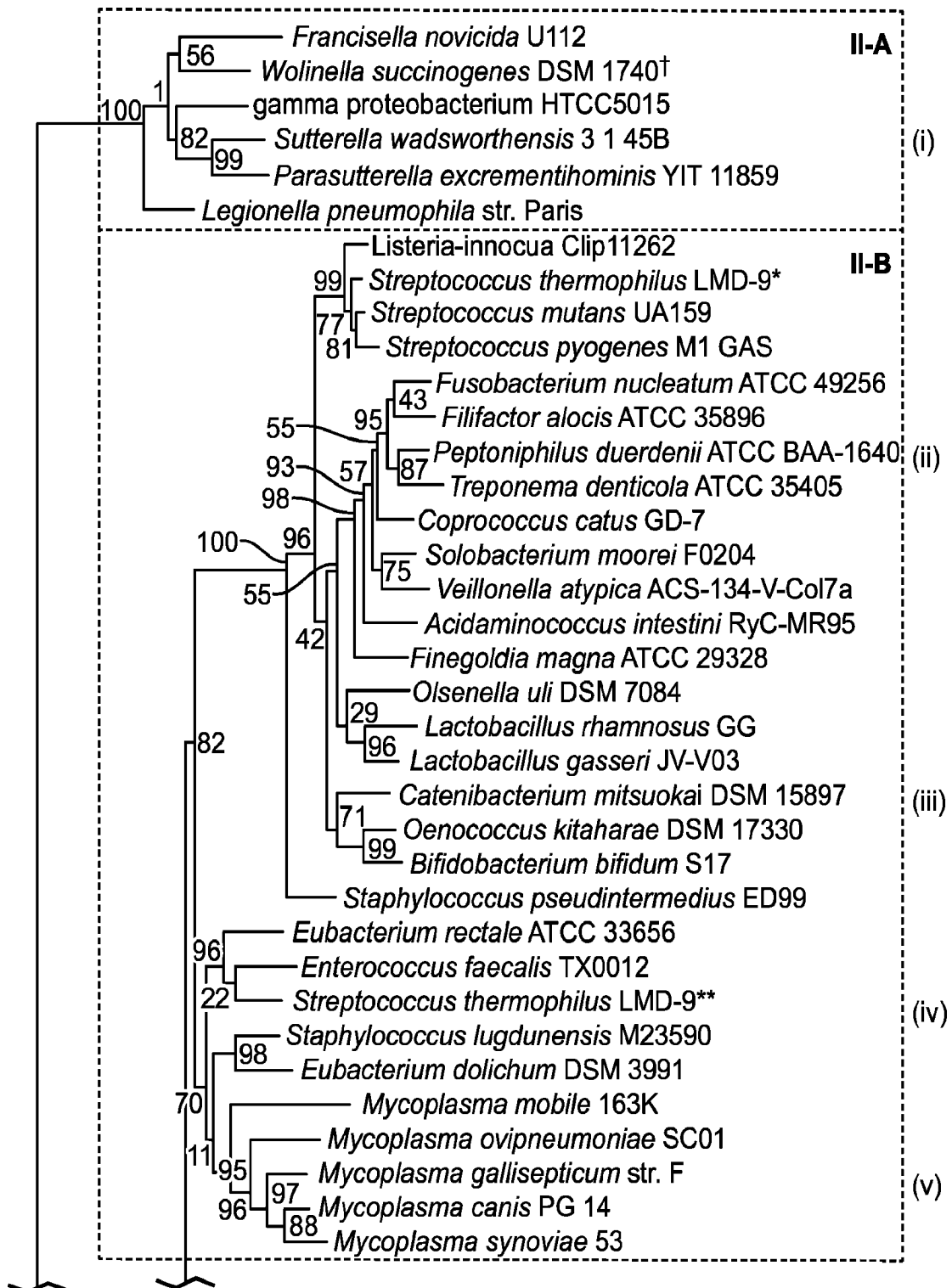


FIG. 32A

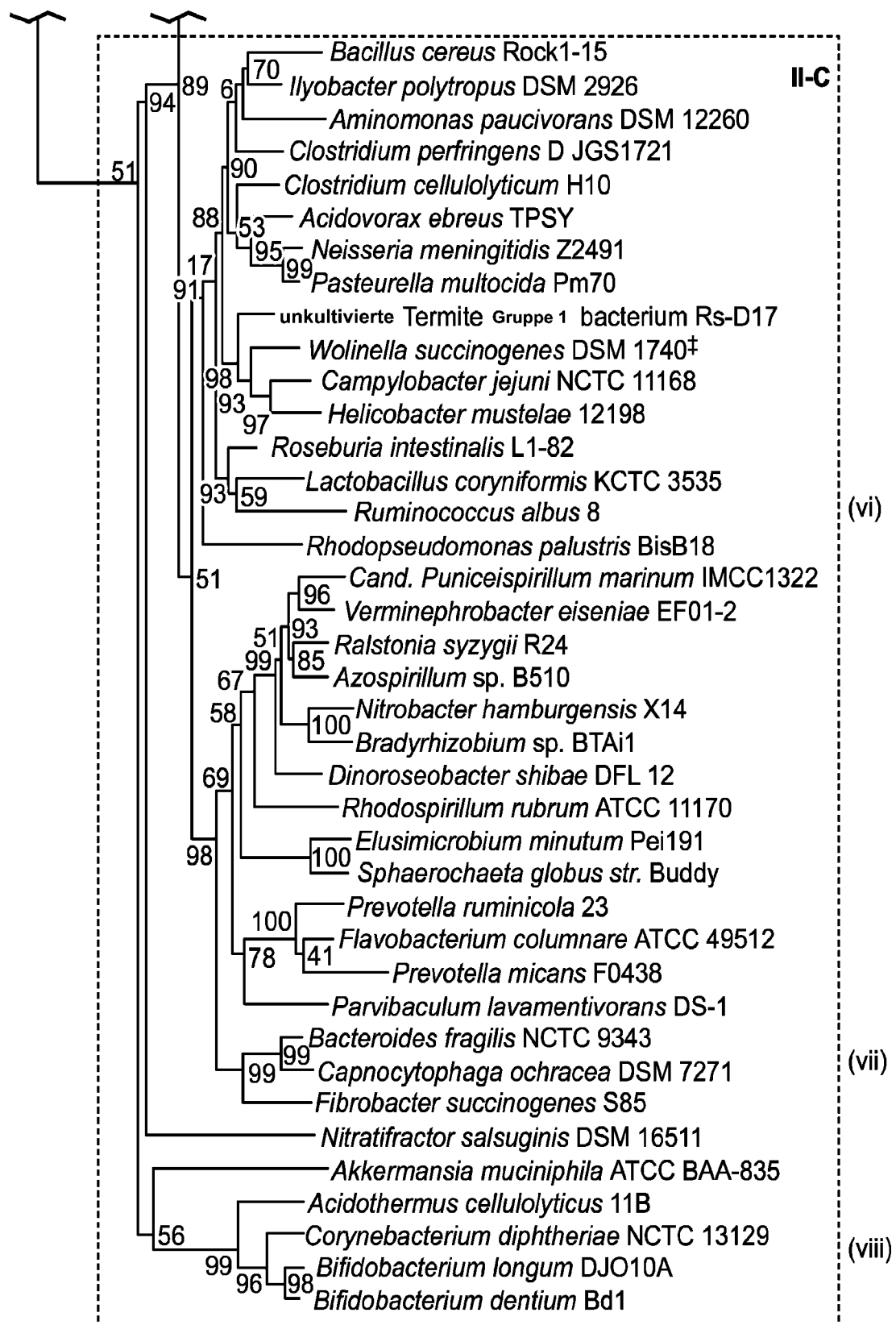


FIG. 32A (Fortsetz.)

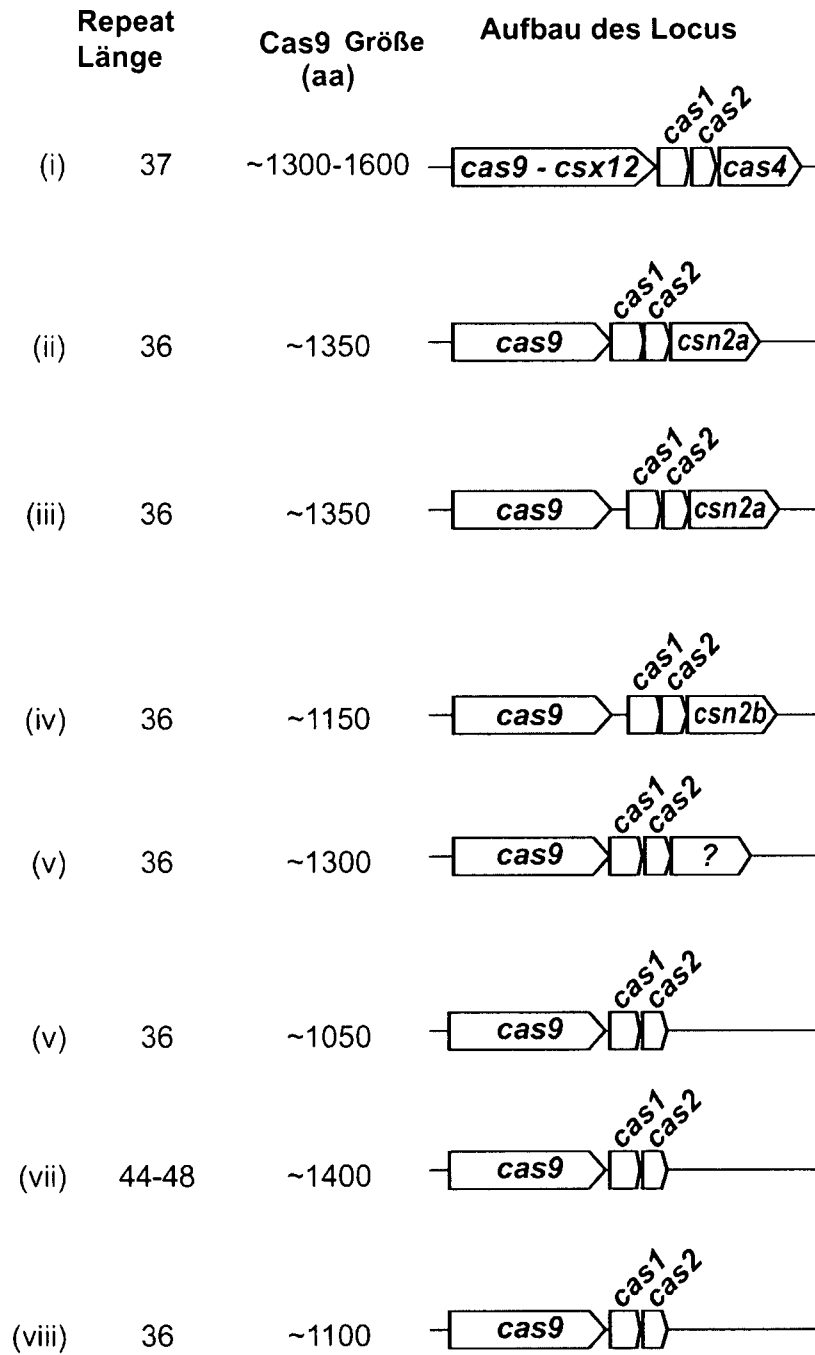


FIG. 32B

Fig. 33A

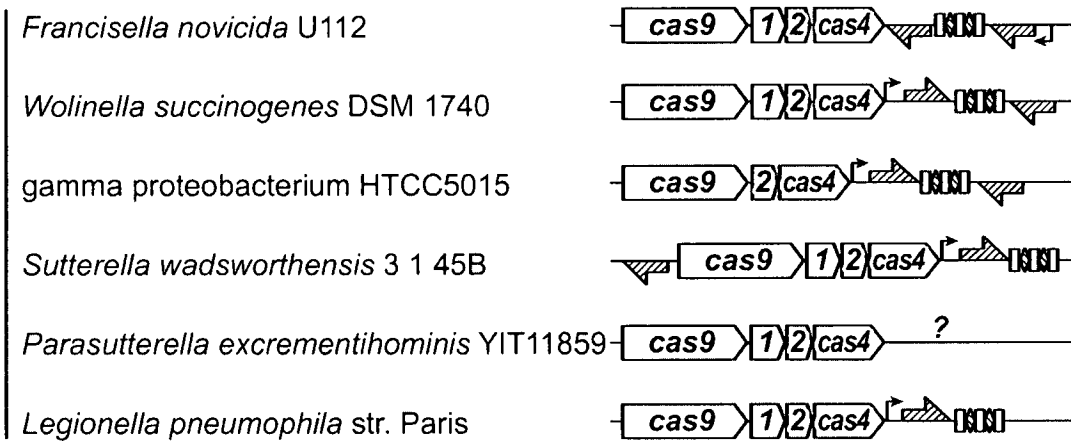
Typ II-A

Fig. 33B (Fortsetzung)

Typ II-B

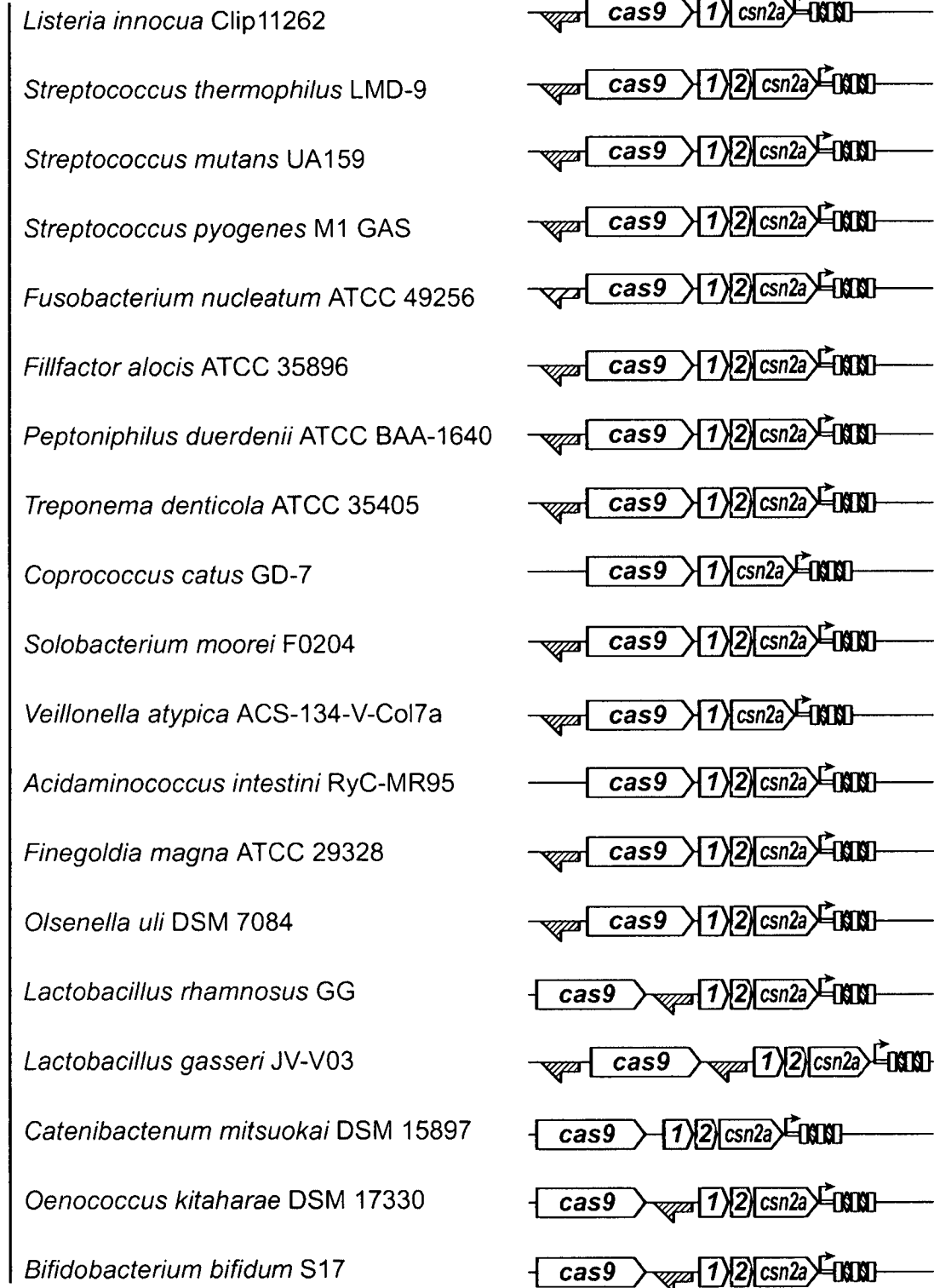


Fig. 33C (Fortsetzung)

Typ II-B

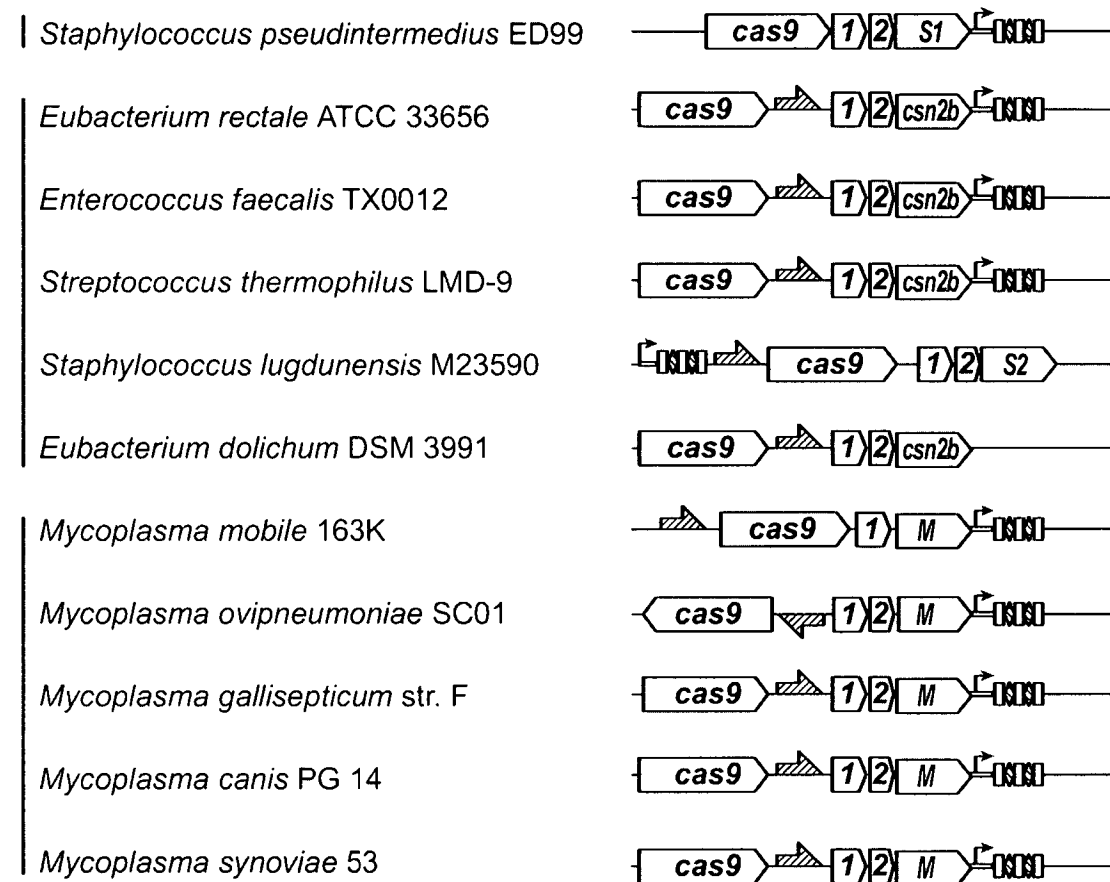


Fig. 33D Typ II-C (Fortsetzung)

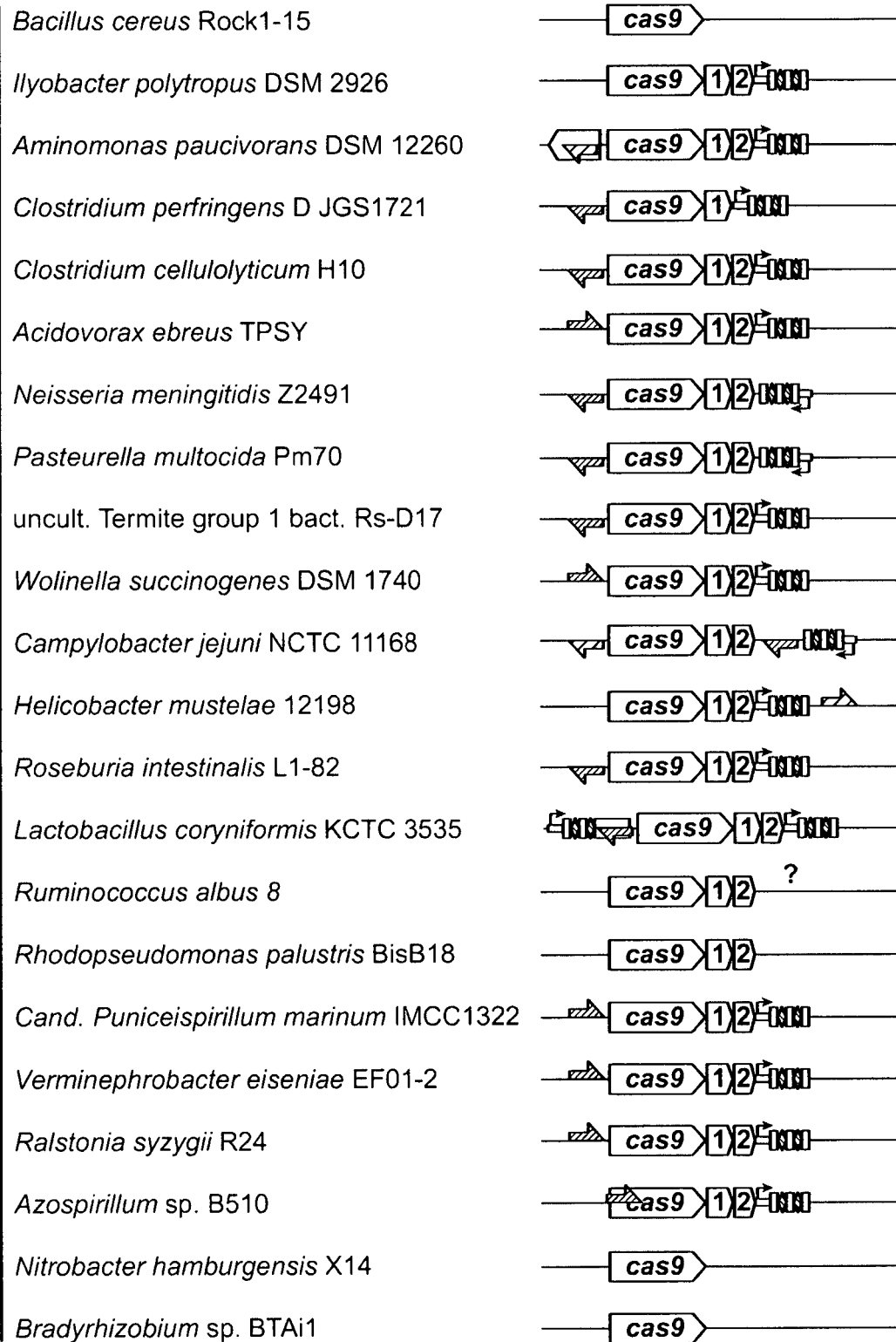


Fig. 33E

Typ II-C (Fortsetzung)

