

(12) Ausschließungspatent

(11) DD 284 053 A5



Ertelt gemäß § 17 Absatz 1
Patentgesetz der DDR
vom 27. 10. 1983
in Übereinstimmung mit den entsprechenden
Festlegungen im Einigungsvertrag

5(51) C 12 Q 1/68
C 07 H 21/04
C 12 Q 1/34

DEUTSCHES PATENTAMT

In der vom Anmelder eingereichten Fassung veröffentlicht

(21)	DD C 12 Q / 331 212 0	(22)	27.07.89	(44)	31.10.90
(31)	8818020.3	(32)	28.07.88	(33)	GB

(71) siehe (73)
(72) Markham, Alexander F.; Smith, John. C.; Anwar, Rashida, GB
(73) Imperial Chemical Industries PLC, London SW1P 3JF, GB
(74) Internationales Patentbüro Berlin, Wallstraße 23/24, Berlin, 1020, DD

(54) Ein Verfahren zur Verstärkung von nucleotiden Sequenzen

(55) Amplifikation Nucleinsäurefragment; Sequenz unbekannt; Primerextension; Sequenzierung Nucleotidsequenzen, schnell; Bildung Targetnucleinsäurefragment/Vektoretteneinheiten; Spaltung Targetnucleinsäure; Ligation

(57) Die Erfindung betrifft eine Methode für die Amplifikation eines Nucleinsäurefragments mit unbekannter Sequenz durch Primerextension, und Kits dafür, die die schnelle und wirksame Sequenzierung langer Nucleotidsequenzen ermöglichen. Die Methode umfaßt die Bildung von Targetnucleinsäurefragment/Vektoretteneinheiten durch Spaltung von Targetnucleinsäure und anschließende Ligation. Eines der Nucleinsäurefragments enthält eine Initiierungsstartregion mit bekannter Sequenz für die Hybridisierung mit einem Initiierungsprimer, und Targetnucleinsäurefragment/Vektoretteneinheiten enthalten eine Vektorettestartregion mit bekannter Sequenz für die Hybridisierung mit einem Vektoretteprimer. Die Amplifikation wird durch Primerextension eines Initiierungsprimers, der an die Initiierungsstartregion der Targetnucleinsäurefragment/Vektoretteneinheit hybridisiert wird, bewirkt.

Patentansprüche:

1. Methode für die Amplifikation eines Nucleinsäurefragments, das eine unbekannte Sequenz aufweist, durch Primerextension, dadurch gekennzeichnet, daß sie umfaßt:

das Spalten einer Targetnucleinsäure zur Gewinnung von Targetnucleinsäurefragmenten, wobei eines der Fragmente eine Initiierungsstartregion mit bekannter Sequenz für die Hybridisierung mit einem Initiierungsprimer enthält; das Herstellen von Targetnucleinsäurefragment/Vektoretteneinheiten aus den Targetnucleinsäurefragmenten durch Ligation jeder Einheit mit einer Vektorettestartregion bekannter Sequenz für die Hybridisierung mit einem Vektorettprimer; und Behandeln der Targetnucleinsäurefragment/Vektoretteneinheiten zusammen und nacheinander mit geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation von Nucleosidtriphosphaten unter Hybridisierungsbedingungen, so daß ein Extensionsprodukt eines Initiierungsprimers komplementär zu einer einzelsträngigen Targetnucleinsäure/Vektoretteneinheit synthetisiert wird, die eine Initiierungsstartregion aufweist, an die ein Initiierungsprimer hybridisiert wird, der so ausgewählt wird, daß er im wesentlichen komplementär zu der Initiierungsstartregion ist, während kein derartiges Extensionsprodukt synthetisiert wird, das komplementär zu einzelsträngigen Targetnucleinsäurefragment/Vektoretteneinheiten ist, die keine solche Initiierungsstartregion haben.

2. Methode nach Anspruch 1, dadurch gekennzeichnet, daß das Extensionsprodukt der Amplifikation in Anwesenheit eines Vektorettprimers unterzogen wird, der so ausgewählt wird, daß er im wesentlichen komplementär zu einer Vektorettestartregion ist.