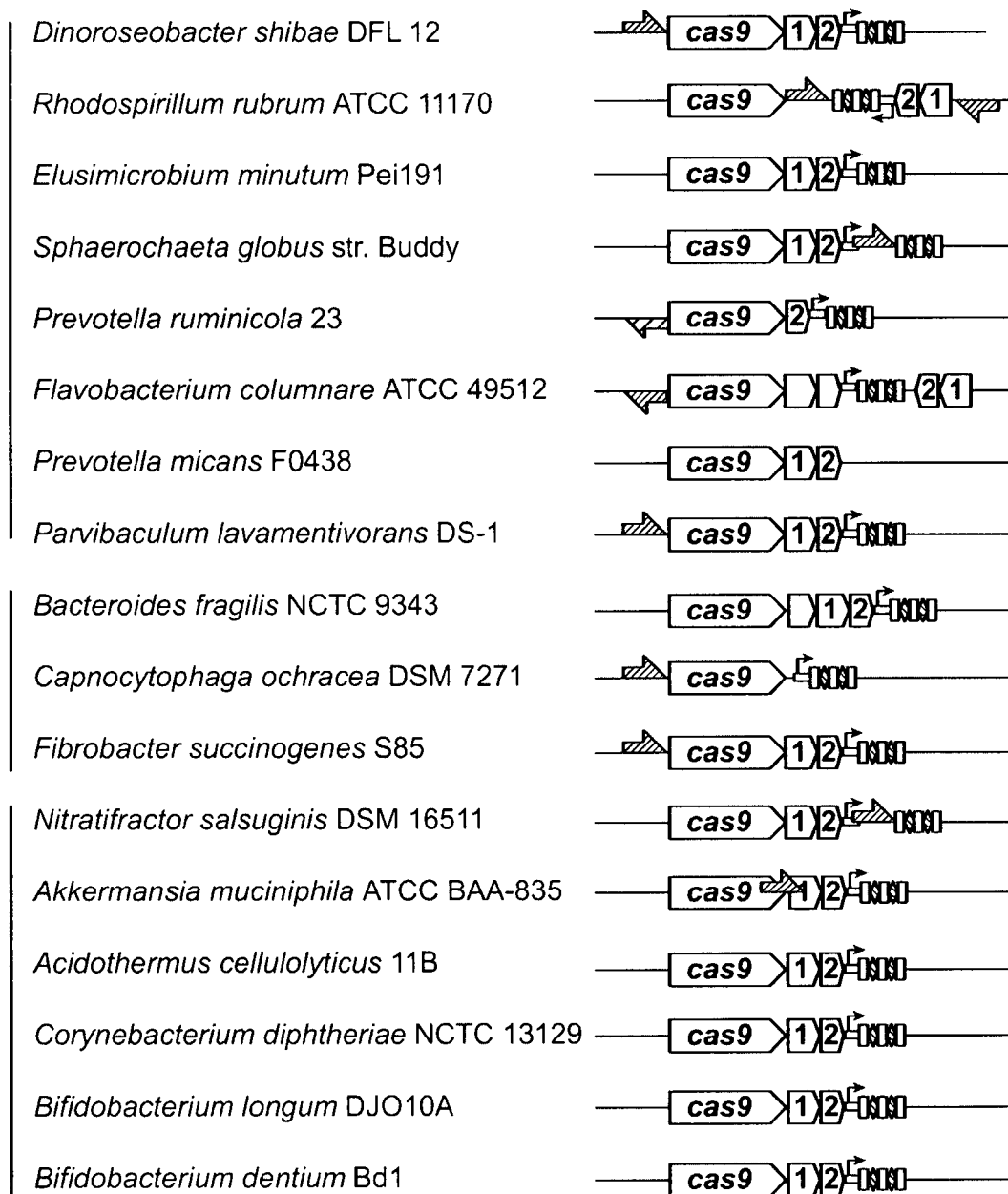


Fig. 34A

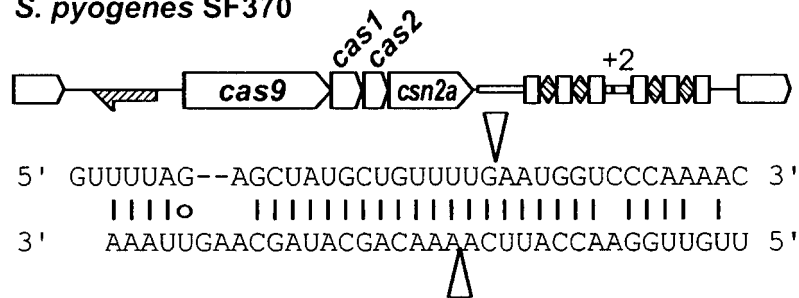
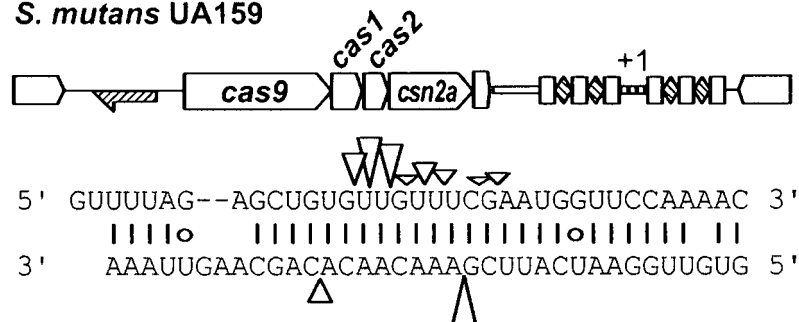
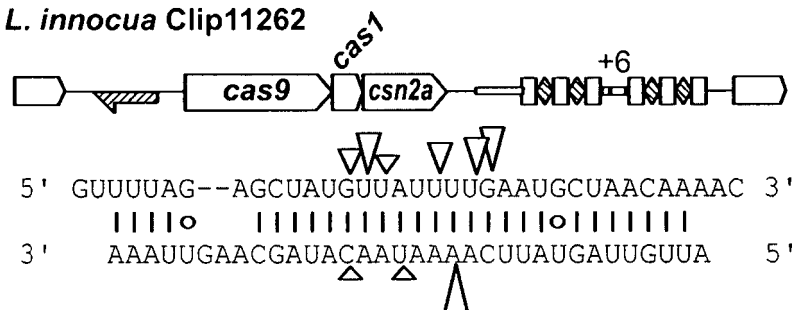
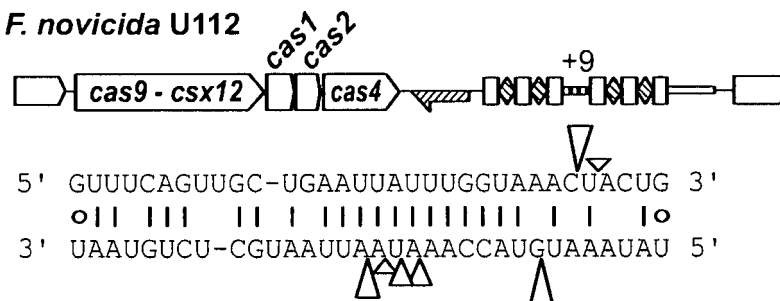
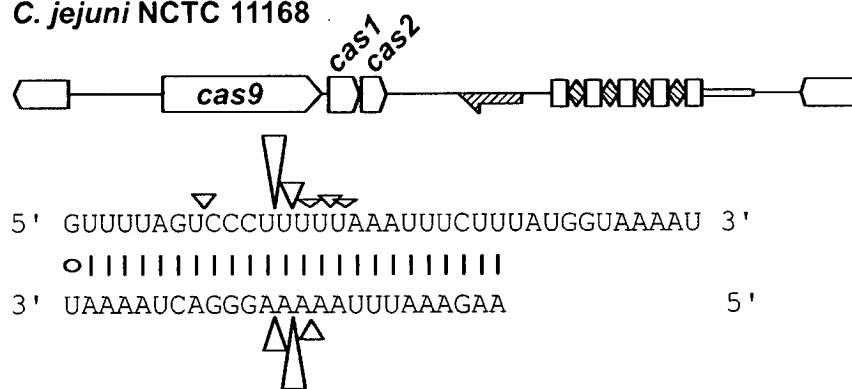
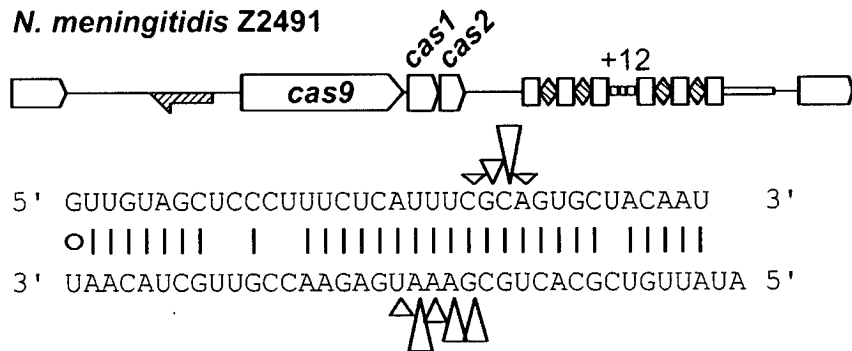
S. pyogenes* SF370**S. mutans* UA159*****L. innocua* Clip11262*****F. novicida* U112**

Fig. 34B

C. jejuni NCTC 11168*N. meningitidis* Z2491

Konsensus-Struktur	
<i>C. jejuni</i>
<i>F. novicida</i>	-----AAAGAAUUAUAAAGGGAUAUAAUAAAGA-----
<i>S. thermophilus2</i>	AUCUAAAUUAUAAUUAACCAAUAUUAUAAUGCUCUGUAUUAUUAUAAAGUAUUUUGA
<i>M. mobile</i>	-UGUA-AGGACGCGUUAACACAGUUAUUA-AAUCUUGCGAAGCUACAAAGAUAAAGGCU
<i>L. innocua</i>	-UGUAUUUUGAAAUACAGAUUGACAGUUAAGAAUAC-AUAAGAAUGAUACAUCACUAAAA
<i>s. pyogenes</i>	-----AUUGUUAAGUAUUCAAUUAACAUAGCAAGUUAUAAUAAAGGC-----
<i>s. mutans</i>	-----GUUGGAACCAUUCAAACAGCAUAGCAAGUUAUAAUAAAGGC-----
<i>s. thermophilus</i>	-----GUUGGAUUCAUUCCGAAACAACAACAGCAAGUUAUAAUAAAGGCAGUGAUUUUUAAU
<i>N. meningitidis</i>	UUGUGGUUGAAACCAUUCGAAACAACAACAGCGAGUUAUAAUAAAGGC-----
<i>P. multocida</i>	--ACAUUUUGUCGACUGCCGAAAUAGAGAACCGUUGCUUACAUAUAAAGGC--CGUCUGAAAAAG
	--GCAUUAUUGUUGCACUGCCGAAAUAGAGAGACCGUUGCUUACAUAUAAAGGC---UUCUGAAAAAG

FIG. 35

Konsensus-Struktur

(.....)

C. jejuni
F. novicida
*S. thermophilus*2
M. mobile
L. innocua
s. pyogenes
s. mutans
s. thermophilus
N. meningitidis
P. multocida

FIG. 35 (Fortsetz.)

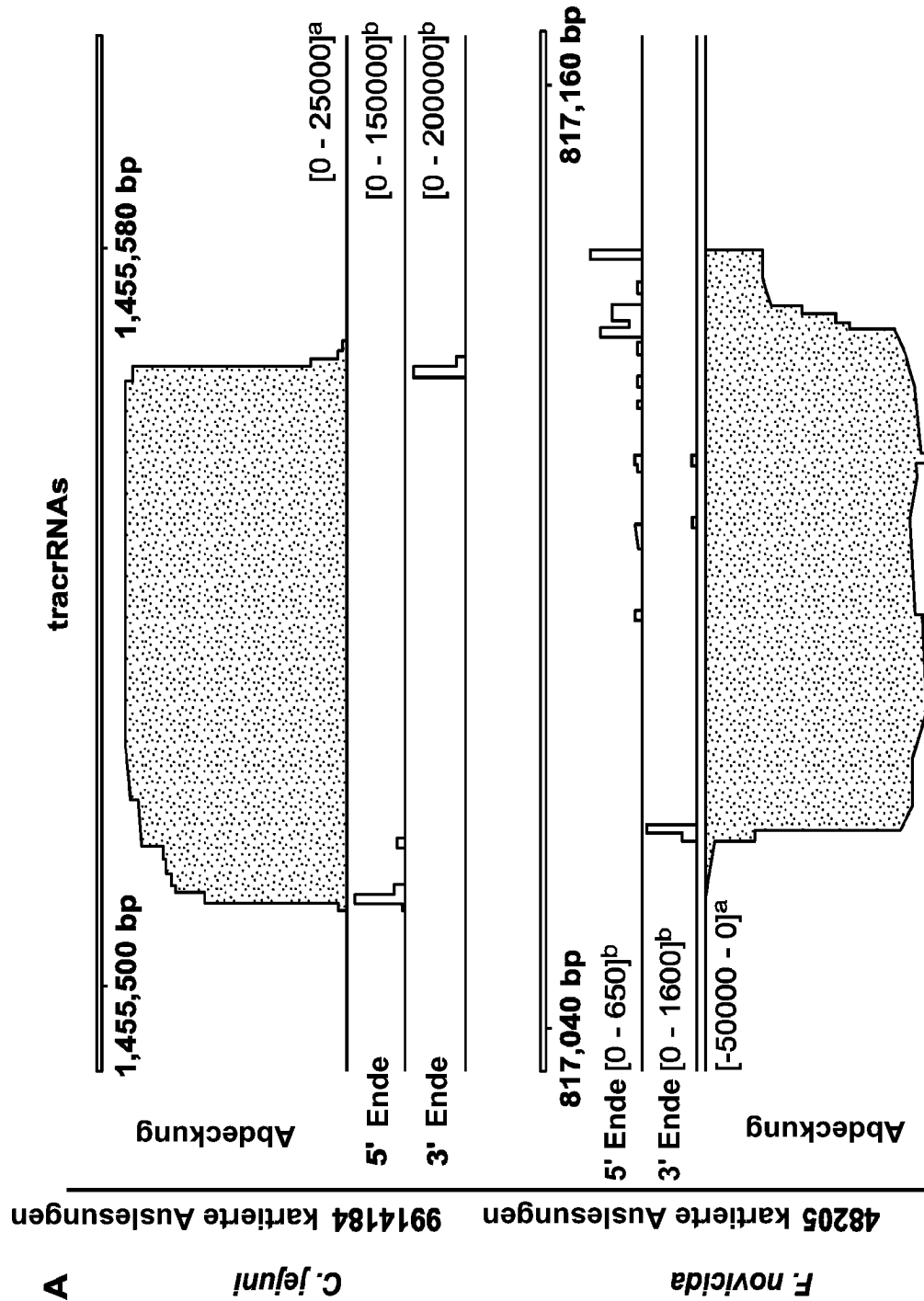


FIG. 36

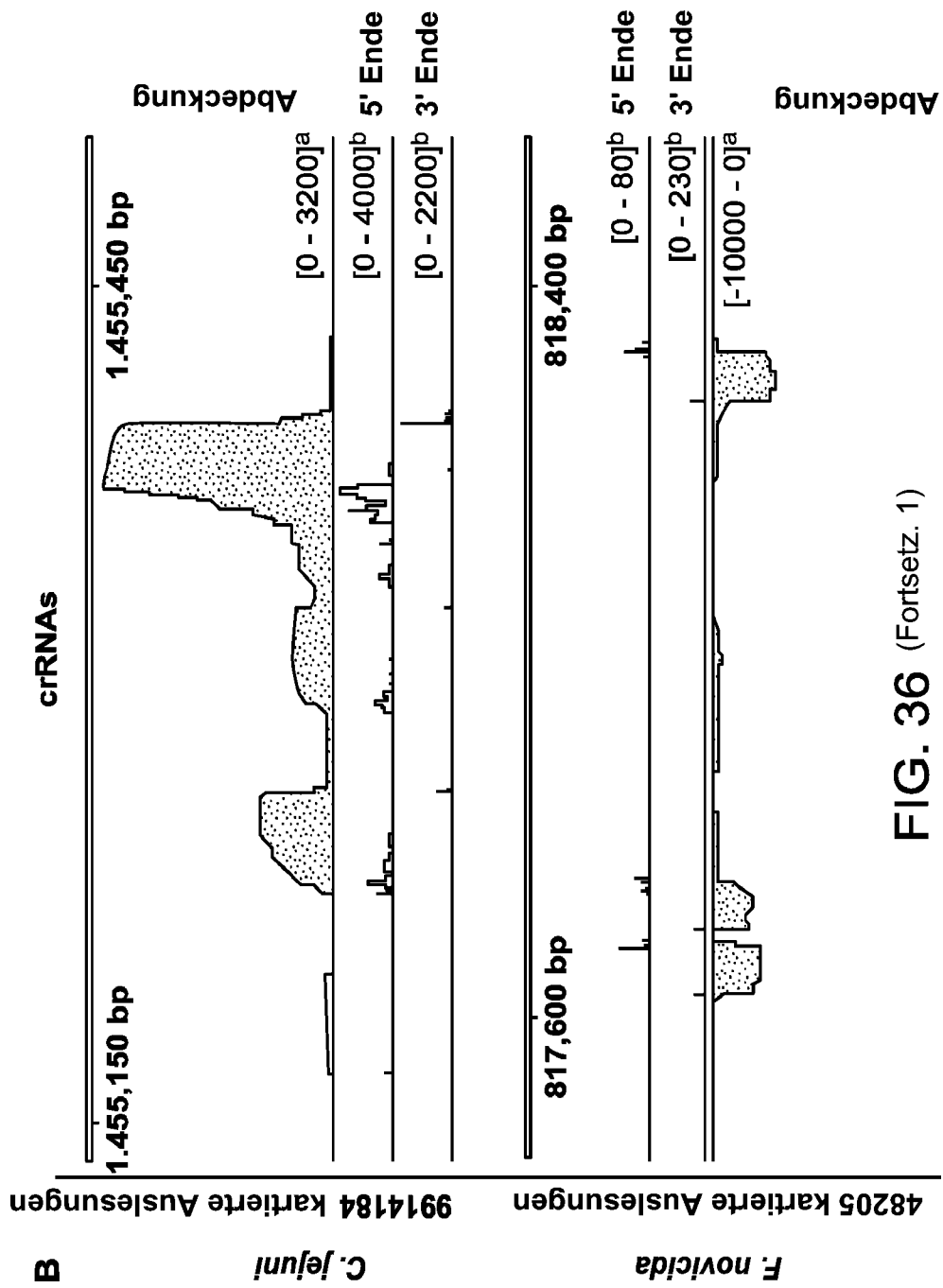


FIG. 36 (Fortsetz. 1)

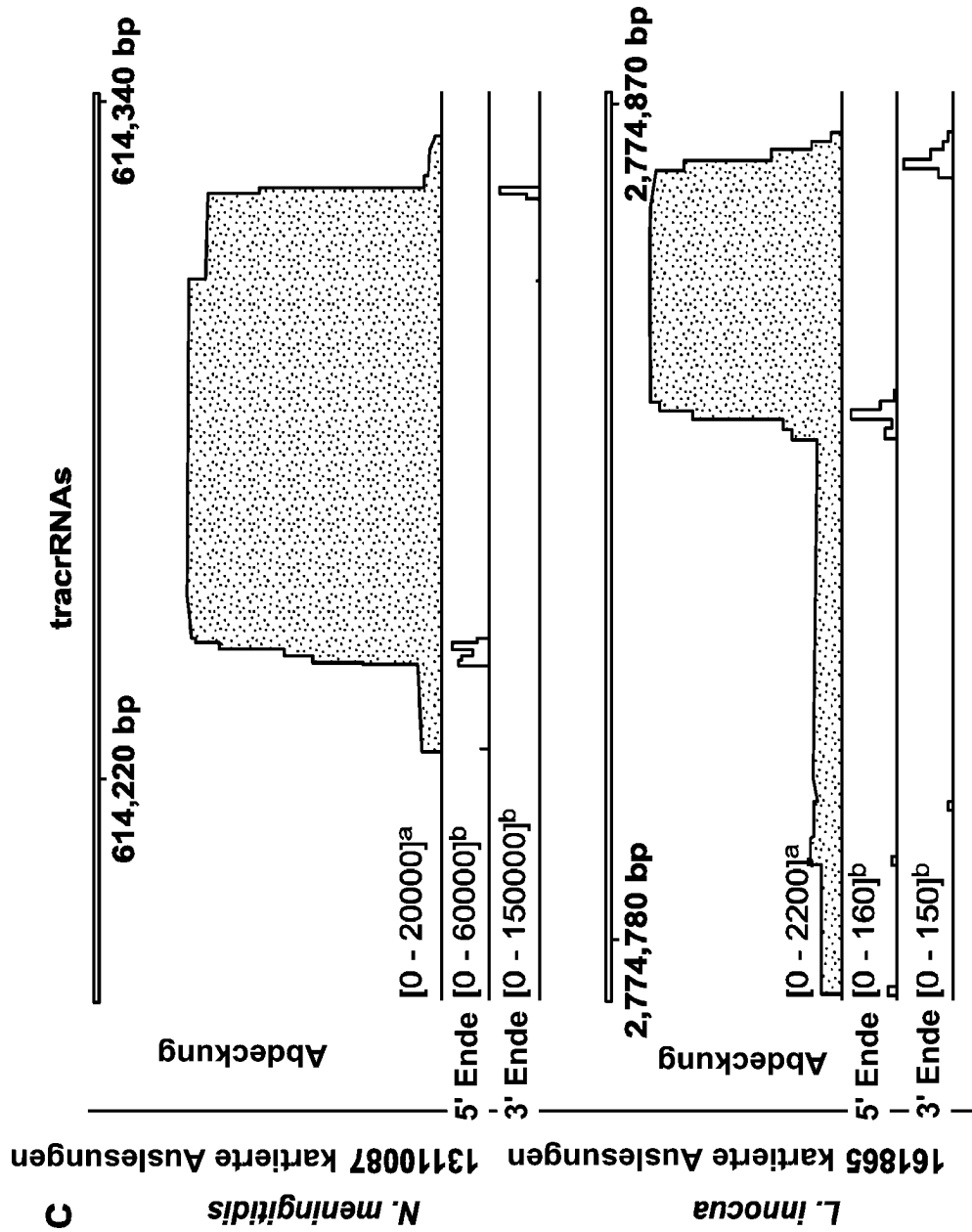
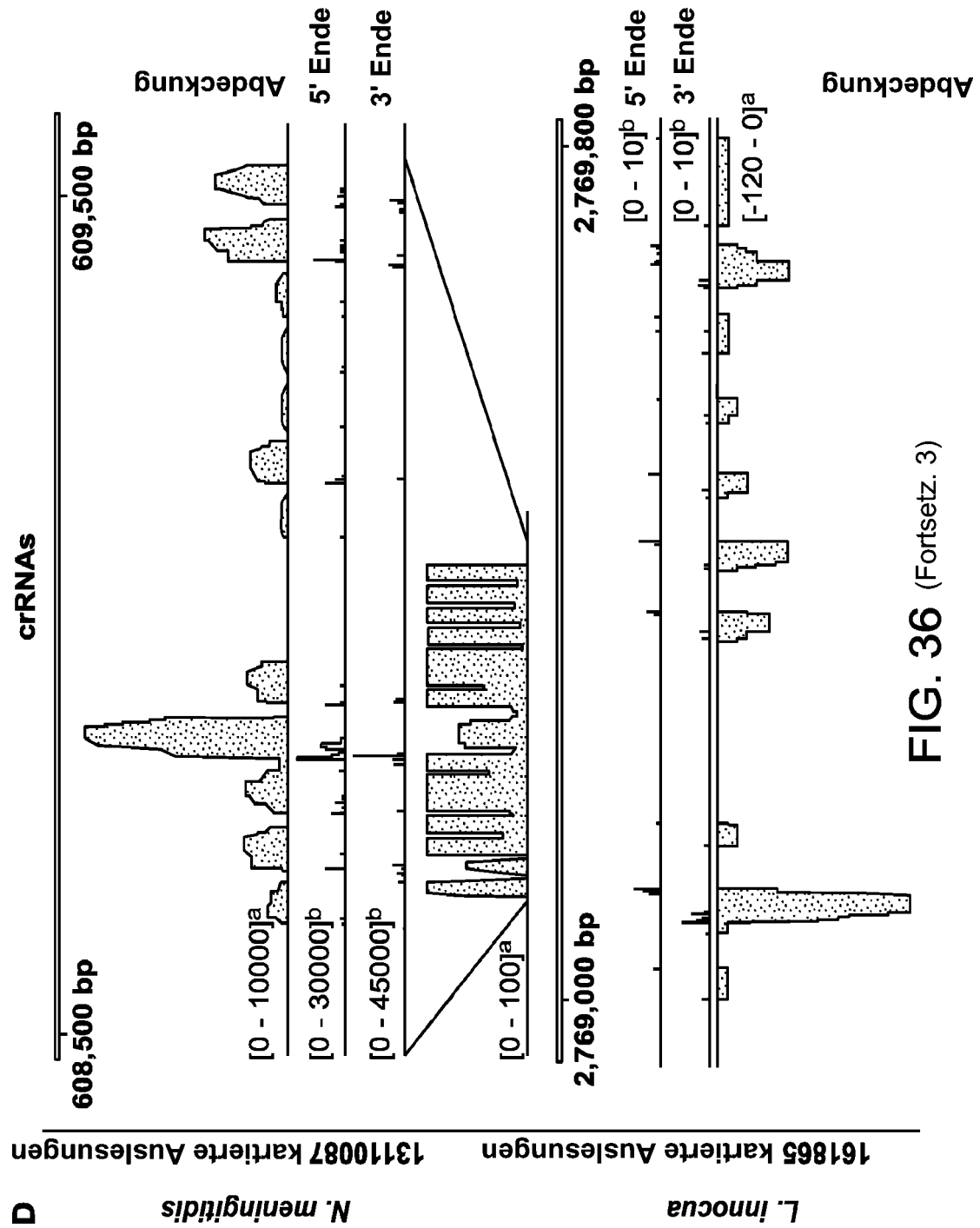


FIG. 36 (Fortsetz. 2)



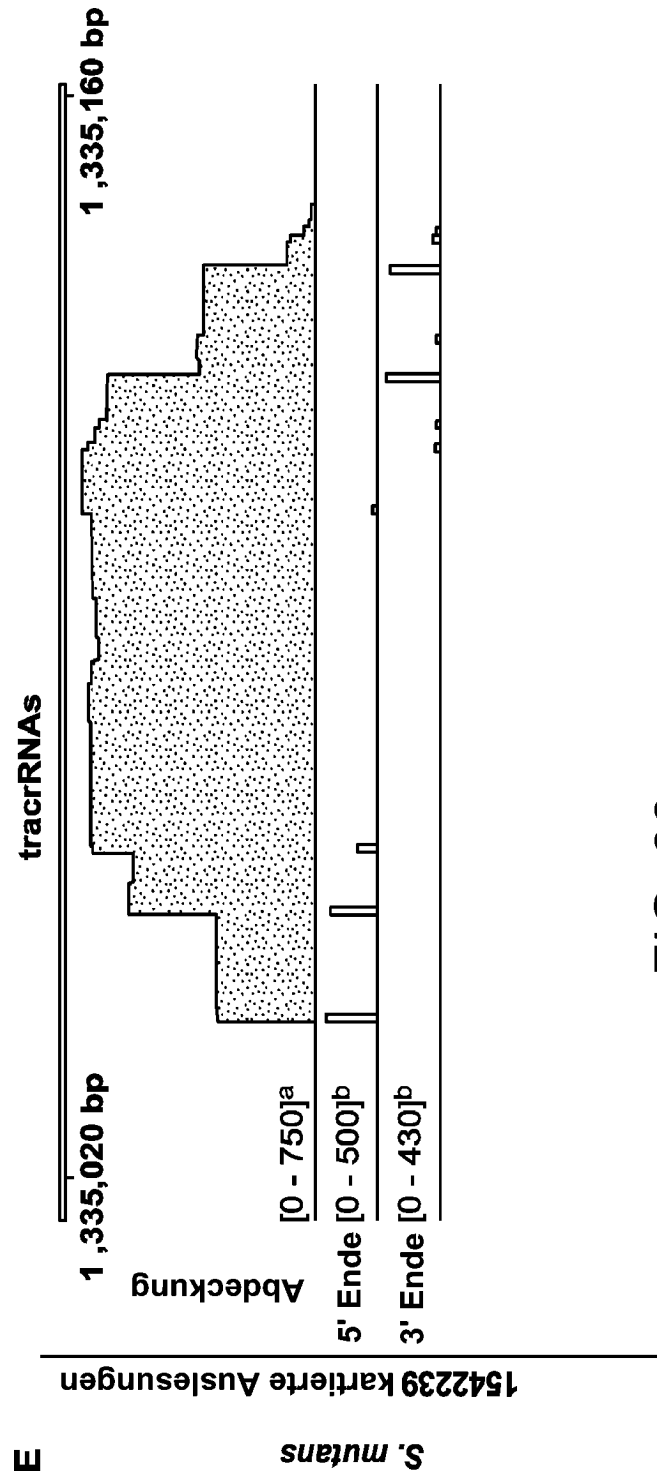


FIG. 36 (Fortsetz. 4)

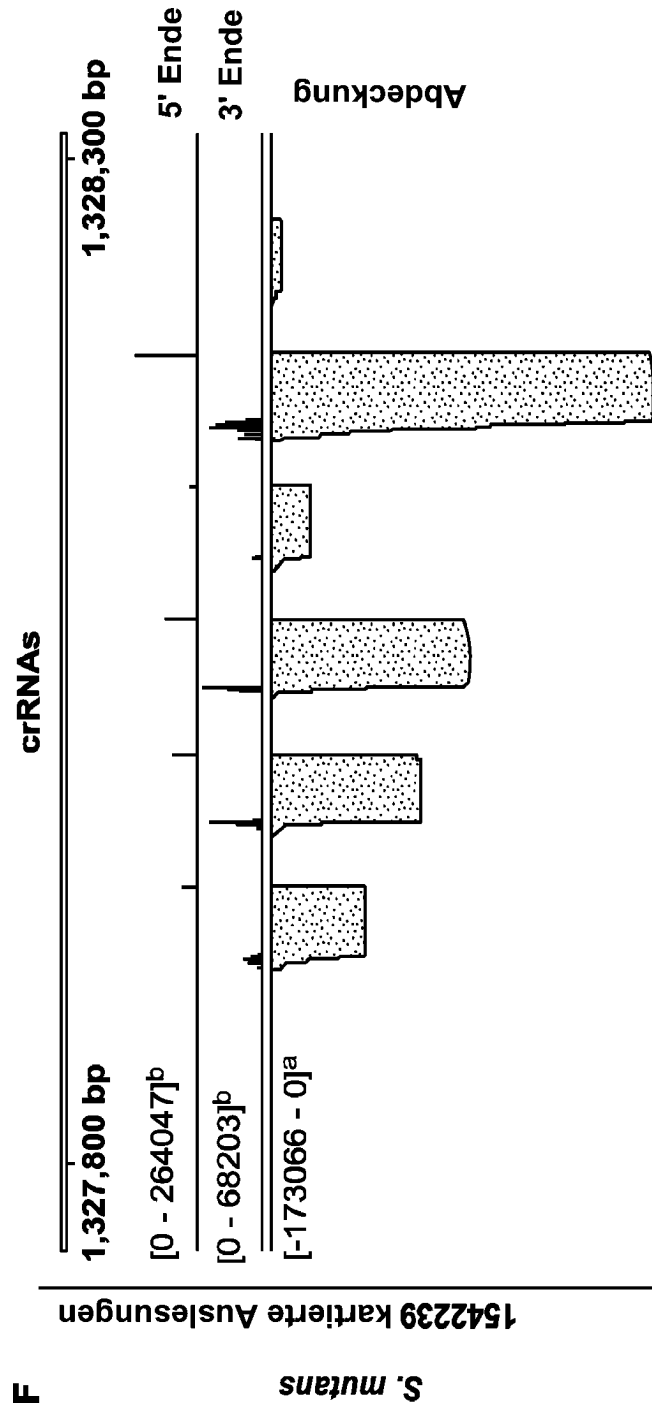


FIG. 36 (Cont. 5)

A

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
C. jejuni NCTC 11168 (NC_002163.1), insg. kartierte Reads : 9914184								
crRNA 4 Spacer	→	crRNA1	36 1455167 1455202	781	0,0079	AGTTT T T A A A A G A G C T T G G C G G T T G T T T T A G T C C C T T T T T	1455162 A 1 1455198 T 5 1455165 G 2 1455199 C 5 1455166 C 3 1455200 C 16 1455167 A 419 1455201 C 72 1455168 G 19 1455202 T 385 1455169 T 7 1455203 T 82 1455170 T 75 1455204 T 30 1455171 T 24 1455205 T 112 1455172 T 9 1455206 T 5 1455173 T 11 1455174 A 4	
		crRNA2	38 1455231 1455268 2658 29 1455240 1455268	2658	0,0268	C A A G T T T C A T T A G T T G A A T T T A A C T G T T T T A G T C C C T T T T T	1455227 T 4 1455264 T 3 1455228 T 1 1455265 C 14 1455230 C 37 1455266 C 60 1455231 C 206 1455267 C 240 1455232 A 72 1455268 T 1364 1455233 A 34 1455269 T 348 1455234 A 59 1455270 T 29 1455235 G 280 1455271 T 44 1455236 T 168 1455272 T 15 1455237 T 130 1455273 A 2 1455238 T 36 1455239 C 259 1455240 A 317 1455241 T 148 1455242 T 298 1455243 A 157 1455244 G 40	

FIG. 37

B

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
C. jejuni NCTC 11168 (NC_002163.1), insg. kartierte Reads : 9914184								
crRNA 4 Spacer	crRNA3	35	1455300 1455334	4729	0,0477	AAGAAATGAGGATGATGATATTTTACAGTTTTA GTCCCTTTTT	1455292 T 16 1455329 G 21 1455293 C 62 1455330 T 479 1455294 T 5 1455331 C 23 1455295 A 12 1455332 C 48 1455296 C 23 1455333 C 179 1455297 A 367 1455334 T 1226 1455298 A 143 1455335 T 220 1455299 G 563 1455336 T 54 1455300 A 1014 1455337 T 49 1455301 A 415 1455338 T 55 1455302 T 242 1455339 A 1 1455303 G 589 1455304 A 55 1455305 G 2	
	crRNA4	26	1455376 1455401	17404	0,1755	GTGTGCTAAAAAATGGACTTAAATGTTTTA GTCCCTTTTT	1455360 T 6 1455398 C 63 1455361 G 2 1455399 C 645 1455362 A 9 1455400 T 9844 1455363 G 261 1455401 T 1518 1455364 T 51 1455402 T 481 1455365 G 598 1455403 T 763 1455366 T 77 1455404 T 578 1455367 G 552 1455405 A 4 1455368 C 579 1455369 T 901 1455370 A 816 1455371 A 227 1455372 A 262 1455373 A 1072 1455374 A 1269 1455375 A 1991 1455376 A 2204 1455377 A 1978 1455378 T 371	

FIG. 37 (Fortsetz. 1)

C

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
C. jejuni NCTC 11168 (NC_002163.1), insg. kartierte Reads : 9914184								
tracrRNA →		65	1455502	833829	8,4105	AAAGAAATTTAAAAAGGGGACTTAAAAATAAAAGAGT	T	1 1455565 G 12371
		58	1455509	1455566		TTGCGGGGACTCTGCGGGGTTACAAATCCCTAA	A	31 1455566 C 713292
						AAACCGCTTTT	A	27 1455567 T 74594
							G	24 1455568 T 10412
							A	19 1455569 T 9580
							A	232 1455570 T 426
							A	435 1455571 A 91
							T	369 1455572 A 97
							T	253 1455573 A 542
							T	85
							A	65
							A	193
							A	33472
							A	615001
							A	131879
							G	16444
							G	9390
							G	1053

FIG. 37 (Fortsetz. 2)

D

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
F. novicida U112 (NC_008601.1), insg. kartierte Reads : 48205								
crRNA ← 13 Spacer	crRNA5	52	817556	817607	117	0,2427	ATAACTCGACCAATATTGACAAAGTTTCAGT TGCTGAATTATTGGTAAACCT	817612 T 10 817557 A 3 817611 T 10 817556 C 55 817607 A 53 817555 T 13 817606 T 7 817553 C 12 817552 T 1
	crRNA6	56	817627	817682	116	0,2406	GGCAGGTTTTGTATGGTCATATAGGAGTGT CAGTTGCTGAATTATTGGTAAACCT	817685 A 4 817629 A 4 817682 G 23 817628 A 3 817681 G 13 817627 C 53 817679 A 2 817626 T 20 817677 G 4
	crRNA7	66	817699	817764	11	0,0228	AGCTATAGGGTTACCTATCTTTGAGTGTGG CAAAATAAGTTTCAGTTGCTGAATTATTGGTA AACCT	817764 A 4 817699 A 2 817763 G 1 817754 T 4 817745 T 2
	crRNA9	53	817845	817897	24	0,0498	ATGCTTTTAACTACTGATATATACGTTTCAG TTGCTGAATTATTGGTAAACT	817897 A 3 817845 C 7
tracrRNA ←	tracrRNA1	74	817065	817138	2808	5,8251	GTACCAATAATAATGCTCTGTAATCAATTA AAAGTATTTTGAACGGACCTCTGTTTGACACG TCTGAATAACTAAAA	817140 A 2 817066 A 28 817139 T 2 817065 C 1523 817138 G 615 817064 T 440 817136 A 9 817063 A 7 817135 C 2 817062 A 10 817134 C 37 817061 A 19 817133 A 26 817060 A 14 817132 A 7 817059 A 6 817131 A 328 817058 G 5 817130 T 355 817057 C 32 817129 A 165 817055 A 10 817128 A 484 817127 T 24 817126 T 55 817125 A 32 817124 A 19
	tracrRNA2	64	817065	817128				

FIG. 37 (Fortsetz. 3)

E

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende		
N. meningitidis AZ2491 (NC_003116.1), insg. kartierte Reads : 13110087										
crRNA → 16 Spacer	crRNA1	48	608456	608503	1346	0,0103	TATCCATTCCCAGCCGGAAATTAAAGTTGTAGC TCCCTTTCTCATTTTCGCAGT	608453		
								T		
								30	608500	T
								105	608501	G
								77	608502	T
								64	608503	A
								259	608504	T
								4	608505	C
								21	608506	A
								118	608507	C
								175	608508	A
								39	608509	T
								608461		
								608517		
								608518		
								608519		
								608520		
								608521		
								608522		
								608523		
								608524		
608525										
crRNA2	50	608520	608569	685	0,0052	GCCTTTTTACAAGCTCGCTTTCTTTGTTGTAG CTCCCTTTCTCATTTTCGCAGT	608517			
							T			
							5	608564	T	
							41	608565	T	
							7	608566	T	
							61	608567	T	
							21	608568	C	
							21	608569	C	
							44	608570	C	
							21	608571	A	
							31	608572	G	
							17	608573	T	
							8	608574	T	
							16	608575	T	
							8	608576	T	
31	608577	C								
101	608578	C								
173	608579	C								
43	608580	A								
6	608581	G								
2	608582	T								

FIG. 37 (Fortsetz. 4)

F

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende				
N. meningitidis AZ2491 (NC_003116.1), insg. kartierte Reads : 13110087												
crRNA3	50	608586	608635	12402	0,0946	TAAAGGTTTCTGTTGCGACCCGAAATGTTGTAG	608583	T	5	608631	T	75
						CTCCCTTTCTCATTTTCGCAGT	608584	G	3	608632	T	114
							608585	G	513	608633	C	414
							608586	T	2243	608634	G	1510
							608587	A	2091	608635	C	2219
							608588	A	188	608636	A	297
							608589	A	233	608637	G	52
							608590	G	744	608638	T	8
							608591	G	1152			
						crRNA4	49	608653	608701	26361	0,2011	TAACTTTGACCGTGTGCAATCCAGTTAGTTGT
AGCTCCCTTTCTCATTTTCGCAGT	608647	C	127	608698	T							569
	608648	T	203	608699	C							1167
	608649	T	377	608700	G							4910
	608650	T	2751	608701	C							4968
	608651	A	477	608702	A							764
	608652	A	1334	608703	G							62
	608653	C	11694	608704	T							4
	608654	T	3665									
	608655	T	382									

FIG. 37 (Fortsetz. 5)

G

sRNA	Strang	Größe reife Form	interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5' Ende	Anz. Reads 3' Ende								
N. meningitidis AZ2491 (NC_003116.1), insg. kartierte Reads : 13110087																
crRNA5		49	608719	608767	28747	0,2193	<u>AACCCACTAAATTTTGCAAAATGCGGTGTAGC</u> <u>TCCCTTTCATTTTCGCAGT</u>	608717 G 20 608764 T 426								
								608718 C 1171 608765 C 1316								
								608719 A 8128 608766 G 3040								
								608720 A 116 608767 C 6431								
								608721 C 59 608768 A 1049								
								608722 C 420 608769 G 102								
								608723 C 1089 608770 T 11								
								608724 A 557								
								crRNA6		50	608784	608833	121014	0,9231	<u>TTTTTTTGTAAGTTTGAACGAGTTGTAG</u> <u>CTCCCTTTCATTTTCGCAGT</u>	608781 G 42 608830 T 2295
																608782 C 630 608831 C 7060
608783 C 10039 608832 G 13864																
608784 T 26026 608833 C 44472																
608785 T 11430 608834 A 6853																
608786 T 8248 608835 G 802																
608787 T 3648 608836 T 69																
crRNA7		52	608848	608899	24611	0,1877	<u>TTCGTTTCAGATAGCAAAACGCAGTAGTGTGT</u> <u>AGCTCCCTTTCATTTTCGCAGT</u>									608846 A 28 608896 T 369
																608847 A 631 608897 C 1114
																608848 T 11039 608898 G 3441
								608849 T 753 608899 C 9015								
								608850 C 239 608900 A 1558								
								608851 G 2205 608901 G 137								
								608852 T 1910 608902 T 8								
								608853 T 826								
								84	608916	608999	901	0,0069	<u>ATATGACGGTGGGCAACTGGTACAGGTTGTAG</u> <u>CTCCCTTTCATTTTCGCAGTGCTACAAATGCC</u> <u>GGATATGACGGTGGGCAACT</u>	608913 C 6 608996 A 6		
														608914 G 3 608997 A 2		
608915 G 2 608998 C 2																
608916 A 266 608999 T 259																
608917 T 17 609000 G 171																
608918 A 96 609001 G 24																
608919 T 26 609004 T 16																

FIG. 37 (Fortsetz. 6)

H

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
<i>N. meningitidis</i> AZ2491 (NC_003116.1), insg. kartierte Reads : 13110087								
crRNA10	49	609049	609097	5027	0,0383	CTTTTGTATTGATTCAAGGTGCTTGTGTAGC <u>TCCCTTTCTCATTTGCGAGT</u>	609046	T 89 609094 T 81
							609047	C 37 609095 C 218
							609048	G 647 609096 G 845
							609049	C 1614 609097 C 1237
							609050	T 583 609098 A 219
							609051	T 208 609099 G 36
							609052	T 104 609101 G 7
							609053	T 18
							609109	A 58 609160 T 540
							609110	G 331 609161 C 1263
crRNA11	52	609112	609163	22711	0,1732	ATTGTCGATGATGAAACTCGAGCATGTTGT <u>AGCTCCCTTTCTCATTTGCGAGT</u>	609111	T 99 609162 G 2850
							609112	A 10627 609163 C 4882
							609113	T 1234 609164 A 552
							609114	T 66 609165 G 52
							609115	C 191 609166 T 15
							609116	G 4906
							609117	T 136
							609175	G 4 609226 T 57
							609176	G 3 609227 C 189
							609177	G 42 609228 G 576
crRNA12	52	609178	609229	5067	0,0386	TAGCCAGTGCTAAAAACCGCACCCGCTTGTGT <u>AGCTCCCTTTCTCATTTGCGAGT</u>	609178	T 897 609229 C 1348
							609179	A 72 609230 A 321
							609180	G 584 609231 G 33
							609181	C 154 609232 T 2
							609182	C 314
							609183	A 38

FIG. 37 (Fortsetz. 7)

I

sRNA	Strang	Größe reife Form	interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
<i>N. meningitidis</i> AZ2491 (NC_003116.1), insg. kartierte Reads : 13110087								
crRNA13	51	609245	609295	4666	0,0356	ATAGAAATACATACGCCGAGTAATTAGTTGTA GCTCCCTTTCTCATTTTCGCAGT	609243	A 6 609292 T 65
							609244	A 62 609293 C 201
							609245	A 1311 609294 G 475
							609246	T 141 609295 C 1039
							609247	A 116 609296 A 229
crRNA14	51	609311	609361	7147	0,0545	TTTTTGTAATTGTTCTGCCTTTTCTTAGTTGTA GCTCCCTTTCTCATTTTCGCAGT	609248	G 44 609297 G 25
								609298 T 1
							609308	T 12 609358 T 136
							609309	T 22 609359 C 442
							609310	C 207 609360 G 858
crRNA15	50	609378	609427	49818	0,3800	ACGGCGAAACCATTTGCCACAAACGTTGTAG CTCCCTTTCTCATTTTCGCAGT	609311	T 1190 609361 C 2335
							609312	T 774 609362 A 540
							609313	T 577 609363 G 69
							609314	T 195 609364 T 5
							609315	T 37
crRNA15	50	609378	609427	49818	0,3800	ACGGCGAAACCATTTGCCACAAACGTTGTAG CTCCCTTTCTCATTTTCGCAGT	609375	C 319 609424 T 532
							609376	C 7253 609425 C 1414
							609377	C 7249 609426 G 2553
							609378	A 19015 609427 C 7448
							609379	C 547 609428 A 1854
crRNA15	50	609378	609427	49818	0,3800	ACGGCGAAACCATTTGCCACAAACGTTGTAG CTCCCTTTCTCATTTTCGCAGT	609380	G 307 609429 G 210
							609381	G 190 609430 T 10

FIG. 37 (Fortsetz. 8)

J

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
<i>N. meningitidis</i> AZ2491 (NC_003116.1), insg. kartierte Reads : 13110087								
	crRNA16	50	609444	609493 42398	0,3234	AATAATAACCCCAATACACGATGTTAAGTTGTA GCTCCCTTCTCATTTCCGAGT	609441 A 8 609442 A 548 609443 A 3686 609444 A 4638 609445 T 1428 609446 A 2861 609447 A 1322 609448 T 287	609490 T 579 609491 C 2197 609492 G 6466 609493 C 13217 609494 A 2438 609495 G 200 609496 T 12
tracrRNA →	tracrRNA1 tracrRNA2 tracrRNA3	163 100 82	614162 614225 614243	614324 208318 614324 614324	1,5890	TTGCTTTATTTATATACAAATAGATTATTGAC TTATCATCTCACAAACGGCTAGAAATCCCAACA TATTGTCGCACTCGAAATGAGAACCGTTGCT ACAATAAGGCCGCTGTGAAAAGATGTGCCGCAA CGCTCTGCCCCCTTAAAGCTTCTGCTTTAAGGG GCATCGTTTAT	614158 C 1 614161 A 1 614162 T 2 614164 G 1 614223 A 596 614224 C 39 614225 A 14761 614226 T 169 614227 A 276 614239 C 208 614240 G 42037 614241 A 39965 614242 A 22890 614243 A 51186 614244 T 17711 614245 G 4192 614246 A 488	614319 G 311 614320 G 109 614321 G 258 614322 G 150 614323 C 41244 614324 A 128531 614325 T 167 614326 C 4197 614327 G 36 614328 T 122 614329 T 865 614330 T 1132 614331 A 1061 614332 T 2788 614333 T 3758 614334 T 77 614335 C 18 614336 G 3 614337 G 3 614338 T 1