3. Methode nach Anspruch 2, dadurch gekennzeichnet, daß die Synthese von Vektorettprimeramplifikationsprodukten von der anfänglichen Synthese eines Extensionsproduktes des Initiierungsprimers abhängig ist.

4. Methode nach Anspruch 3, dadurch gekennzeichnet, daß der Vektoretteabschnitt der Targetnucleinsäurefragment/Vektoretteeinheit einen doppelsträngigen Abschnitt mit einem ersten und zweiten Strang umfaßt, wobei der zweite Strang des Vektoretteabschnitts an jenen Strang des Targetnucleinsäurefragments ligiert ist, der die Initiierungsstartregion enthält, und die Nucleotidsequenz des ersten Stranges, zweiten Stranges und Vektoretteprimers so ausgewählt werden, daß der Vektoretteprimer unter den gleichen Hybridisierungsbedingungen zur Hybridisierung an das Komplement des zweiten Stranges, aber nicht an den ersten Strang in der Lage ist.

5. Methode nach Anspruch 3, dadurch gekennzeichnet, daß der Vektoretteabschnitt der Targetnucleinsäurefragment/Vektoretteeinheit einen doppelsträngigen Abschnitt mit einem ersten und zweiten Strang umfaßt, wobei der erste Strang eine terminale Polymerisationsblockierkomponente aufweist und der zweite Strang, der im Einsatz an jenen Strang des Targetnucleinsäurefragments ligiert wird, der die Initiierungsstartregion enthält, einen einzelsträngigen Abschnitt trägt, wobei die terminale Polymerisationsblockierkomponente dahingehend wirkt, die Extension des ersten Stranges zur Bildung eines Komplements zu dem einzelsträngigen Abschnitt des zweiten Stranges in Anwesenheit geeigneter Nucleosidtriphosphate und eines Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen zu verhindern.

6. Methode nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß eine Vielzahl von unterschiedlichen Vektorettebibliotheken zum Einsatz mit dem gleichen einzelnen Initiierungsprimer hergestellt wird, wobei jede Vektorettebibliothek Targetnucleinsäurefragment/Vektoretteeinheiten umfaßt, die durch Ligation von Nucleinsäurefragmenten gewonnen werden, die hergestellt werden durch: Spalten von Targetnucleinsäure an unterschiedlichen Spaltstellen; Behandeln jeder Vektorettebibliothek, entweder separat oder zusammen, mit geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation von Nucleosidtriphosphaten un-

ter Hybridisierungsbedingungen, um dadurch eine Vielzahl von Initiierungsprimerextensionsprodukten auf der Basis des Einsatzes eines einzelnen Initiierungsprimers zu gewinnen.

7. Methode nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß eine Nucleinsäure, die zur Ligation aneinander fähige Termini aufweist, zirkularisiert wird, wobei die Nucleinsäure einen Abschnitt aufweist, der in der Lage ist, als eine Initiierungsstartregion für die Hybridisierung mit einem Initiierungsprimer zu dienen; die zirkularisierte Nucleinsäure außerhalb der bekannten Nucleotidsequenz gespalten wird, um ein lineares Molekül zu bilden, das die bekannte Nucleotidsequenz enthält und an mindestens einem Terminus ein bekanntes Spaltstellenmuster für die Ligation zur Bildung einer Targetnucleinsäurefragment/Vektoretteneinheit aufweist; die Targetnucleinsäurefragment/Vektoretteneinheit durch Ligation gebildet wird und die Targetnucleinsäurefragment/Vektoretteneinheit zusammen oder nacheinander mit Initiierungsprimer, geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen behandelt wird.

8. Methode nach Anspruch 7, dadurch gekennzeichnet, daß die Targetnucleinsäurefragment/Vektoretteneinheit zusammen oder nacheinander mit Initiierungsprimer, Vektoretteneinheit, geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate behandelt wird.

9. Methode nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß die Targetnucleinsäurefragment/Vektoretteneinheiten aus Targetnucleinsäurefragmenten durch Ligation hergestellt werden, so daß die Vektorette von den daraus gewonnenen Targetnucleinsäurefragment/Vektoretteneinheiten nicht durch das gleiche Agens abgespalten werden kann, das zur Spaltung der Targetnucleinsäure zur Gewinnung der Targetnucleinsäurefragmente verwendet worden ist.

10. Methode nach Anspruch 9, dadurch gekennzeichnet, daß die Targetnucleinsäure mit einer Restriktionsendonuclease gespalten wird,

um ein Targetnucleinsäurefragment für die Ligation an eine Vektorette zu gewinnen, wobei die Sequenz der Vektorette so ausgewählt wird, daß die Restriktionsendonucleaseerkennungssequenz der Restriktionsendonuclease in der Targetnucleinsäurefragment/Vektoretteeinheit abwesend ist.

11. Methode nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß eines oder alle der gewonnenen Initiierungsprimerextensionsprodukte mindestens an dem (den) Ende(n) sequenziert wird (werden), das (die) distal zu einem gegebenen Initiierungsprimer ist (sind), um die Sequenz eines weiteren Initiierungsprimers zu bestimmen, um dadurch weitere Initiierungsprimerextensionsprodukte auf der Basis von Primerextension des weiteren Initiierungsprimers zu gewinnen.

12. Methode nach einem der vorstehenden Ansprüche, dadurch gekennzeichnet, daß ein Initiierungsprimerextensionsprodukt oder ein Teil davon sequenziert wird, um dadurch das Extensionsprodukt oder einen Teil davon zu charakterisieren.

13. Methode nach Anspruch 11 oder 12 für die Identifizierung eines für einen gegebenen Phänotyp verantwortlichen Genotyps, dadurch gekennzeichnet, daß die Sequenz der den Genotyp enthaltenden Nucleinsäure mit der Sequenz von Nucleinsäure verglichen wird, die keinen solchen Genotyp enthält, um dadurch den für einen gegebenen Phänotyp verantwortlichen Genotyp zu identifizieren.

14. Methode nach Anspruch 13, dadurch gekennzeichnet, daß der Vergleich durch die Verwendung einer Einzelprobe von Nucleinsäure von einer für einen gegebenen Phänotyp obligaten Heterozygoten ausgeführt wird.

15. Methode nach Anspruch 11 oder 12 für die Identifizierung eines Genotyps, der für eine Prädisposition für einen gegebenen Phänotyp, sofern vorhanden, verantwortlich ist oder dazu beiträgt, dadurch gekennzeichnet, daß die Nucleinsäuresequenzen von

einer Vielzahl von durch den zu untersuchenden Phänotyp befallenen Individuen mit den Nucleinsäuresequenzen von einer Vielzahl von Individuen verglichen werden, bei denen kein Zeichen für den zu untersuchenden Phänotyp vorhanden ist, um dadurch einen Genotyp zu identifizieren, der für eine Prädisposition für einen zu untersuchenden Phänotyp, sofern vorhanden, verantwortlich ist oder dazu beiträgt.

16. Methode nach Anspruch 15, dadurch gekennzeichnet, daß die Sequenzunterschiede zwischen 1) einem Pool von Nucleinsäuren von einer Vielzahl von durch einen zu untersuchenden Phänotyp befallenen Individuen und 2) einem Pool von Nucleinsäuren von einer Vielzahl von Individuen, die kein Zeichen des zu untersuchenden Phänotyps zeigen, verglichen werden.

17. Vektorettbibliothekskit für die Amplifikation eines Nucleinsäurefragments mit unbekannter Sequenz durch Primerextension, dadurch gekennzeichnet, daß er umfaßt:

1) mindestens eine Vektorettbibliothek, wobei jede Bibliothek eine Gruppe von Targetnucleinsäurefragment/Vektoretteneinheiten umfaßt, die von Nucleotidsequenzen eines einzelnen Vertreters einer Tier-, Pflanzen- oder Mikroorganismenspecies gewonnen wurden, und

2) einen oder mehrere Initiierungsprimer für die Hybridisierung an eine Initiierungsstartregion der Targetnucleinsäurefragment/Vektoretteneinheiten.