FIG. 37 (Fortsetz. 9)

K

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
<i>L. innocua</i> Clip11262 (NC_003212.1), insg. kartierte Reads : 161865 (Beachte: schlechte Qualität der RNA-Bibliothek)								
crRNA	←	35	2769606 2769640	2	0,0012	GGTAACTTTGCCTAGGATAGTTTATAGAGCTAT	2769640 G	1 2769606 T
10 Spacer						GTT		
crRNA2		22	2769540 2769561	2	0,0012	CATTATGTTTTAGAGCTATGTT	2769561 C	1 2769540 T
							2769560 A	1
crRNA3		24	2769468 2769491	3	0,0019	GAGTTTTAGAGCTATGTTATTTTG	2769491 G	3 2769468 G
crRNA4		27	2769402 2769428	7	0,0043	TTATAGTTTTAGAGCTATGTTATTTTG	2769428 T	5 2769407 A
							2769427 T	2 2769406 T
								2769405 T
								2769403 T
								2769402 G
crRNA5		26	2769337 2769362	5	0,0031	TAAATGTTTTAGAGCTATGTTATTTT	2769362 T	3 2769339 T
							2769360 A	2 2769337 T
crRNA8		23	2769142 2769164	2	0,0012	TACAAAGTTTTAGAGCTATGTTAT	2769164 T	1 2769143 A
							3769163 A	1 2769142 T
crRNA9		30	2769072 2769101	19	0,0017	TTCAATGTTGTTTTAGAGCTATGTTATTTTG	2769101 T	6 2769079 T
							2769100 T	1 2769078 T
							2769099 C	4 2769075 T
							2769098 A	3 2769073 T
							2769097 T	3 2769072 G
							2769096 G	2
crRNA10		28	2769000 2769027	19	0,0017	GTTTTAGAGCTATGCTATTTTCGAATACT	2769027 G	1 2769000 T

FIG. 37 (Fortsetz. 10)

sRNA	Strang	Größe reife Form	interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads						
							5'-Ende	3'-Ende					
L. innocua Clip11262 (NC_003212.1), insg. kartierte Reads : 161865 (Beachte: schlechte Qualität der RNA-Bibliothek)													
tracrRNA →	tracrRNA1	90	2774774	2774863	367	0,2267	ATTGTTAGTATTCAAAATAACATAGCAAGTTA	2774774	A	34	2774861	T	2
	tracrRNA2	76	2774788	2774863			AAATAAGGCTTTGTCCGTTATCAACTTTTAAT	2774783	A	1	2774862	T	47
	tracrRNA3	68	2774796	2774863			TAAAGTAGCGCTGTTTCGGCGCTTTTTTT	2774786	C	1	2774863	T	150
								2774787	A	1	2774864	T	67
								2774788	A	22	2774865	T	30
								2774794	C	1	2774866	G	15
								2774795	A	1	2774867	T	6
								2774796	T	5			
								2774797	A	2			
								2774799	C	5			
								2774801	A	1			

FIG. 37 (Fortsetz. 11)

M

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
S. mutans UA159 (NC_004350.2), insg. kartierte Reads : 1542239								
crRNA 5' Spacer	←	38	1328162	1328199	267104	17,3192	GCCATTAAATTAATATGGTGAGTGGTTTATAGAGCTG <u>TGTTGTTTCGA</u>	1328201 A 8 1328166 G 18547
								1328200 C 178 1328165 T 41345
								1328199 G 264047 1328164 T 53386
								1328198 C 191 1328163 G 9084
								1328197 C 167 1328162 T 59197
								1328196 A 117 1328161 T 28333
								1328160 T 4240
								1328159 C 18236
								1328158 G 26742
								1328157 A 5573
crRNA2		36	1328098	1328133	26578	1,7233	GCTAGCGCAGTTAGTGCTCTGTTTATAGAGCTG <u>TGTTGTTTCGA</u>	1328156 A 17
								1328135 C 4 1328101 T 37
								1328134 A 13 1328100 G 800
								1328133 G 25656 1328099 T 6395
								1328132 C 212 1328098 T 11256
								1328131 T 25 1328097 G 301
								1328130 A 62 1328096 T 1453
								1328129 G 25 1328095 T 1755
								1328094 T 447
								1328093 C 1302
								1328092 G 1996
								1328091 A 670
								1328090 A 5

FIG. 37 (Fortsetz. 12)

N

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
S. mutans UA159 (NC_004350.2), insg. kartierte Reads : 1542239								
crRNA3	34	1328034	1328134	138067	8,9567	TGTTGTGTCATCATAGTTAGGTTTTTAGAGCTG TGTTGTTTCGA	1328069	G 4 1328036 G 2880
							1328068	C 66 1328035 T 733
							1328067	T 134361 1328034 G 68203
							1328066	G 609 1328033 T 37212
							1328065	T 321 1328032 T 24528
								1328031 G 1003
								1328030 T 889
								1328029 T 610
								1328028 T 124
								1328027 C 371
crRNA4	35	1327967	1328001	104705	6,7892	CAATTAGACAAATAGACAAACGTTTTAGAGCTG TGTTGTTTCGA	1328003	T 8 1328969 T 348
							1328002	T 36 1328968 G 8606
							1328001	C 101017 1328967 T 59863
							1328000	A 681 1328966 T 25845
							1327999	A 216 1328965 G 756
								1328964 T 3053
								1328963 T 1433
								1328962 T 1602
								1328961 C 864
								1328960 G 839
								1328959 A 322

FIG. 37 (Fortsetz. 13)

O

sRNA	Strang	Größe reife Form	Interess. Region	Reads	Abdeck. (%)	Sequenz	Anz. Reads 5'-Ende	Anz. Reads 3'-Ende
S. mutans UA159 (NC_004350.2), insg. kartierte Reads : 1542239								
	crRNA5	37	1327899 1327935	63999	4,1497	TTCGGACATGACTTGCCACAGTTTATAGAGCTG TGTTGTTATCGA	1327940 A 4 1327937 T 6 1327936 A 17 1327935 T 62587 1327934 T 1029 1327933 C 108 1327932 G 19 1327931 G 36	1327902 G 902 1327901 T 3296 1327900 T 12381 1327899 G 19843 1327898 T 2864 1327897 T 14180 1327896 T 2061 1327895 C 2079 1327894 G 5259 1327893 A 630
tracrRNA →	tracrRNA1 tracrRNA2 tracrRNA3	102 88 80	1335040 1335141 1335054 1335141 1335062 1335141	1299	0,0842	GTGGAAATCATTCGAAACAACACAGCAAGTTA AAATAAGGCAGTGATTTTAAATCCAGTCCGTA CACAACTTGAAAAAGTGCGCACCGATTCCGGTG CTTTTATTATTT	1335038 G 1 1335040 G 466 1335041 T 6 1335042 T 1 1335051 T 4 1335053 G 3 1335054 A 415 1335055 A 2 1335057 C 1 1335058 A 1 1335062 C 186 1335063 A 15	1335140 T 13 1335141 T 64 1335142 T 29 1335143 A 13 1335144 T 6 1335145 T 13 1335146 T 1 1335149 T 1

FIG. 37 (Fortsetz. 14)

A

Cluster	SEQ ID NO:
1	2, 3, 4, 5, 6, 7, 8, 15, 23, 24, 25, 36, 37, 38, 39, 41, 71, 74, 105, 116, 136, 138, 166, 177, 180, 183, 193, 204, and 232
2	83, 75, 156, 96, 121, 235, 208, 127, 182, 134, 119, 246, 153, 202
3	101, 168, 48, 226, 216, 210, 120, 102, 176, 57, 108, 79, 1, 245
4	219, 135, 53, 62, 240, 165, 217, 82, 212, 19, 40, 18, 194
5	84, 21, 150, 221, 111, 76, 47, 59, 77, 112, 198, 147
6	90, 91, 214, 92, 152, 98, 243, 197, 32, 227, 162
7	103, 187, 223, 151, 158, 126
8	88, 167, 13, 164, 184, 123
9	58, 73, 195, 148, 31, 33
10	206, 188, 211, 161, 205, 44

Cluster	SEQ ID NO:
11	50, 54, 78, 106, 174
12	209, 220, 146, 157
13	70, 154, 100, 117
14	128, 144, 118, 129
15	131, 66, 149, 145
16	89, 169, 163
17	141, 49, 72
18	196, 114, 86

Cluster	SEQ ID NO:
19	55, 27, 215
20	228, 234
21	160, 213
22	207, 237
23	230, 94
24	200, 247
25	133, 143
26	64, 68
27	20, 45
28	60, 56
29	99, 52

FIG. 38

B

Cluster	SEQ ID NO:	Cluster	SEQ ID NO:	Cluster	SEQ ID NO:
30	244, 185	45	233	59	218
31	43	46	122	60	65
32	189	47	16	61	171
33	170	48	242	62	97
34	11	49	203	63	63
35	107	50	26	64	46
36	14	51	137	65	225
37	236	52	199	66	10
38	12	53	34	67	173
39	17	54	201	68	51
40	239	55	178	69	142
41	61	56	42	70	69
42	85	57	190	71	28
43	191	58	81	72	139
44	22			73	80
				74	172
				75	115
				76	229
				77	175
				78	181

FIG. 38 (Fortsetz.)

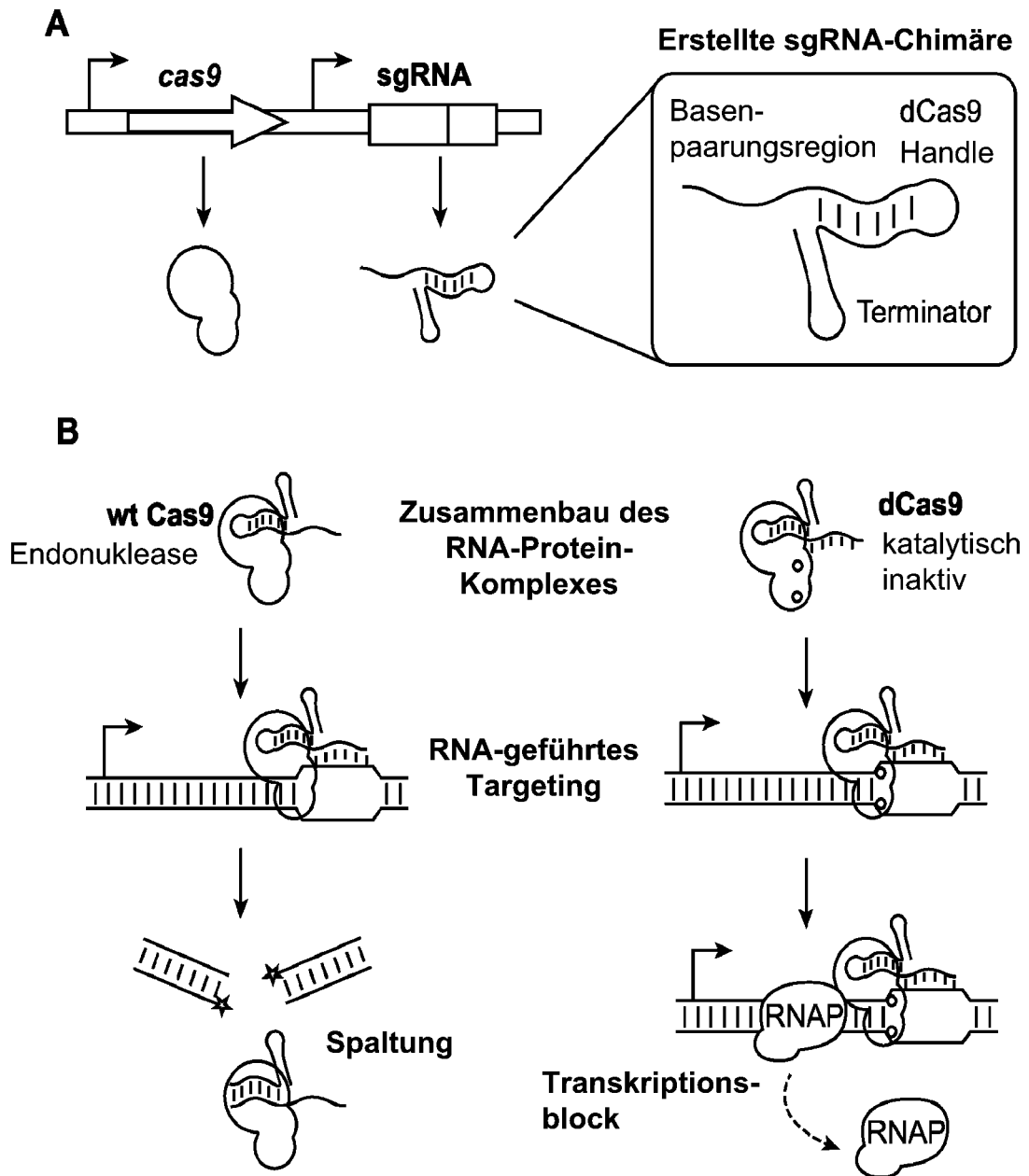


FIG. 39

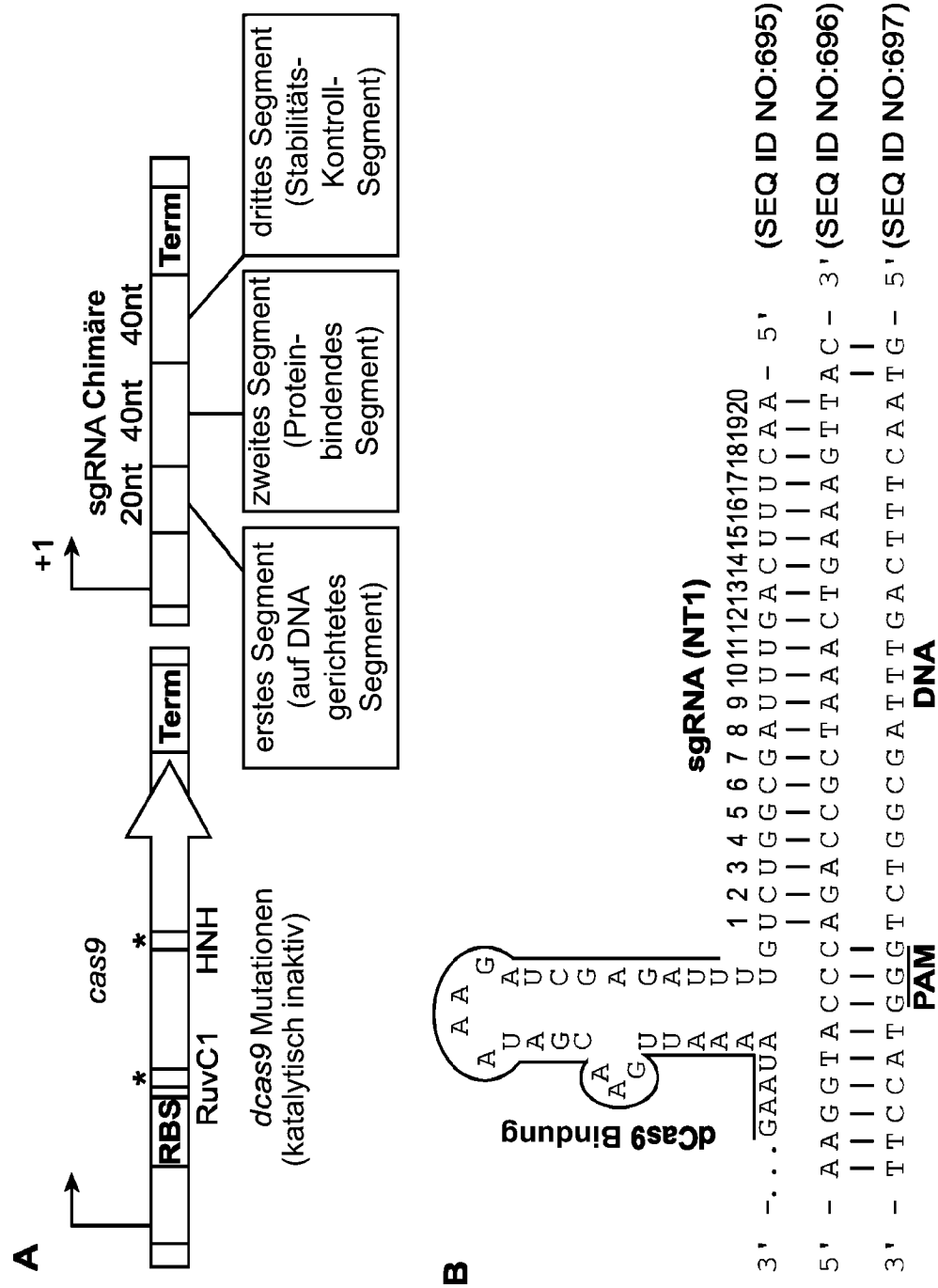


FIG. 40

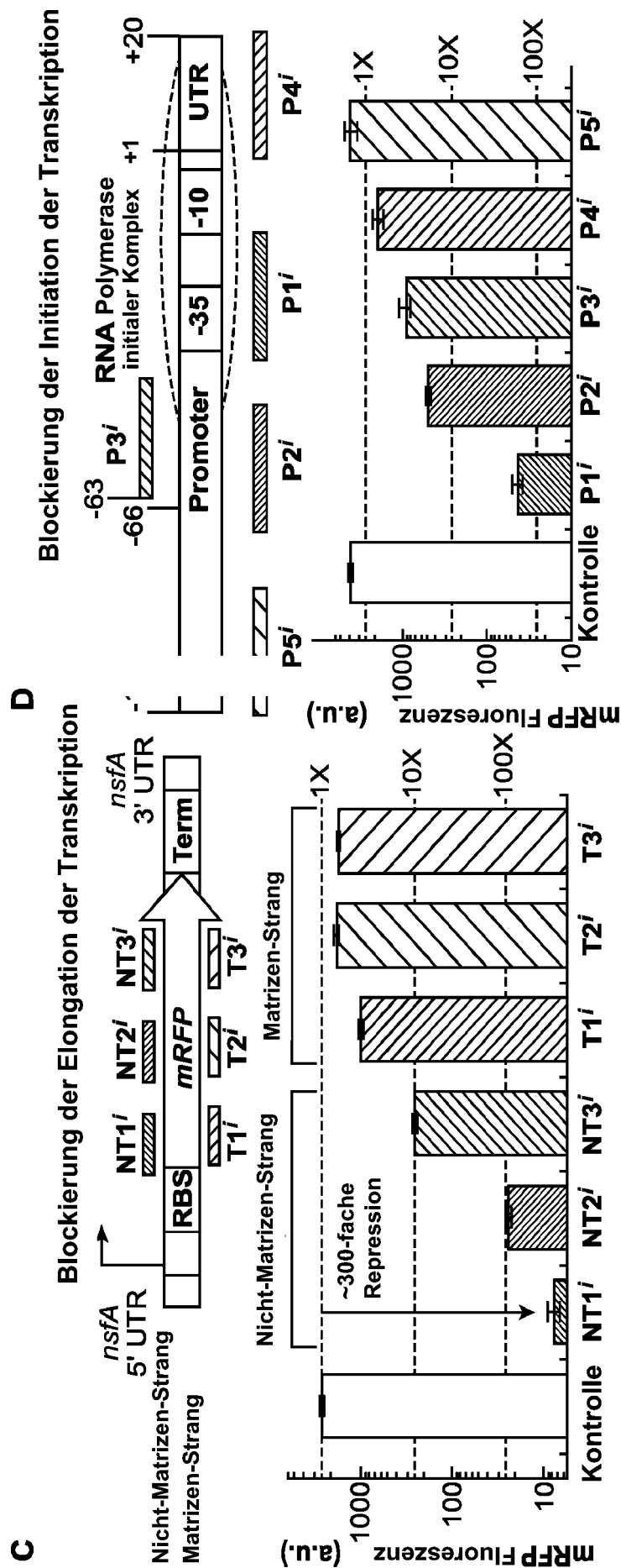


FIG. 40 (Fortsetz. 1)

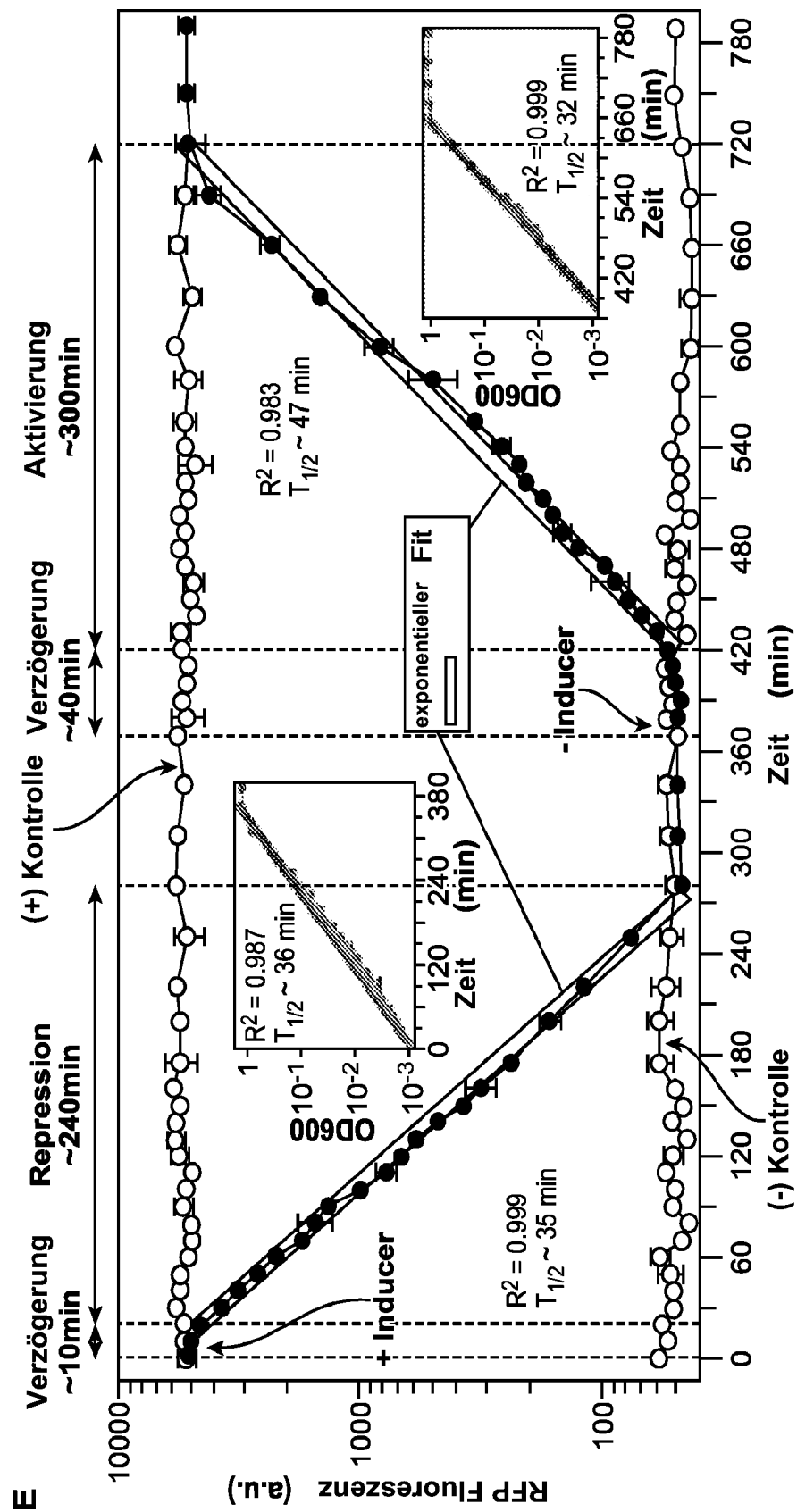


FIG. 40 (Fortsetz. 2)