18. Kit nach Anspruch 17 für die Analyse eines für eine Prädisposition für einen gegebenen Phänotyp verantwortlichen oder dazu beitragenden Genotyps, dadurch gekennzeichnet, daß

die Gruppe von Targetnucleinsäurefragment/Vektoretteneinheiten umfaßt:

- 1) einen Pool von Targetnucleinsäurefragment/Vektoretteneinheiten von einer Vielzahl von durch einen zu untersuchenden Phänotyp befallenen Individuen und
- 2) einen Pool von Targetnucleinsäurefragment/Vektoretteneinheiten von einer Vielzahl von Individuen, die kein Zeichen des zu untersuchenden Phänotyps zeigen.

Hierzu 30 Seiten Zeichnungen

Methode für die Amplifikation eines Nucleinsäurefragments
mit unbekannter Sequenz durch Primerextension und Vektorette-
bibliothekskit

Anwendungsgebiet der Erfindung

Die Erfindung betrifft eine Methode für die Amplifikation von Nucleotidsequenzen und Kits dafür. Eine solche Methode ist von besonderem Interesse in bezug auf die Amplifikation von Sequenzen, von denen nur ein Teil bekannt ist, und ermöglicht die schnelle und wirksame Sequenzierung langer Nucleotidsequenzen. Die Methode vermeidet die bislang für die Sequenzierung unbekannter Nucleotidsequenzen notwendigen Klonierungsverfahren rekombinanter DNA. Dadurch gestattet sie sowohl den Nachweis von Polymorphismen zwischen Nucleotidsequenzen verschiedener Allele an einem Genort als auch die gleichzeitige Analyse von Allelen an einem bestimmten Ort in verschiedenen Individuen.

Charakteristik des bekannten Standes der Technik

Nucleotide können als einzelne Nucleotide oder Basenpaare (nachstehend bp abgekürzt) oder in Strängen von Nucleotiden vorkommen, wobei jeder Strang bis zu 10^8 bp oder noch mehr enthält. Es wird zum Beispiel angenommen, daß das Humangenom ungefähr 3×10^9 bp und ein einzelnes Chromosom ungefähr 10^7 bis 10^8 bp enthält. Techniken für die Sequenzierung relativ kurzer Nucleotidsequenzen existieren seit vielen Jahren und schließen das Verfahren von Maxam und Gilbert/Maxam,

A.M. Gilbert, W. (1977) "A new method for sequencing DNA" (Ein neues Verfahren zur DNA-Sequenzierung), Proc. Natl. Acad. Sci., USA 74, 560 - 564, und "Sequencing end-labelled DNA with base-specific chemical cleavages" (Sequenzierung end-markierter DNA mit basenspezifischen chemischen Spaltungen), Methods in Enzymology 65, 499 - 560 (1980)/ ein, bei dem an einem Ende, mit z. B. ^{32}P , radioaktiv markierte einzelsträngige DNA mehreren chemischen Spaltungsprotokollen (beispielsweise mit Dimethylsulfat oder Hydrazin) unterzogen wird, die selektiv an einer Seite eines bestimmten Nucleotids Brüche herbeiführen. Die gewonnenen Fragmente werden durch Elektrophorese auf Acrylamidgelen nach ihrer Größe getrennt und durch Autoradiografie identifiziert. Relativ kurze Nucleotidsequenzen können auch durch die enzymatische "Dideoxy"-Technik von Sanger et al./Proc. Natl. Acad. Sci., USA 74,