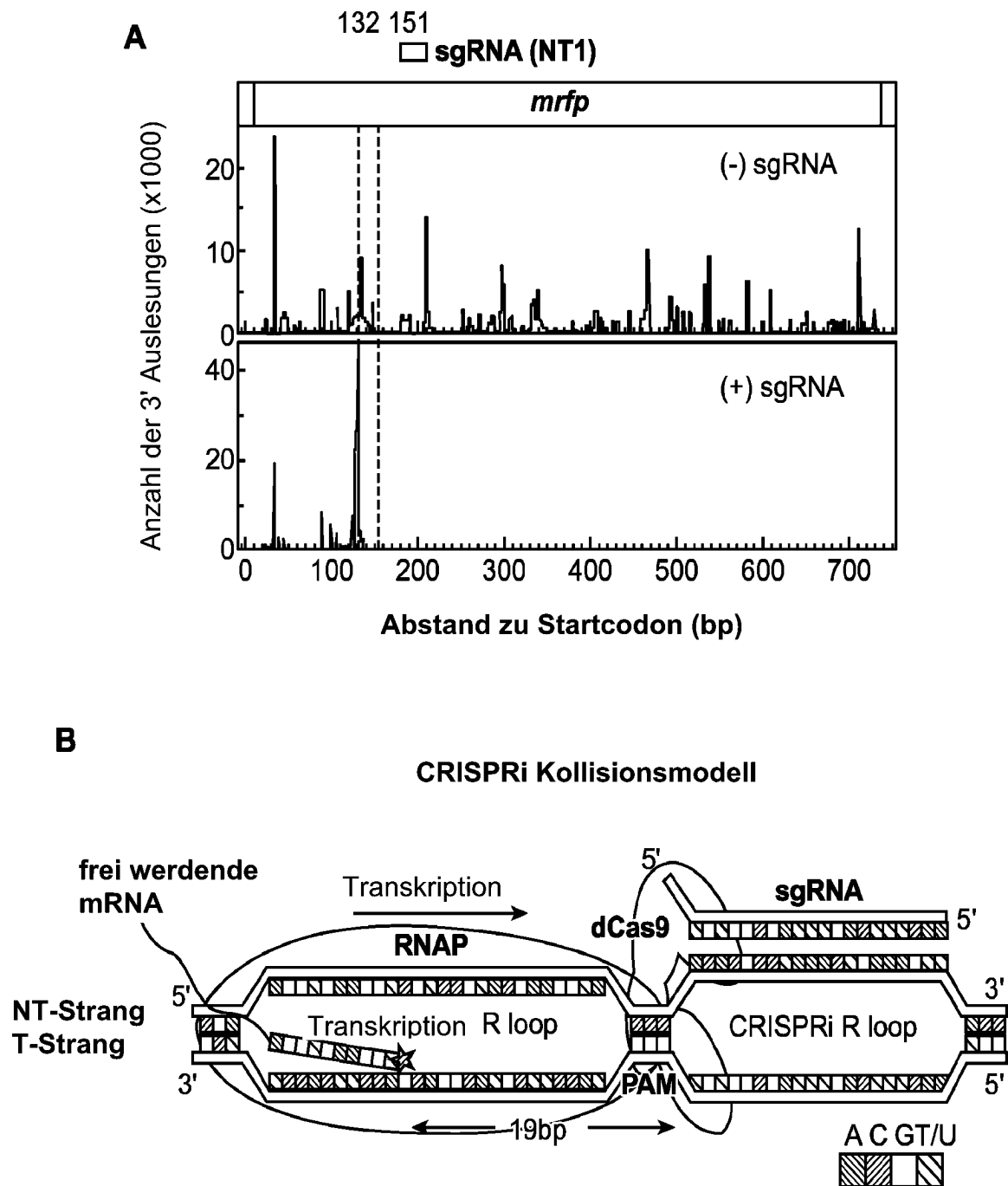


FIG. 41

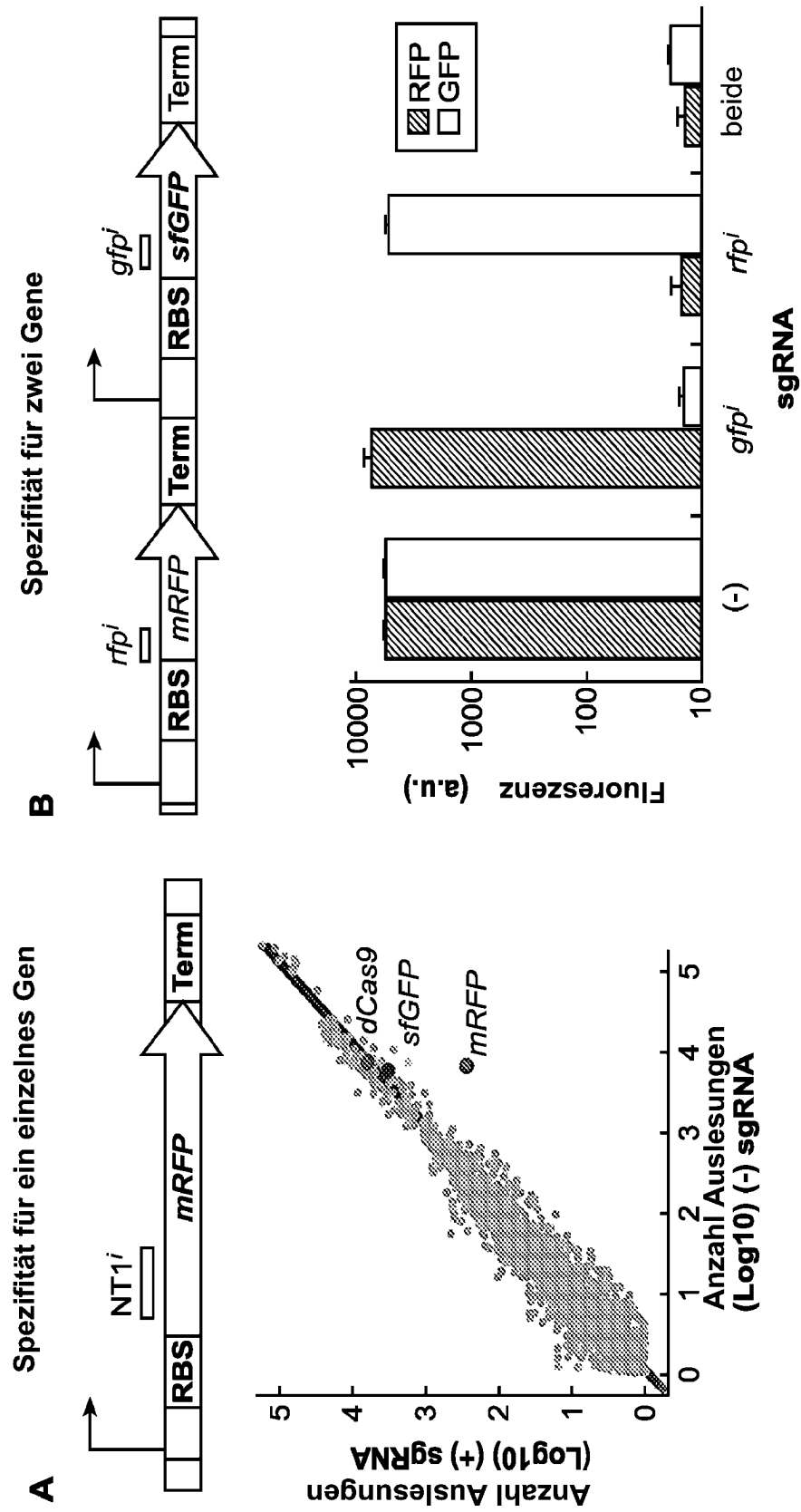


FIG. 42

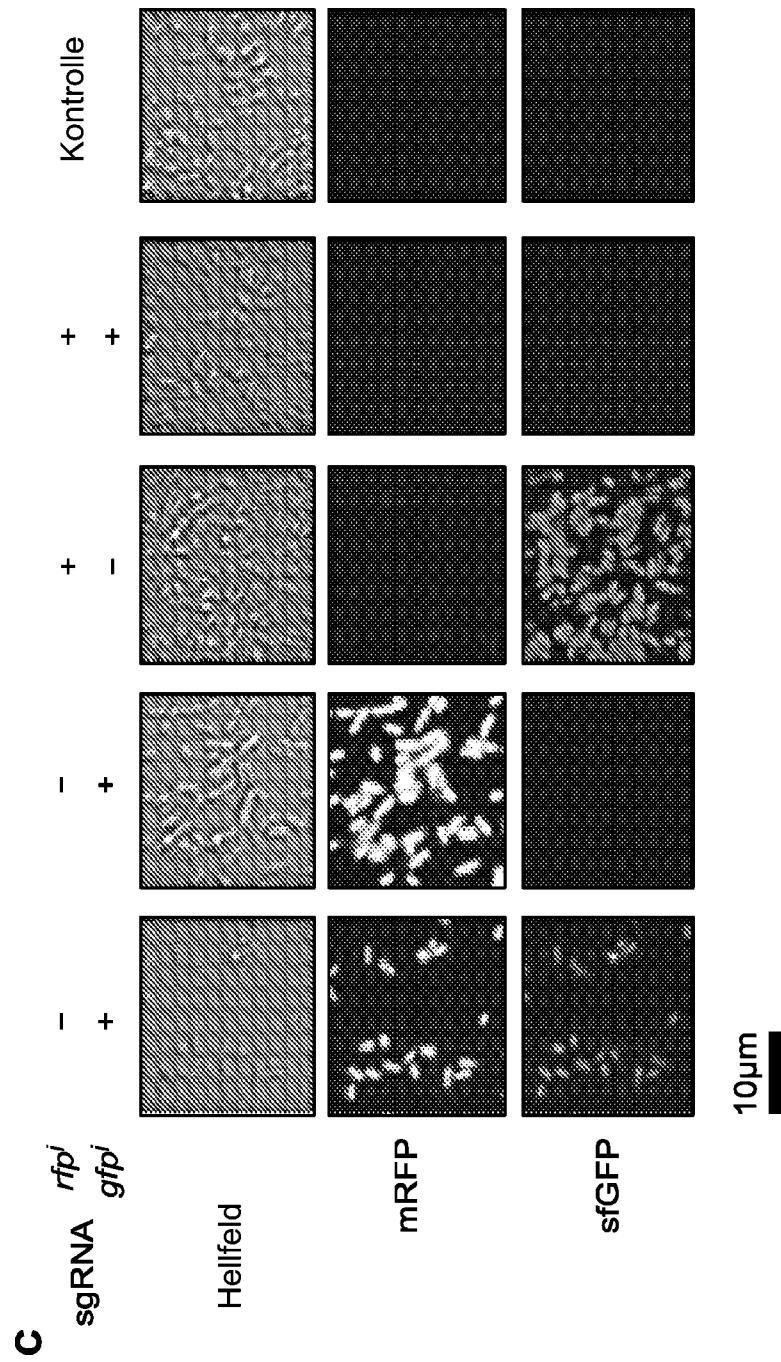


FIG. 42 (Fortsetz.)

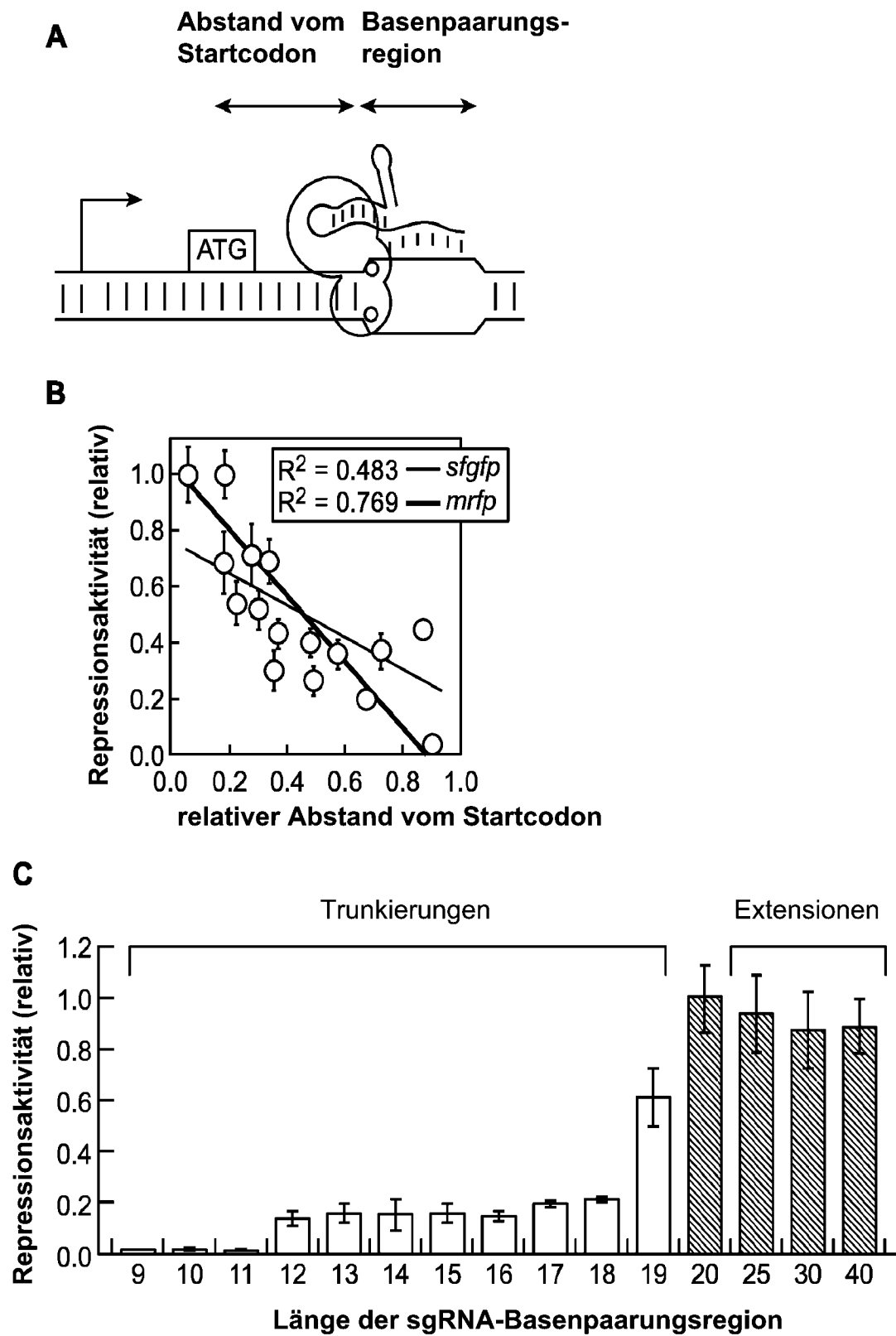


FIG. 43

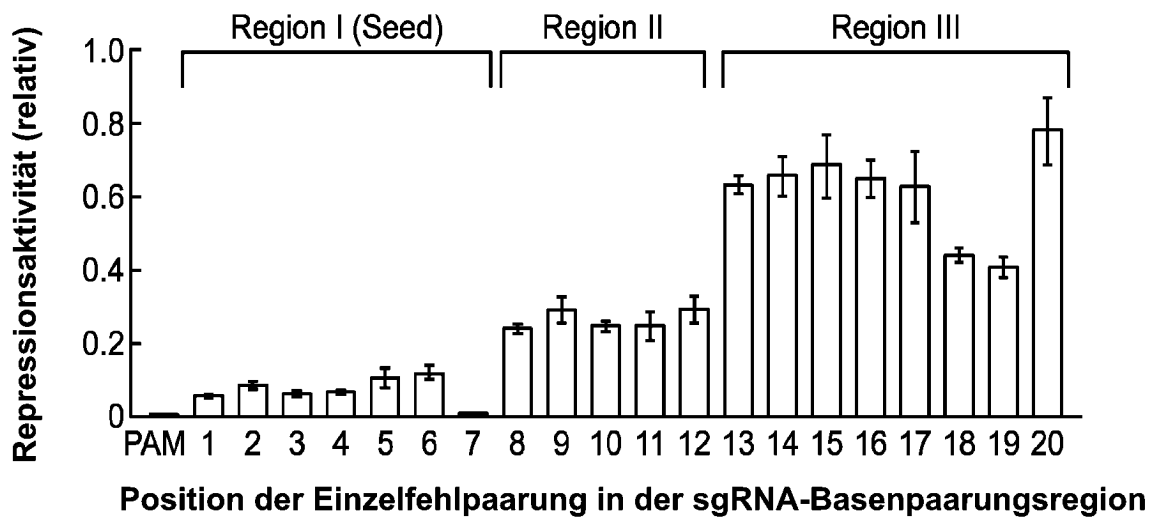
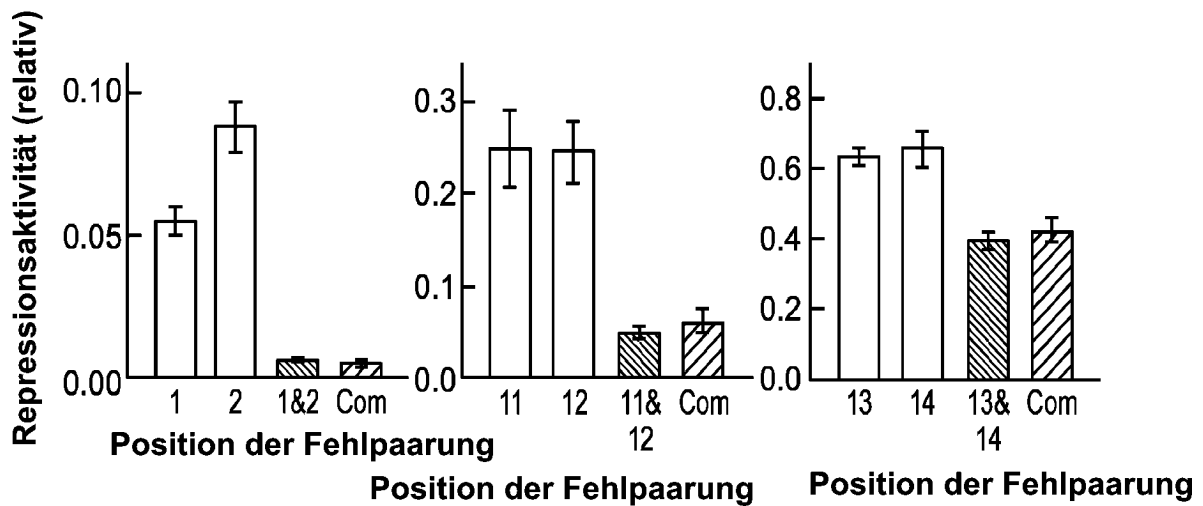
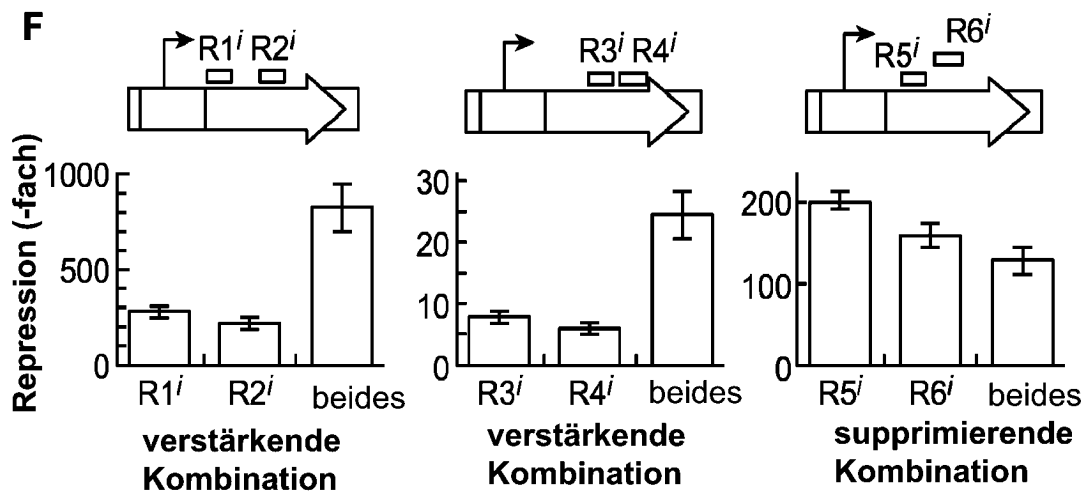
D**E****F**

FIG. 43 (Fortsetz.)

A

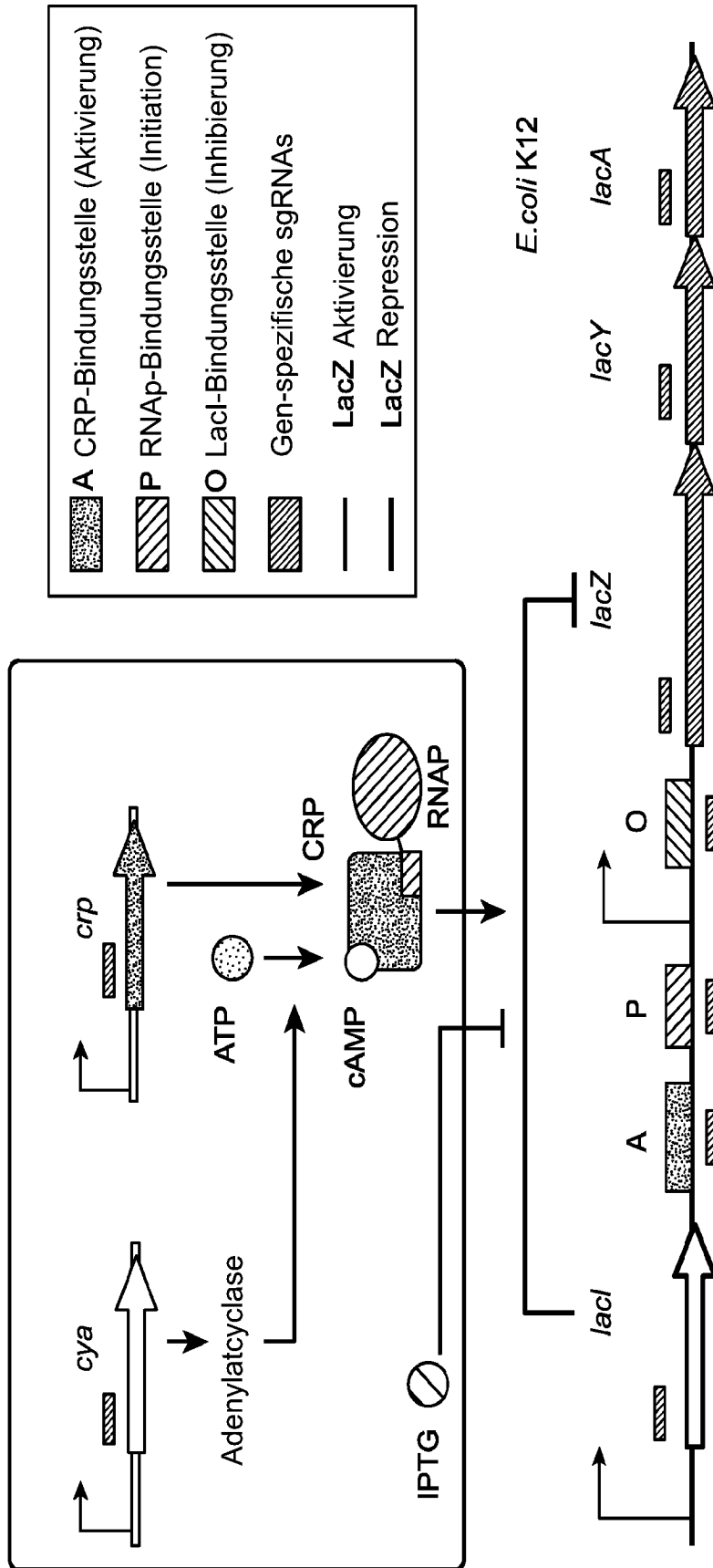


FIG. 44

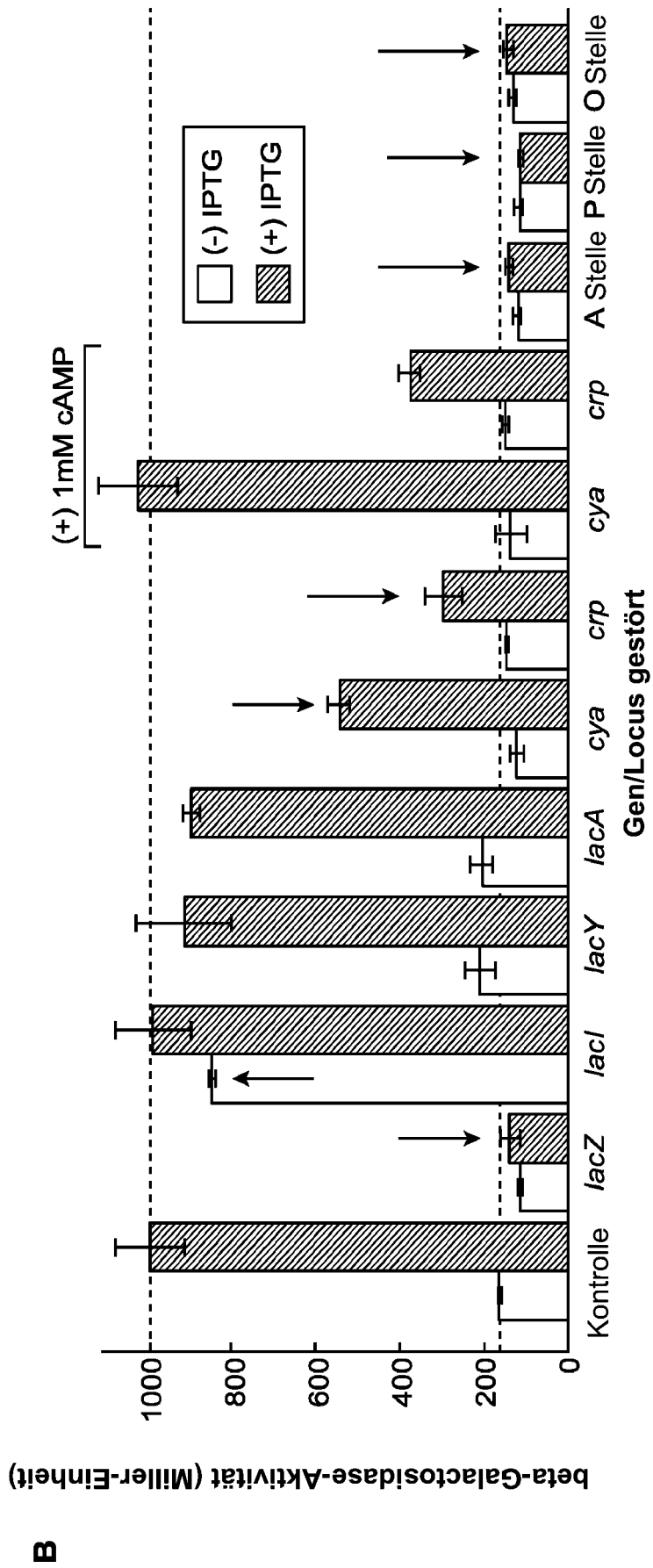


FIG. 44 (Fortsetz. 1)

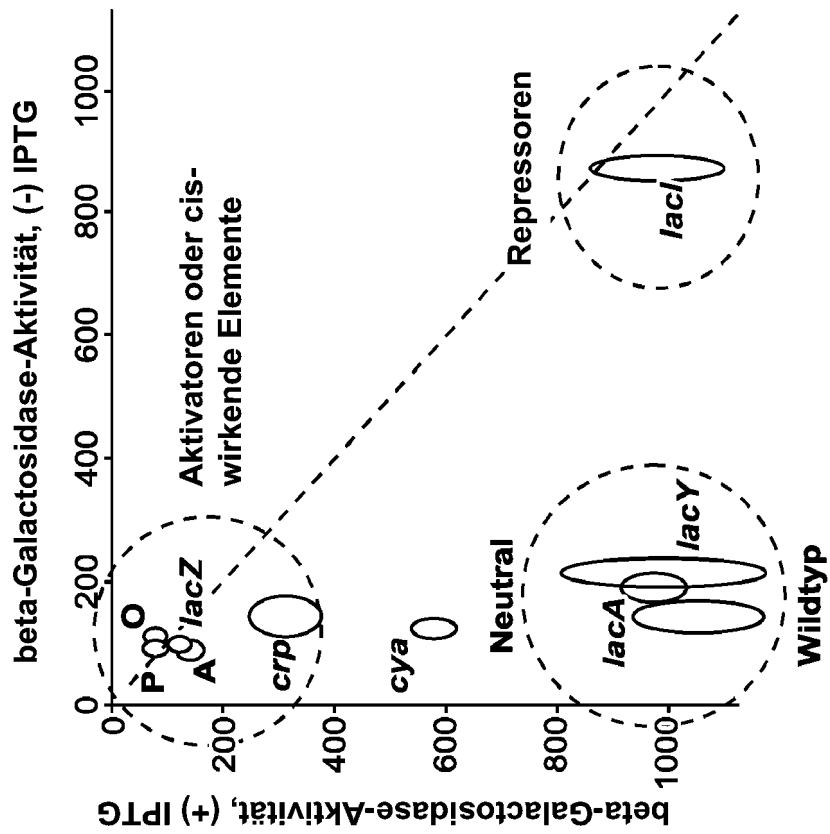
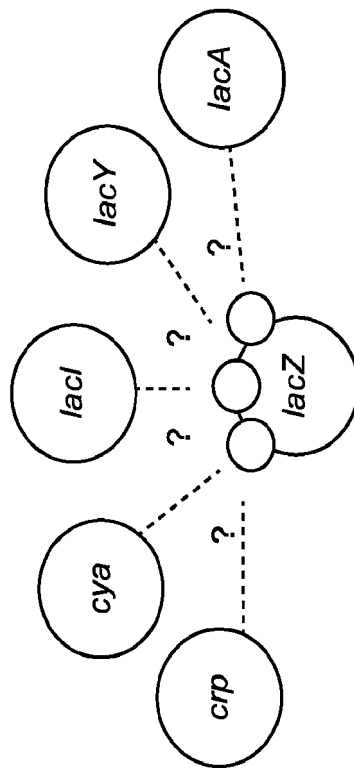


FIG. 44 (Fortsetz. 2)

C



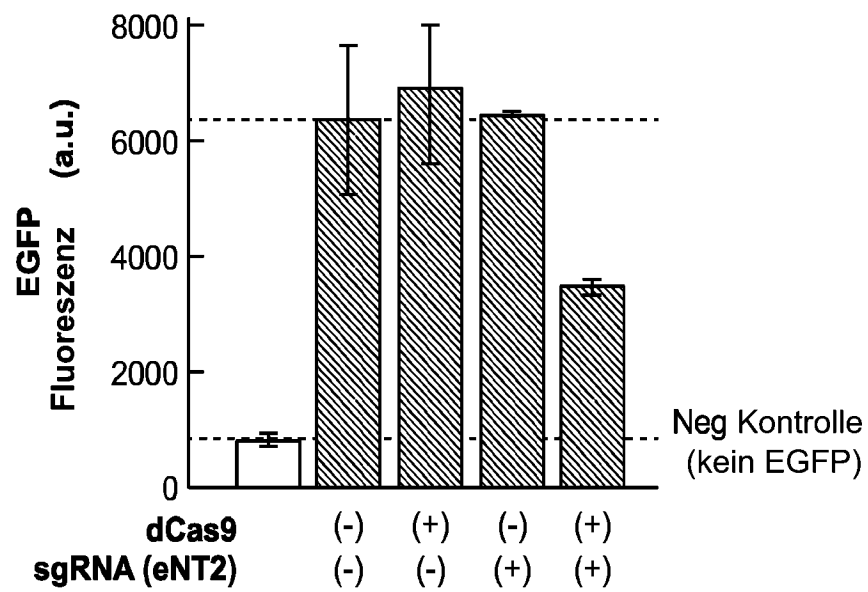
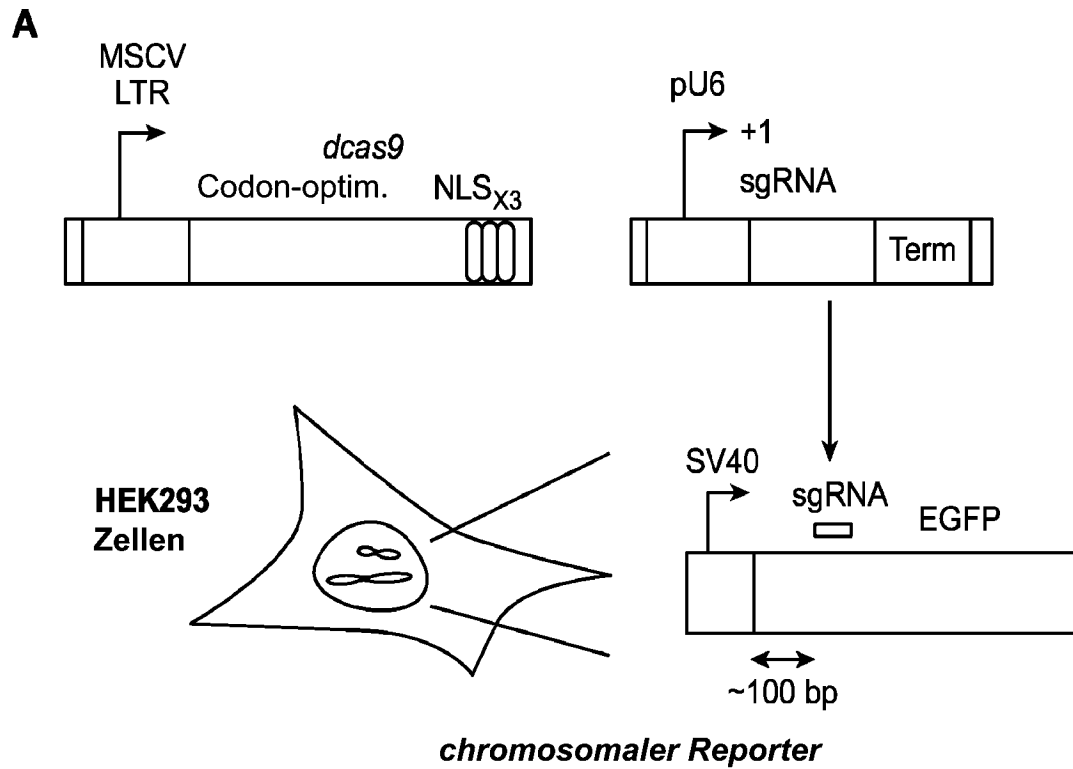


FIG. 45

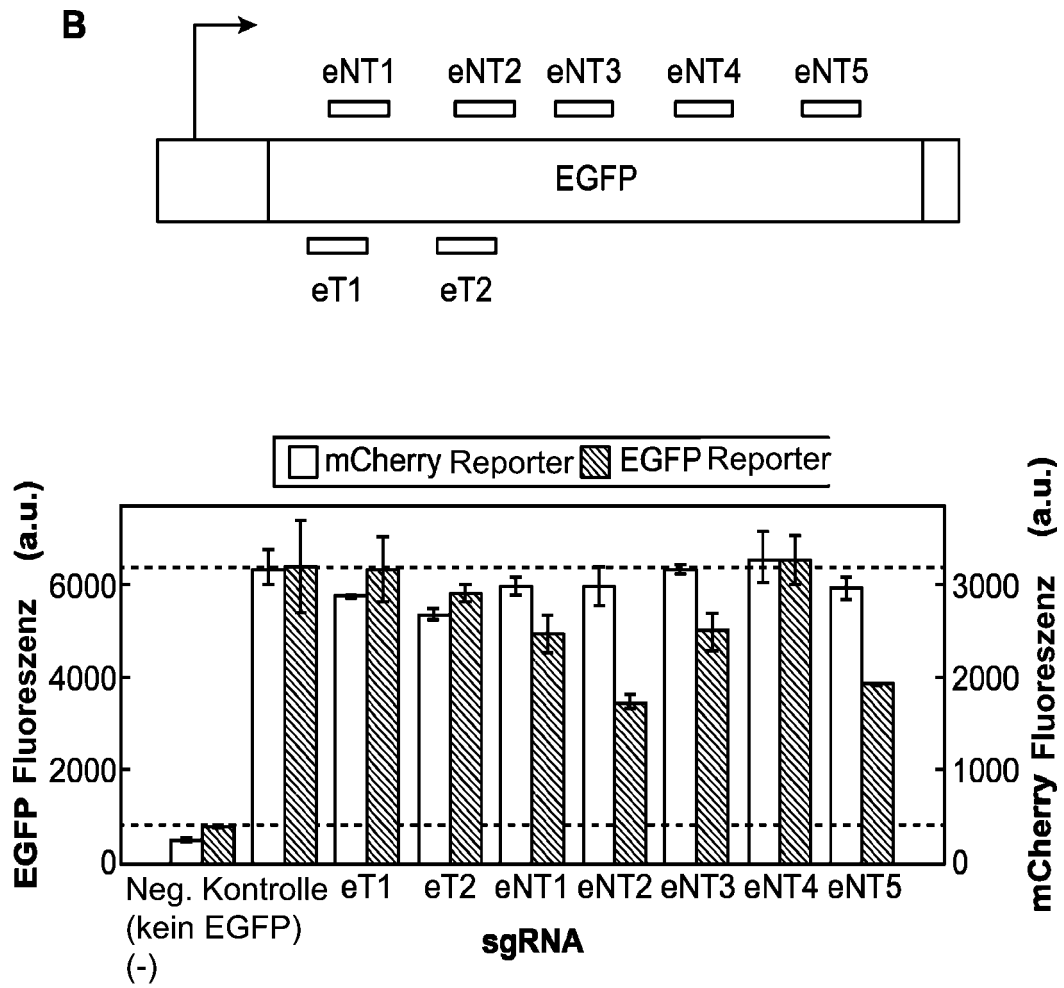


FIG. 45 (Fortsetz.)

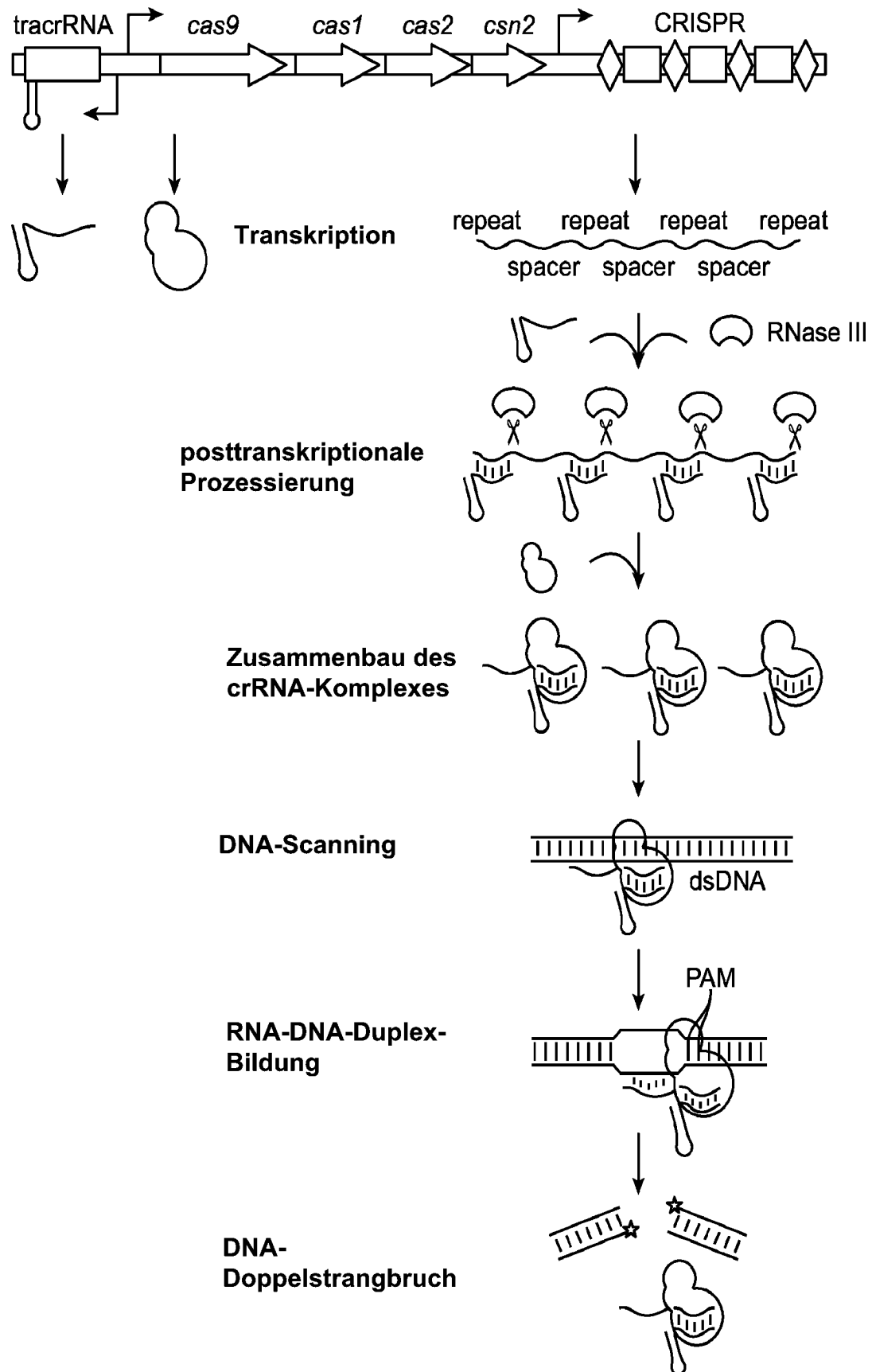
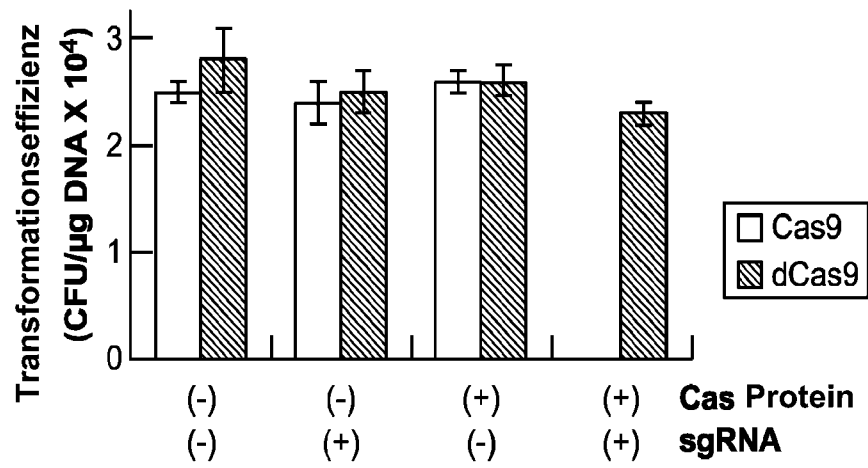
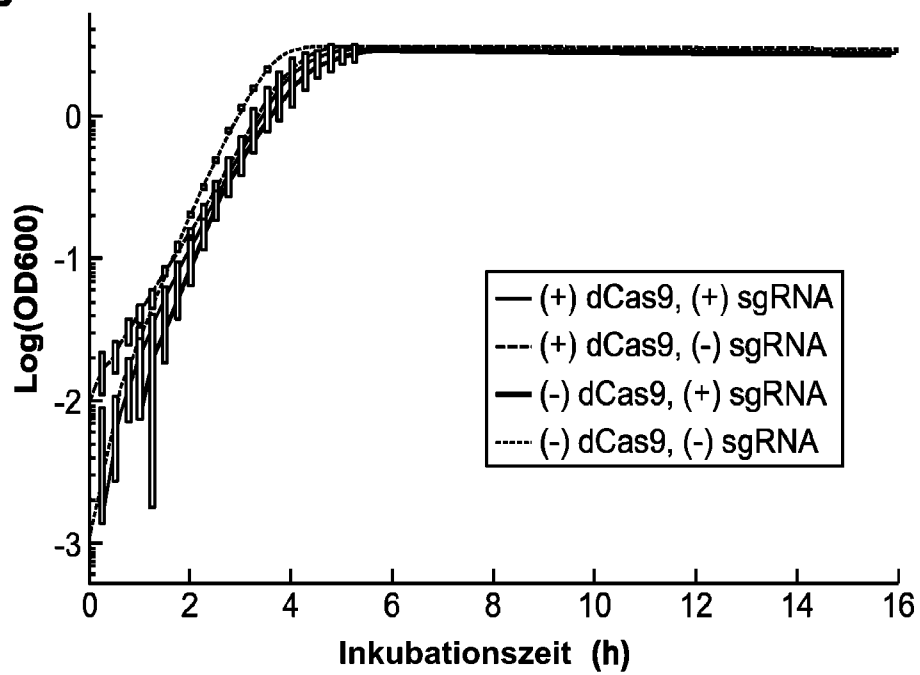