5463-7, (1977)/ bestimmt werden, bei der das Klenow-Fragment von DNA-Polymerase I oder T7-DNA-Polymerase oder Taq-DNA-Polymerase verwendet wird, um in Anwesenheit der vier Desoxynucleotidtriphosphate, von denen eines oder mehrere radioaktiv markiert sind, beispielsweise mit ^{32}P , und in vier separaten Inkubationsgemischen, die eine niedrige Konzentration von jeweils einem der vier Didesoxynucleosidtriphosphate enthalten, eine komplementäre Kopie der einzelsträngigen Targetsequenz zu synthetisieren. Eine Population beschnittener (truncated) radioaktiver DNA-Moleküle, die je ein gemeinsames 5'-Ende haben, aber jeweils in der Länge zu einem basenspezifischen 3'-Ende variieren, wird somit bei jeder Reaktion gewonnen. Nach angemessener Inkubation kann die DNA in jedem Gemisch denaturiert und nebeneinander der Elektrophorese unterzogen werden, und die radioaktiven Bande einzelsträngiger DNA können durch Autoradiografie nachgewiesen werden. Die Sequenz der Target-DNA wird dann direkt aus dem Autoradiogramm abgelesen. Außerdem hat die Firma E.I. Du Pont de Nemours & Company (DuPont) kürzlich ein automatisches Gerät zur Sequenzierung von DNA-Strängen auf den Markt gebracht, welches eine Modifikation der obengenannten "Dideoxy"-Technik verkörpert. Diese Modifikation umfaßt die Verwendung der vier mit Fluoreszenzfarbstoffen markierten Didesoxynucleotidtriphosphatterminatoren, wobei jeder Dideoxykettenterminator bei Anregung durch einen Argonlaser Licht einer etwas anderen Wellenlänge aussendet /Science, 238, 336 (1987)/. Alle vier Didesoxynucleosidtriphosphatterminatoren können im gleichen Gefäß eingesetzt werden, weil die Terminatoren durch ihre Emissionsspektren unterschieden werden können. Das entstehende Gemisch von DNA-Fragmenten kann in einer einzelnen Spur eines Polyacrylamidgels elektrophoretisiert werden. Die Sequenz der Target-DNA kann dann automatisch gelesen werden, da die Identität des jede Bande auf dem Gel terminierenden Nucleotids durch seine charakteristische fluoreszierende Emission bestimmt wird.

Es wird eingeschätzt, daß der DuPont-DNA-Sequenzierungsautomat in der Lage ist, etwa 10 000 Nucleotide täglich unter optimalen

Bedingungen zu identifizieren; dagegen identifiziert eine Fachkraft bei Anwendung der unmodifizierten manuellen "Dideoxy"-Technik schätzungsweise im Durchschnitt etwa 50 000 Nucleotide im Jahr. Während das eine wesentliche Verbesserung hinsichtlich Geschwindigkeit und Effektivität der Bestimmung relativ kurzer Nucleotidsequenzen ist, ist der geschwindigkeitsbestimmende Schritt bei der Bestimmung relativ langer Sequenzen, beispielsweise genomischer Sequenzen, von dem Schritt abhängig, der in der Umgangssprache als "Chromosom-Walking" bekannt ist. "Chromosom-Walking" ist ein Verfahren, das die sequentielle Isolierung von überlappende Fragmente tragenden Klonen zur Überspannung eines Segments von beispielsweise einem Chromosom, das größer ist als es in einem Phagen- oder einem Cosmid- oder einem YAC-(yeast artificial chromosome)-Vektor getragen werden kann, umfaßt. Dieses Verfahren gestattet also die Isolierung eines interessierenden Locus, für den keine Sonde zur Verfügung steht, und ist dort von besonderem Nutzen, wo der interessierende Locus bekanntlich mit einem Locus verbunden ist, der identifiziert und kloniert worden ist, wie z. B. ein Gen oder DNA-Marker. Der identifizierte Locus kann dann als Sonde zum Screenen einer genomischen Bibliothek eingesetzt werden und hybridisiert mit einem Fragment, das komplementäre Nucleotidsequenzen enthält und daher überlappende Klone darstellt.