FIG. 46

A**B****FIG. 47**

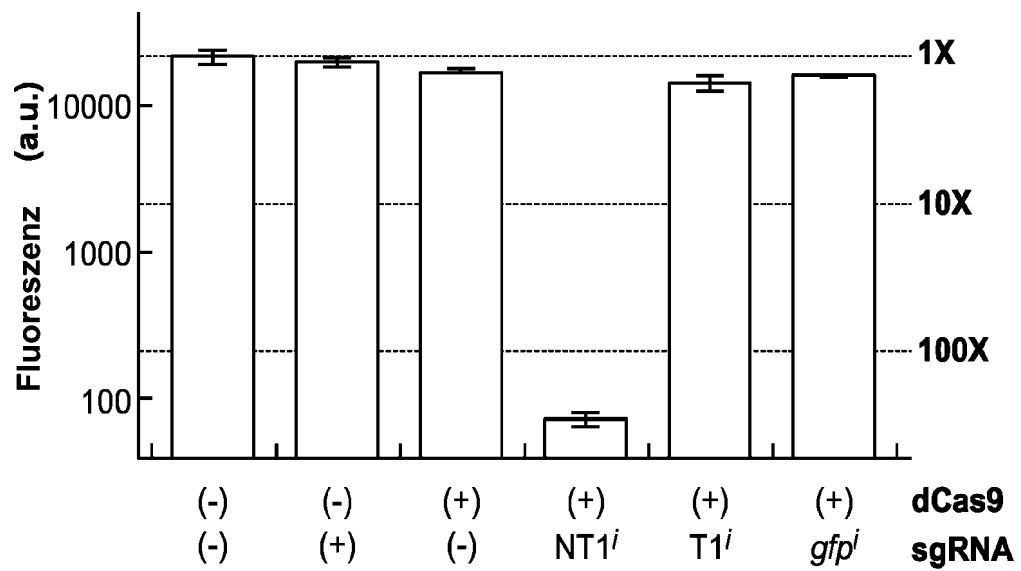


FIG. 48

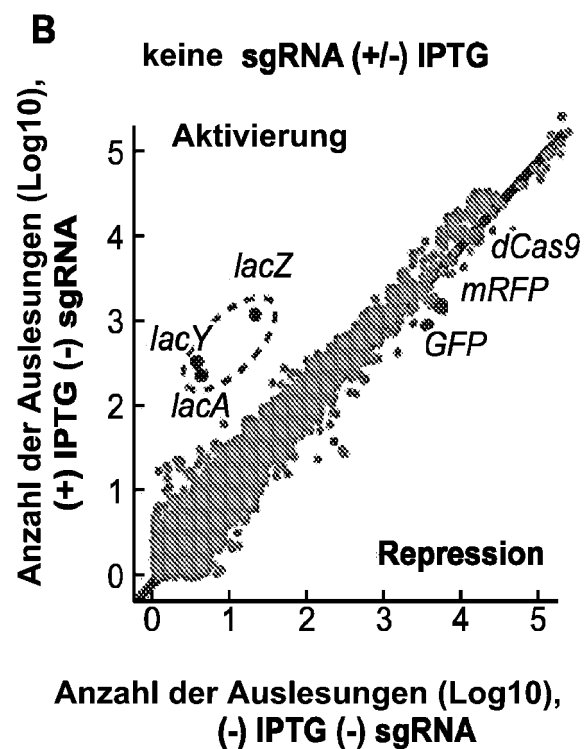
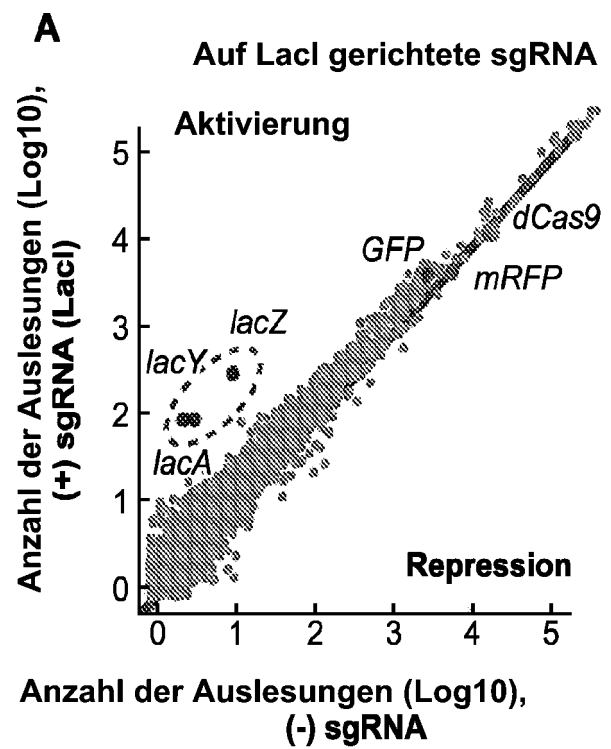


FIG. 49

C

Auf LacZ gerichtete sgRNA

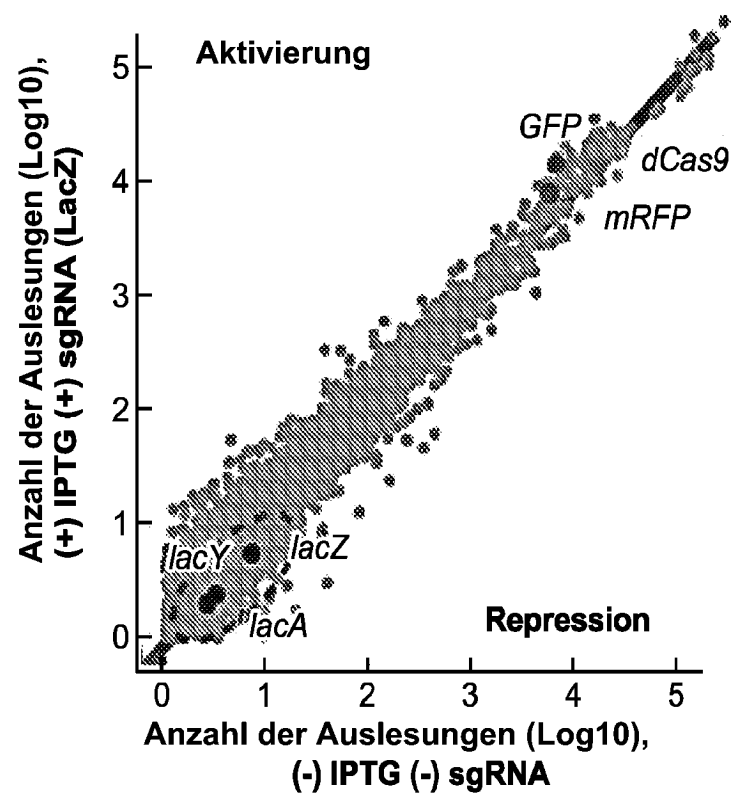


FIG. 49 (Fortsetz.)

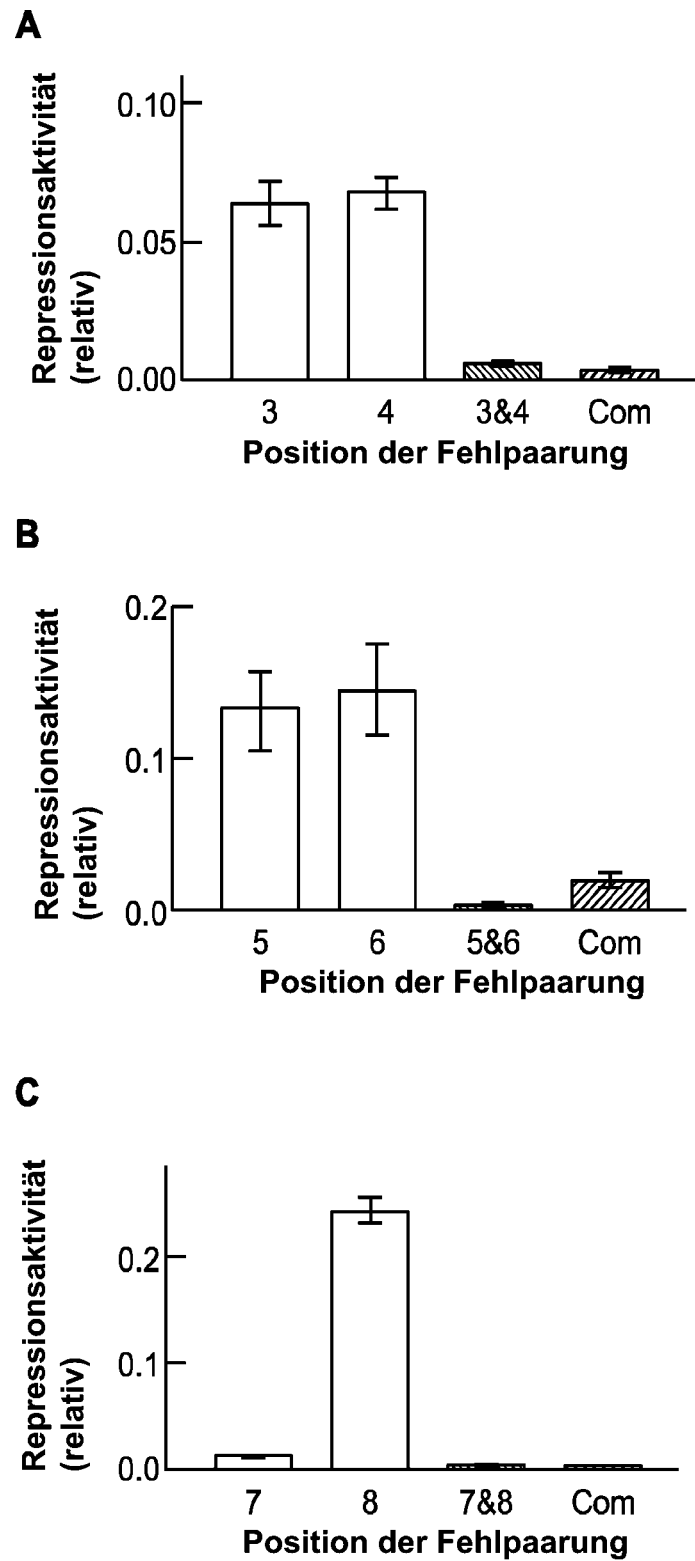
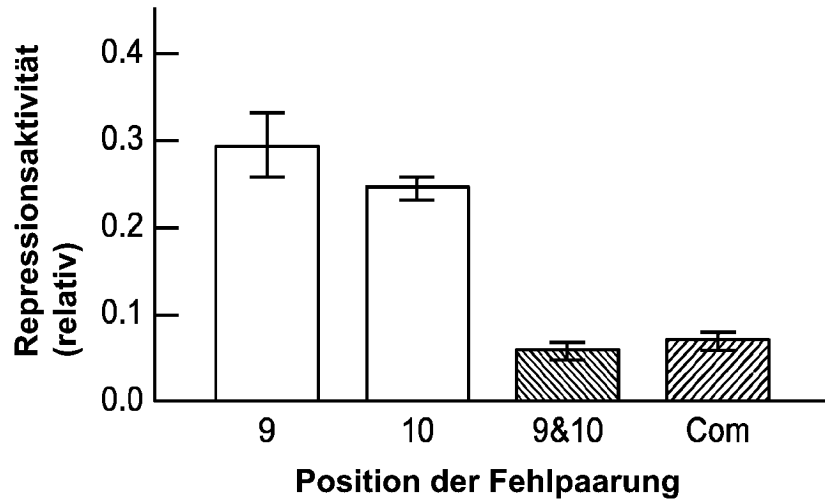


FIG. 50

D



E

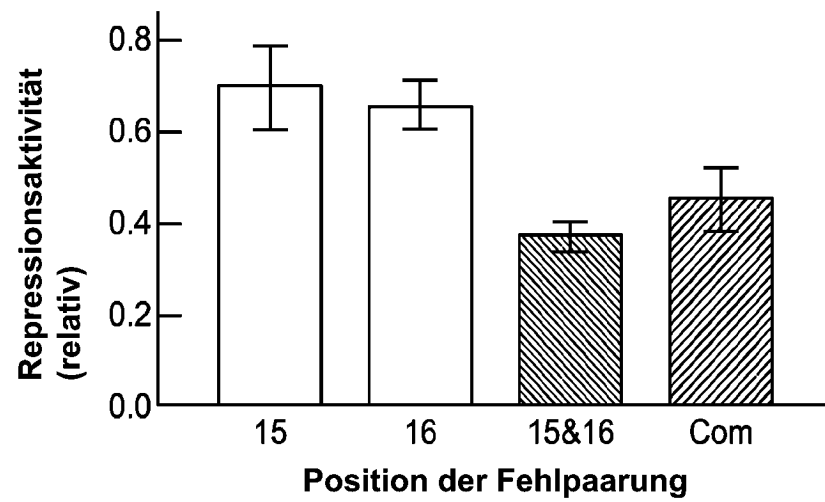
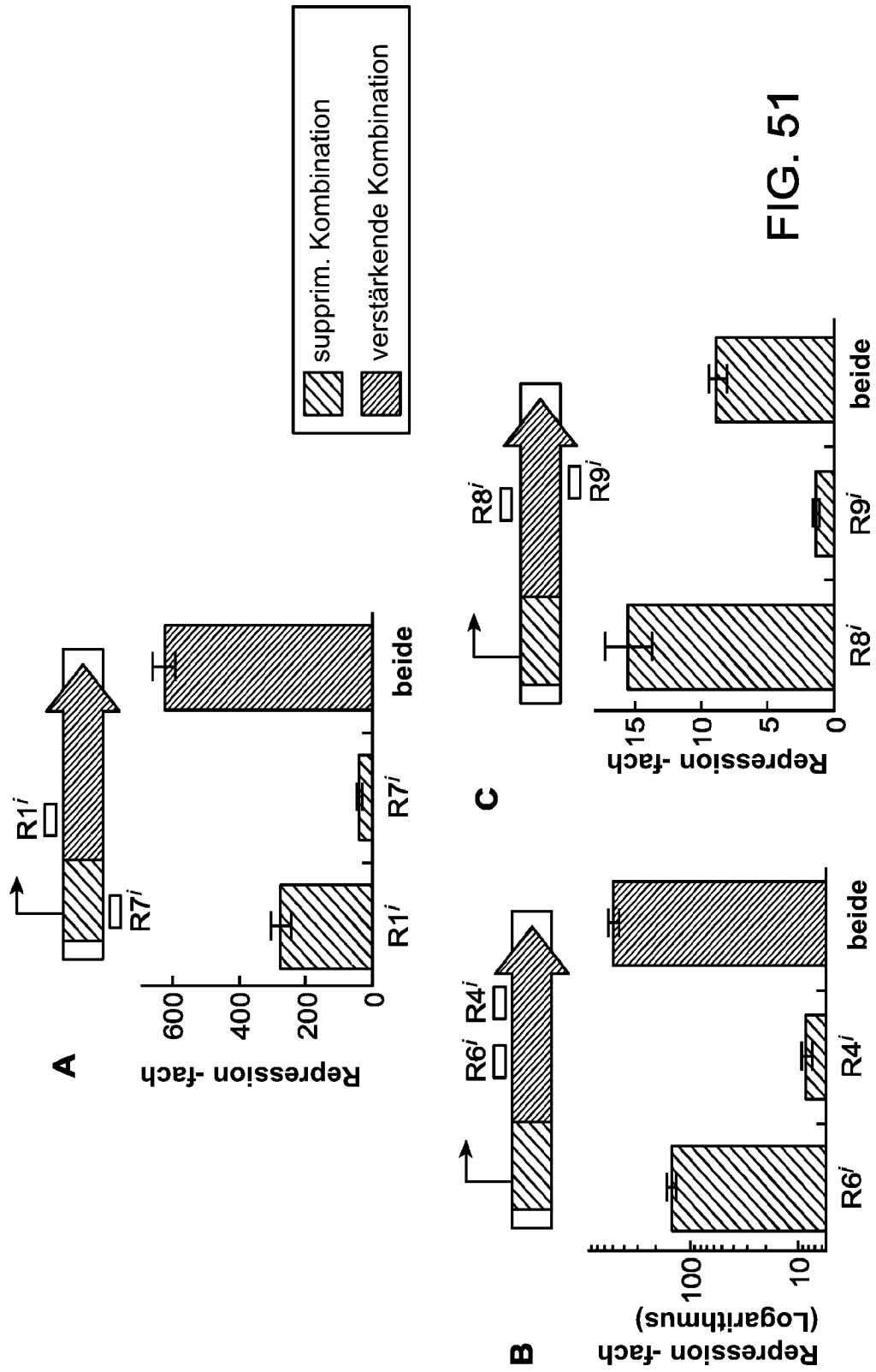


FIG. 50 (Fortsetz.)



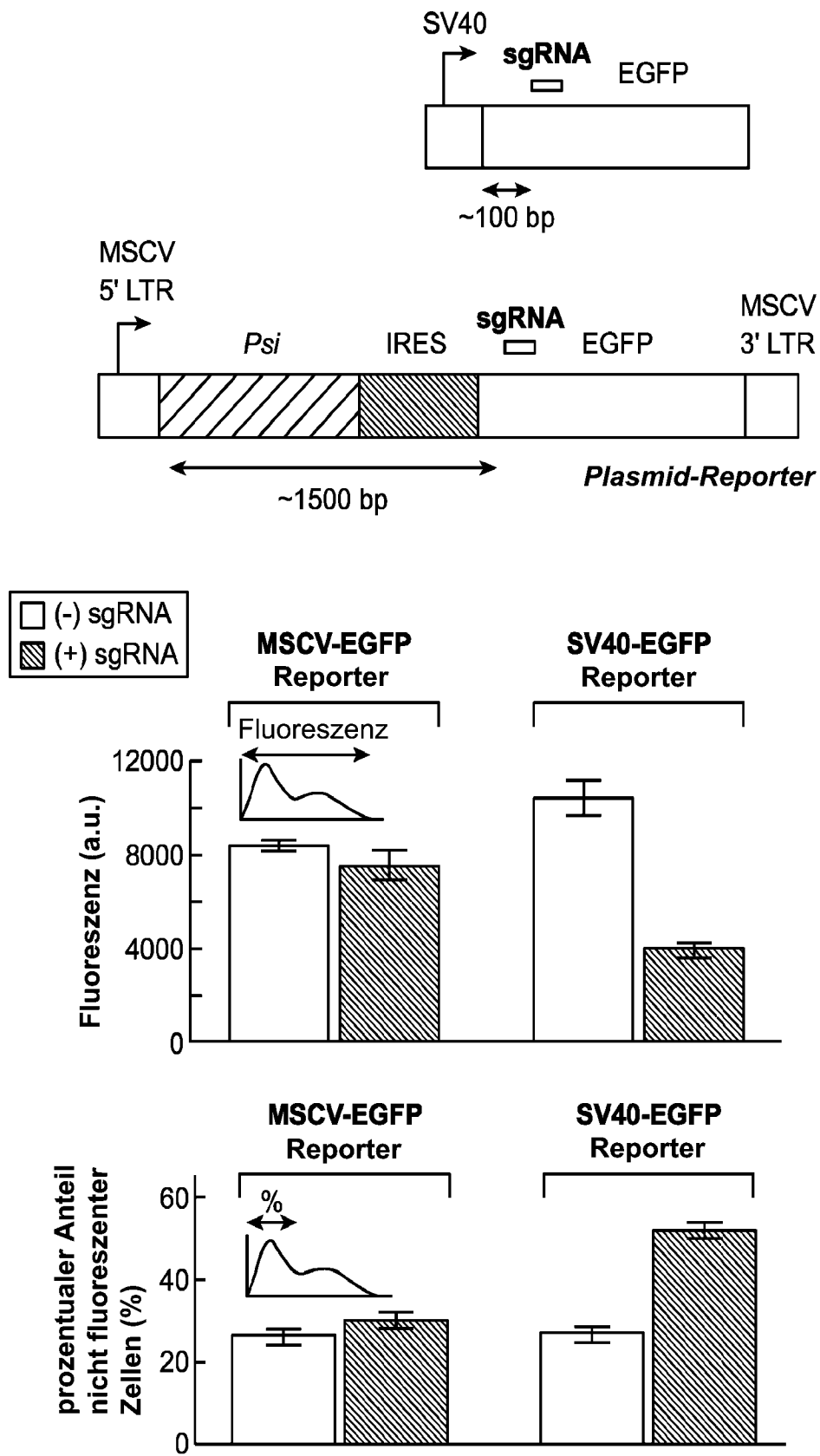


FIG. 52

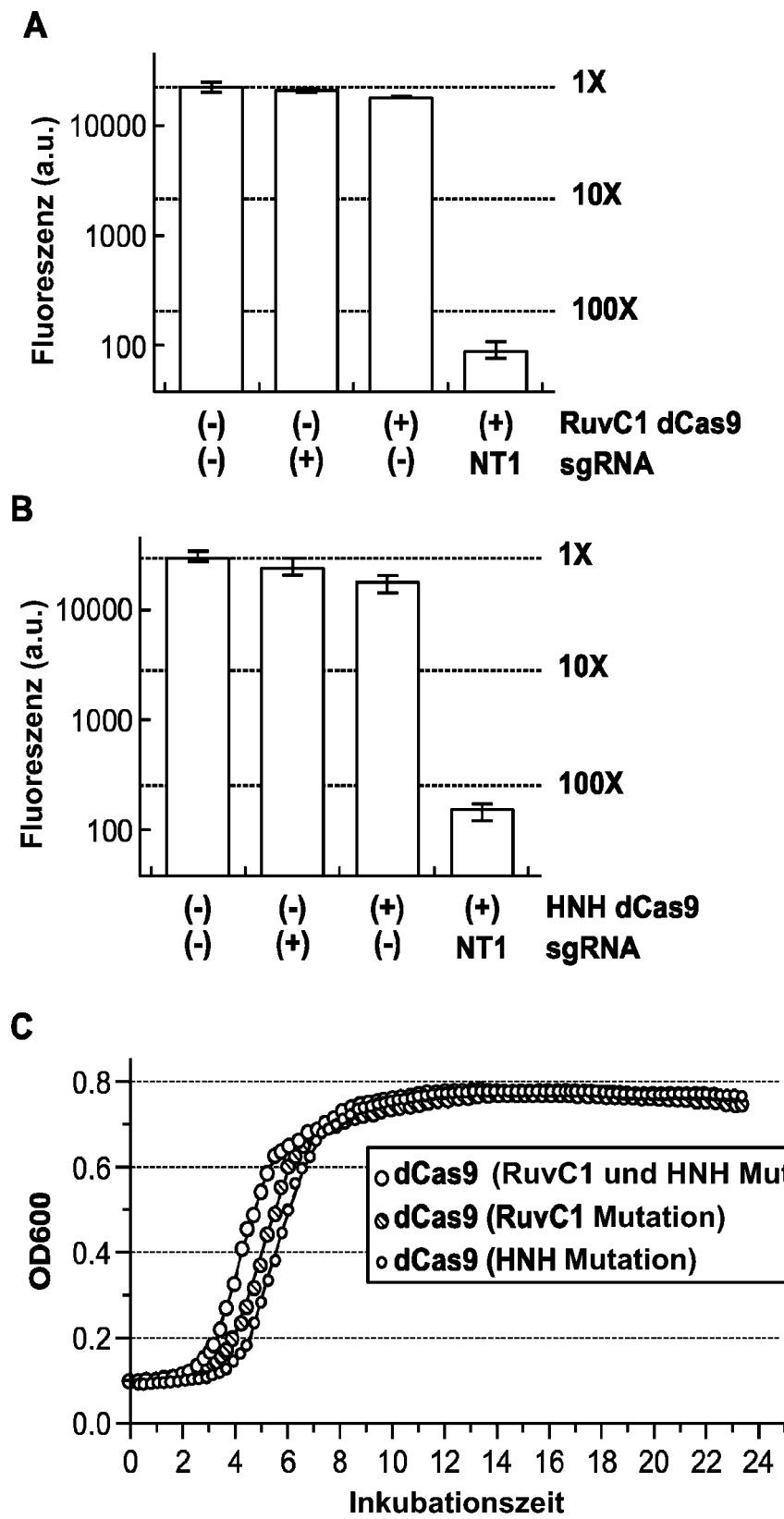


FIG. 53

A

Name des Proteins	Funktion
Transkriptionsaktivatoren	
GAL4	Transkriptionsaktivierung
VP16	Transkriptionsaktivierung
VP64	Transkriptionsaktivierung
p65 Subdomäne (NFkB)	Transkriptionsaktivierung
Transkriptionsrepressoren	
KRAB	Transkriptionsrepression
Mad mSIN3 Interaktionsdomäne (SID)	Transkriptionsrepression
ERF Repressordomäne ERD)	Transkriptionsrepression
Histon-Lysin-Methyltransferasen (KDM)	
KMT1 Familie SUV39H1, SUV39H2, G9A, ESET/SETDB1, und Homologe (Clr4, Su(var)3-9)	Heterochromatin-Bildung/ Transkriptionsrepression
KMT2 Familie hSET1A, hSET1B, MLL1 to 5, ASH1, und Homologe (Trx, Trr, Ash1)	Transkriptionsaktivierung
KMT3 Familie SYMD2, NSD1	Transkriptionsaktivierung
KMT4: DOT1L und Homologe	Transkriptionsaktivierung
KMT5 Familie Pr-SET7/8, SUV4-20H1, und Homologe (PR-set7, Suv4-20, Set9)	DNA-Schadensreparatur, Transkriptionsrepression
KMT6: EZH2	Polycomb-Silencing
KMT8: RIZ1	Transkriptionsrepression
Histon-Lysin-Demethylase (KDM)	
KDM1: LSD1/BHC110 und Homologe (SpLsd1/Swm1/Saf110, Su(var)3-3)	Transkriptionsaktivierung und -repression, Heterochromatin-Bildung
KDM3 Familie JHDM2a/b	Androgenrezeptor-Gen-Aktivierung; Spermatogenese
KDM4 Familie: JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, und Homologe (Rph1)	Transkriptionselongation, Transkriptionsrepression, Heterochromatin-Bildung Genom-Integrität
KDM5 Familie: JARID1A/RBP2, JARID1B/PLU-1, JARID1C/SMCX, JARID1D/SMCY, und Homologe (Lid, Jhn2, Jmj2)	Transkriptionsrepression
KDM6 Familie: UTX, JMJD3	Transkriptionsaktivierung

FIG. 54

B

Name des Proteins	Funktion
Histon-Lysin-Acetyltransferasen (KAT)	
KAT2 Familie hGCN5, PCAF, und Homologe (dGCN5/PCAF, Gcn5)	Transkriptionsaktivierung, DNA-Reparatur
KAT3 Familie CBP, p300, und Homologe (dCBP/NEJ)	Transkriptionsaktiv., DNA-Reparatur
KAT4: TAF1 und Homologe (dTAF1)	Transkriptionsaktivierung
KAT5: TIP60/PLIP, und Homologe	Transkriptionsaktiv., DNA-Reparatur
KAT6: MOZ/MYST3, MORF/MYST4, und Homologe (Mst2, Sas3, CG1894)	Transkriptionsaktivierung und -elongation, DNA-Replikation
KAT7: HBO1/MYST2, und Homologe (CHM, Mst2)	Transkription, DNA-Replikation
KAT8: HMOF/MYST1, und Homologe (dMOF, CG1894, Sas2, Mst2)	Chromatin-Begrenzung, Dosierungsausgleich, DNA-Reparatur
KAT13 Familie SRC1, ACTR, P160, CLOCK, und Homologe	Transkriptionsaktivierung
Histon-Lysin-Deacetylasen	
Class I: HDAC1, HDAC2, HDAC3, HDAC8, und deren Homologe (Rpd3, Hos1, Cir6)	Transkriptionsrepression, Bildung von Heterochromatin
Class IIa: HDAC4, HDAC5, HDAC7, HDAC9, und deren Homologe (Hda1, Cir3, usw.)	Transkriptionsrepression, Bildung von Heterochromatin
Class III: SIRT1, SIRT2, und deren Homologe (Sir2, Hst1, Hst2, Hst3, Hst4)	Transkriptionsrepression, Bildung von Heterochromatin
Class IV: HDAC11	Transkriptionsrepression
DNA-Methylasen (Adenosin- oder Cytosin-Modifikation)	
Dam (E. coli)	Restriktionssystem
Dcm (E. coli)	Restriktionssystem
M. SssI (Spiroplasma sp)	Restriktionssystem
DNMT1	Transkriptionsrepression, Imprinting, Bildung von Heterochromatin
DNMT3a/DNMT3b, MET1, DRM3 (Pflanzen) und Homologe	Transkriptionsrepression, Imprinting, Bildung von Heterochromatin
Chromomethylases z.B. ZMET2, CMT1, CMT2 (Pflanz.)	Transkriptionsrepression, Imprinting, Bildung von Heterochromatin
DNA- Demethylasen	
AID/Apobec Deaminase-Familie A/D	Transkriptionsaktiv., Genom-Integrität
TET Dioxygenase-Familie TET1	Transkriptionsaktiv., Genom-Integrität
DEMETER Glycosylase-Familie : DME, DML1, DML2, ROS1	Transkriptionsaktiv., Genom-Integrität

FIG. 54 (Fortsetz. 1)

C

Name des Proteins	Funktion
Begrenzungselemente	
CTCF	Chromatin-Isolierung, Suppression des Heterochromatin-Spreadings
Elemente der periph. Rekrutierung	
Lamin A	Transkriptionsrepression
Lamin B	Transkriptionsrepression
Protein-Docking-Elemente	
FKBP/FRB (<i>S. pombe</i>)	Rapamycin-abh. Recruitment
Pil1/Aby1 (<i>E. coli</i>)	ABA-abh. Recruitment

FIG. 54 (Fortsetz. 2)

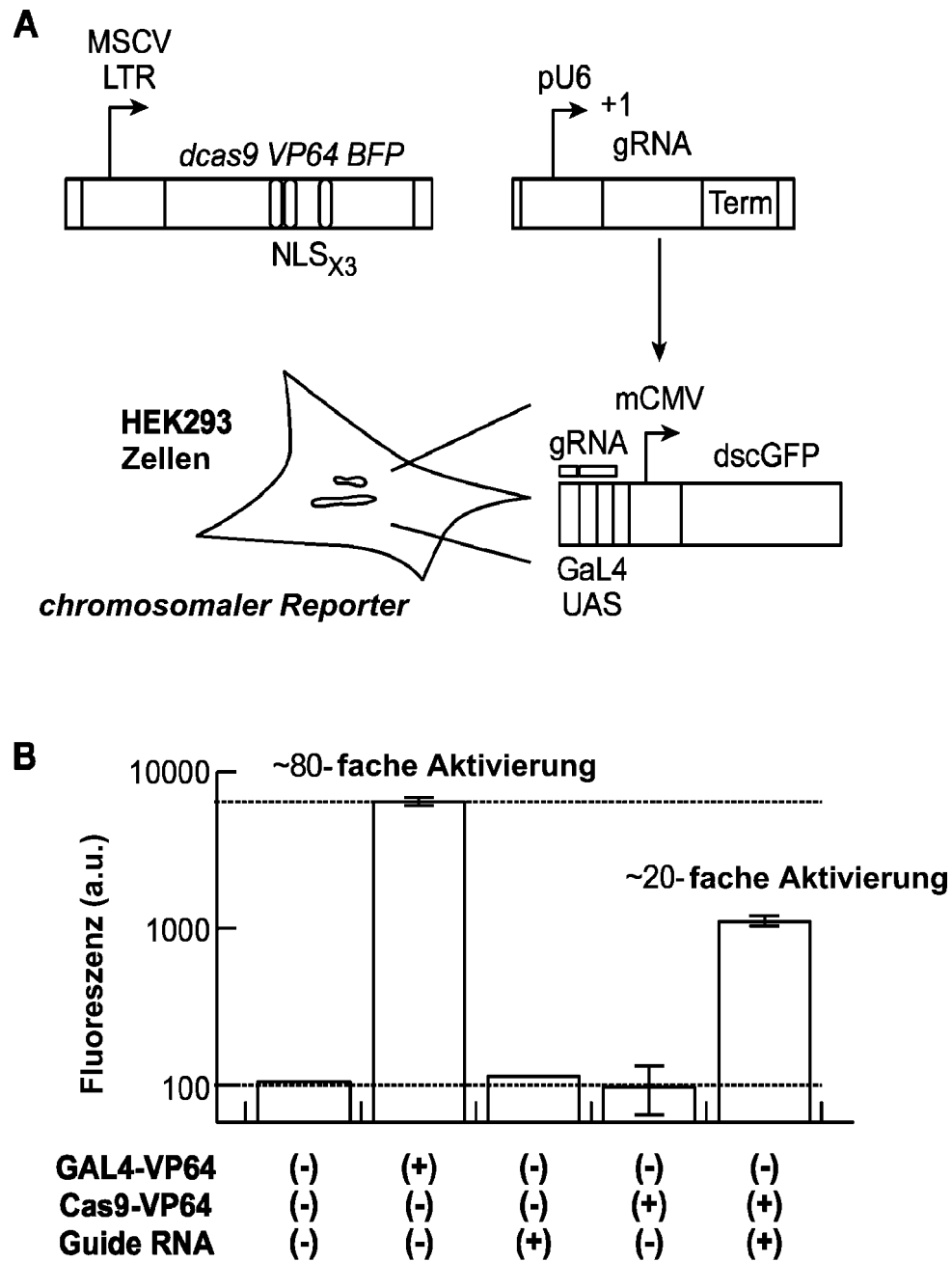


FIG. 55

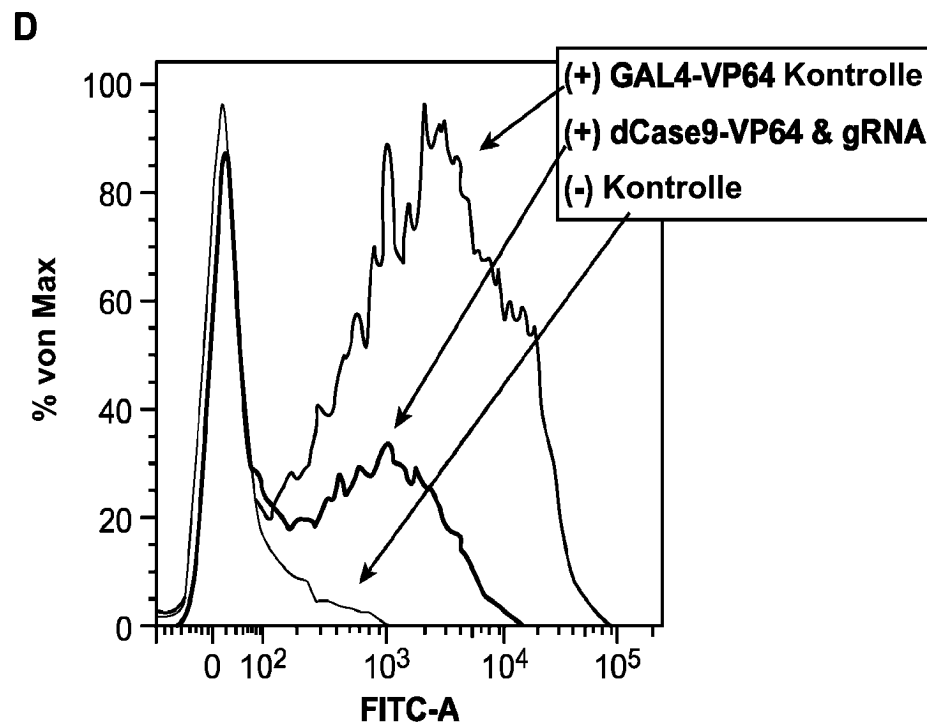
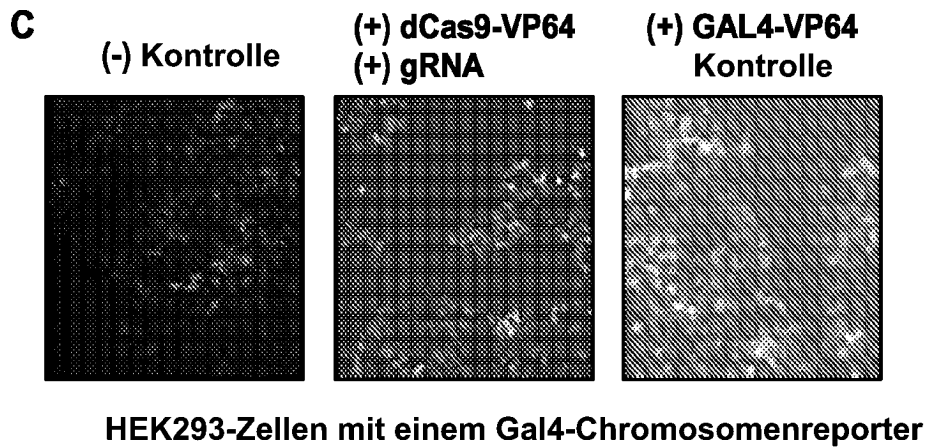


FIG. 55 (Fortsetz.)

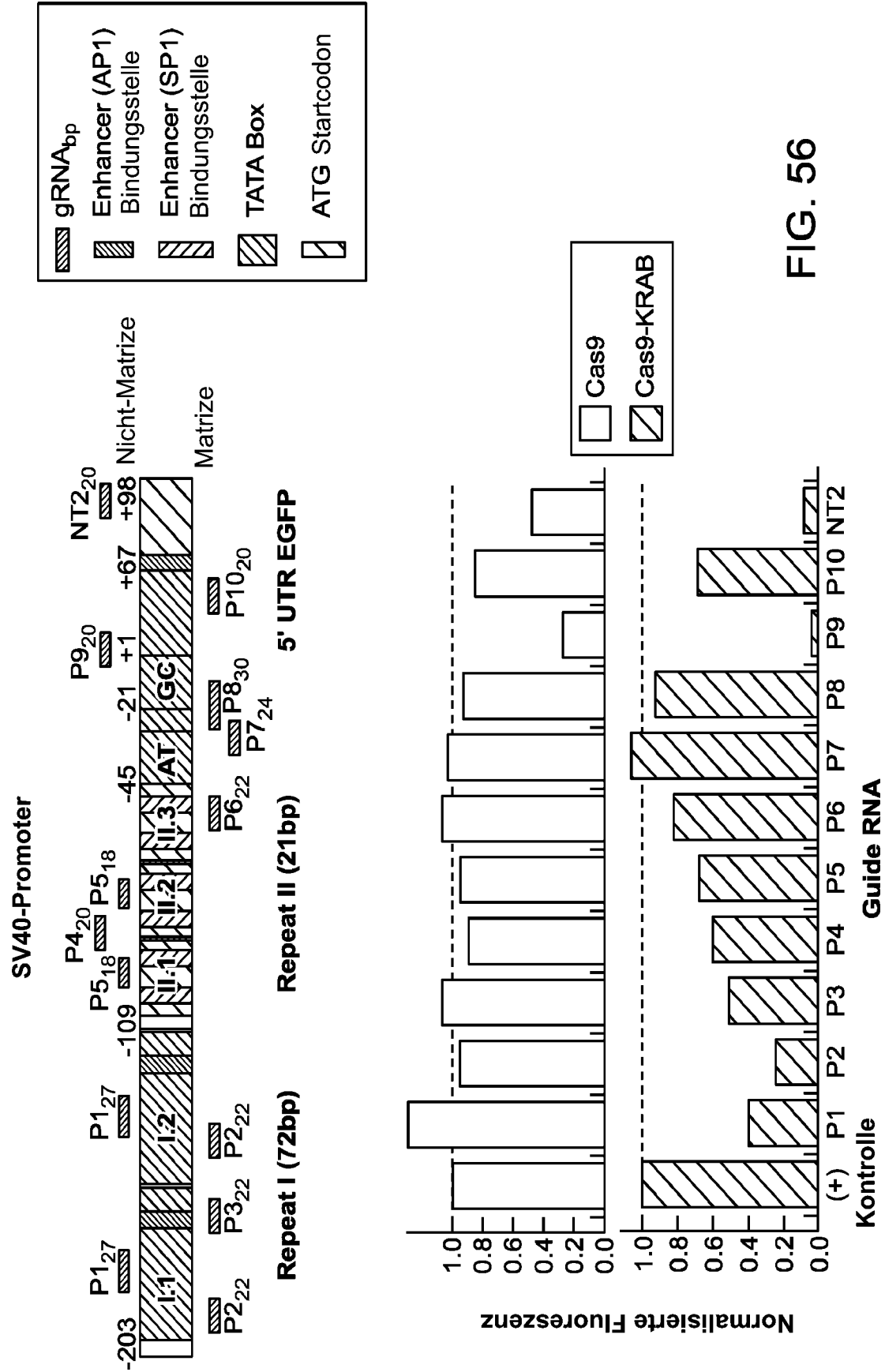


FIG. 56

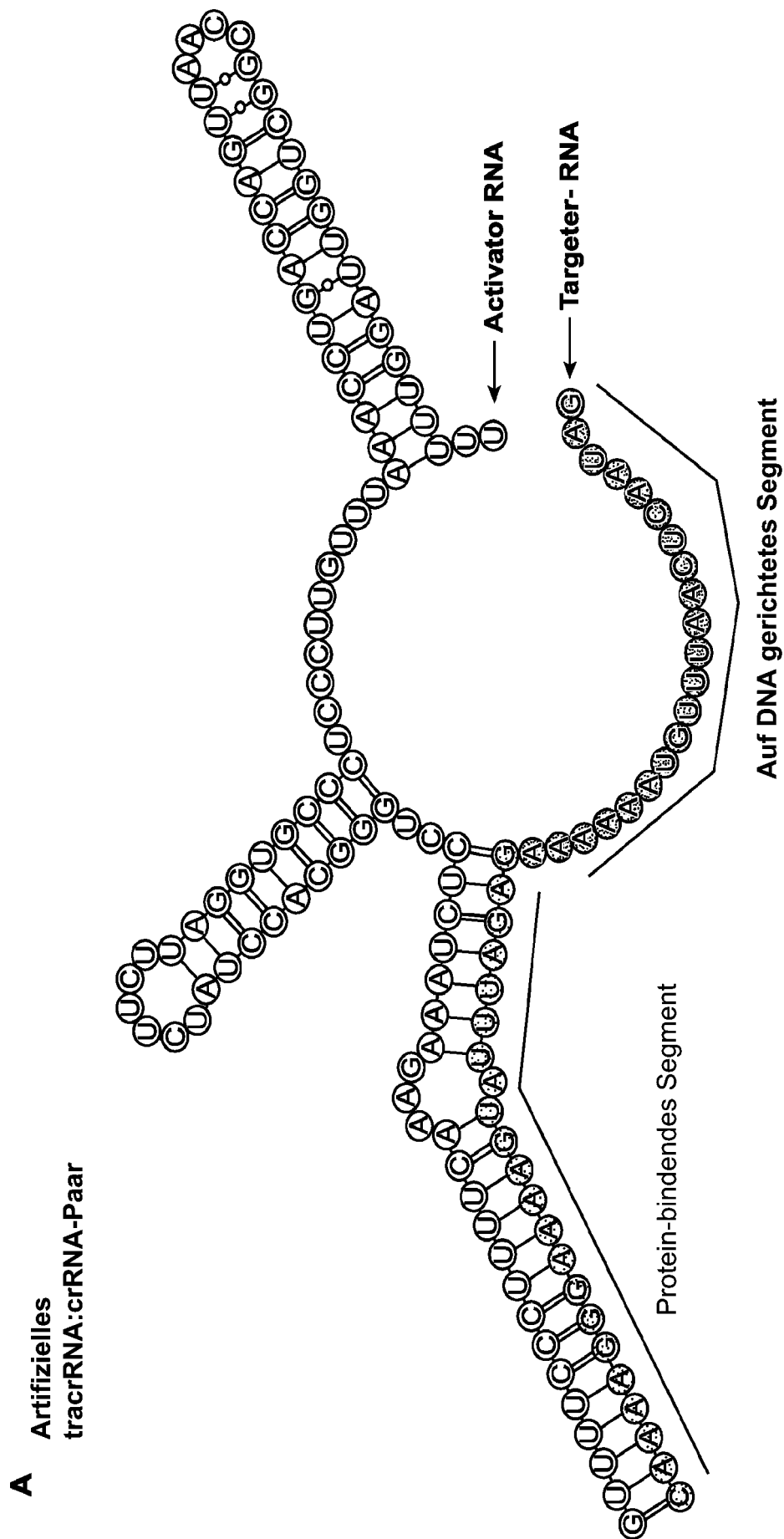


FIG. 57

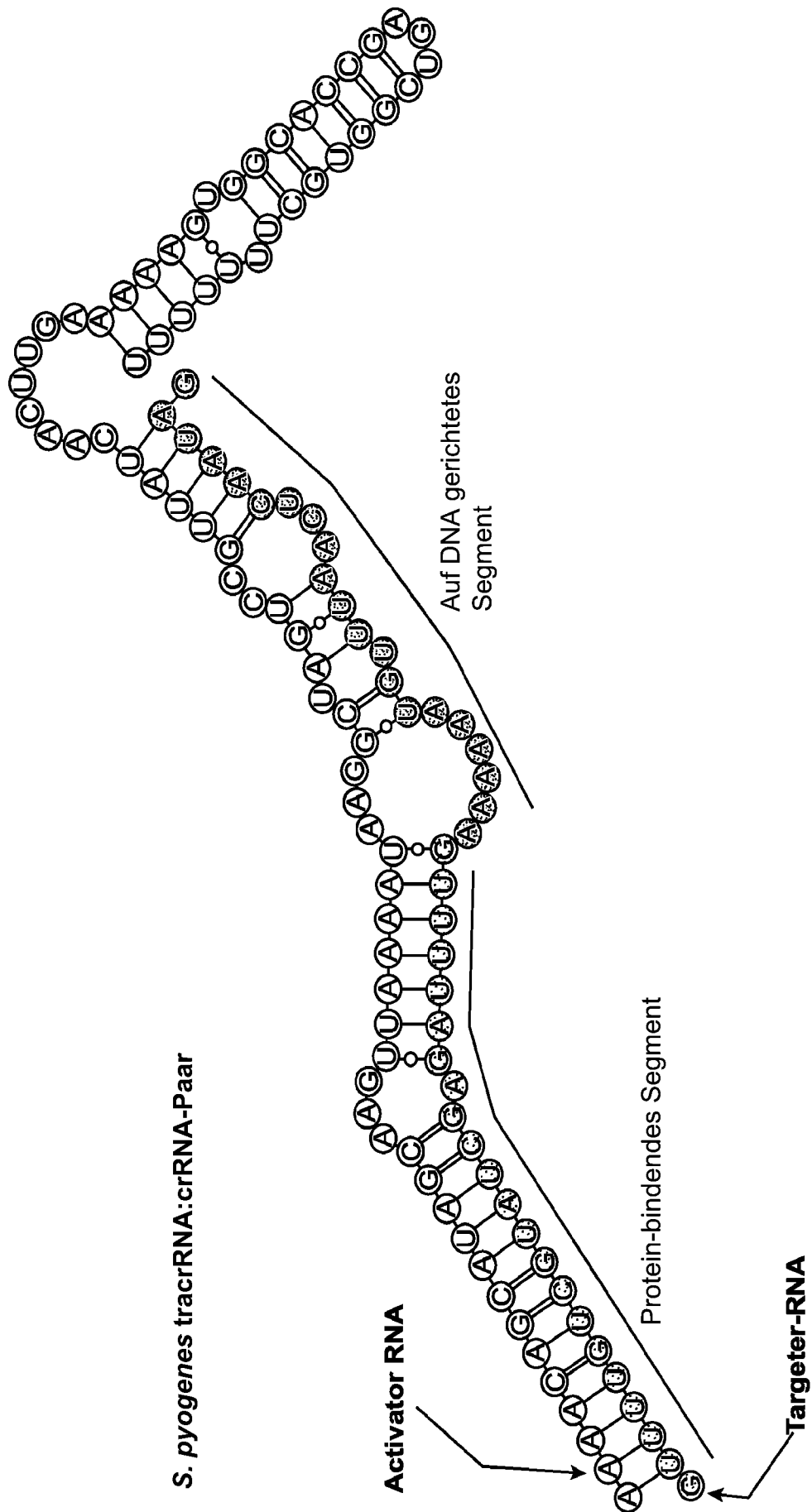


FIG. 57 (Fortsetz. 1)

FIGUR 57 B