Solche überlappenden Sequenzen können entweder im 5'- oder 3'-Sinn vorliegen. Ein Fragment, das gemäß Identifizierung einen überlappenden Klon darstellt, kann dann selbst als Sonde zum Rescreenen der genomischen Bibliothek verwendet werden und mit allen Fragmenten hybridisieren, die komplementäre Nucleotidsequenzen enthalten, die somit weitere überlappende Klone darstellen. Durch Wiederholung dieses Prozesses können die Nucleotidsequenzen von Regionen, die sich immer weiter von dem ursprünglich identifizierten Locus entfernt befinden, bestimmt werden, bis schließlich der interessierende Locus getroffen wird. Die Technik des "Chromosom-Walking" bringt eine Reihe potentieller Schwierigkeiten mit sich. Ein Beispiel dafür ist die Zeit,

gemessen von der Entdeckung eines Markers für eine Erbkrankheit bis zur Entdeckung des spezifischen Genschadens, der für die Krankheit verantwortlich ist. So wurde beispielsweise im Jahre 1983 ein verbundener genetischer Marker für Huntingtonsche Chorea (D4S1) entdeckt, aber noch heute ist der für diese Krankheit verantwortliche spezifische Genschaden nicht bekannt. Ähnliches gilt für viele andere Erbkrankheiten.

Die Technik des "Chromosom-Walking" leidet besonders unter dem Nachteil, daß Klonierung von genomischer DNA eine Voraussetzung ist. Unter einer Reihe von Umständen kann sich die Klonierung als unmöglich oder zumindest als sehr schwierig erweisen, und in solchen Situationen kommt der "Chromosomen-Walk" zu einem vorzeitigen Ende; A. R. Wyman und K. F. Wertman, in: Methods in Enzymology, Bd. 152, Guide to Molecular Cloning Techniques (Leitfaden für molekulare Klonierungsverfahren), S. L. Berger und A. R. Kummel, Hrsg. Academic Press, San Diego, 1987, S. 173 - 180.

Außerdem ist die Analyse der Fragmente, die gemäß Identifizierung überlappende Klone darstellen, unter anderem angesichts der Anzahl solcher Fragmente, die bei jedem Screenen der genomischen Bibliothek lokalisiert werden können, und der Tatsache, daß die überlappenden Sequenzen entweder im 5'- oder 3'-Sinn vorliegen können, kompliziert.

Ein weiteres ernstliches praktisches Problem beim "Chromosom-Walking" ist das weit verbreitete Auftreten von sich wiederholenden Sequenzen innerhalb der menschlichen und anderen Genome. Wenn der identifizierte Locus also solche sich wiederholenden Elemente enthält, können alle solchen Elemente im Humangenom inkorrekt als überlappende Klone identifiziert werden. Die Analyse gestaltet sich dann äußerst kompliziert.

Außerdem erzeugen alle durch "Chromosom-Walking" gewonnenen überlappenden Klone, wenn sie gemäß vorstehender Beschreibung se-

quenziert wurden, eine Sequenzinformation, die von einem einzelnen klonierten Allel von dem identifizierten Locus abgeleitet wurde. Im Zusammenhang mit der Untersuchung der wichtigen genetischen Unterschiede zwischen einzelnen Vertretern einer Species und ihrer Beziehung zum Phänotyp eines Individuums wäre es von Vorteil, wenn mehr als ein einzelnes Allel auf einmal analysiert und charakterisiert werden könnte.

Ziel der Erfindung

Durch die Erfindung wird eine gegenüber dem Stand der Technik weniger aufwendige und breit anwendbare Methode für die Amplifikation eines Nucleinsäurefragments mit unbekannter Sequenz durch Primerextension sowie dafür erforderliche Kits zur Verfügung gestellt.

Darlegung des Wesens der Erfindung

Der vorliegenden Erfindung liegt die Aufgabe zugrunde, eine Methode für die Amplifikation eines eine unbekannte Sequenz enthaltenden Nucleinsäurefragments, wobei vor allem die aufwendige Klonierung entfallen soll, bereitzustellen.

Die Erfindung basiert auf der Entdeckung einer Methode, die zumindest teilweise die obengenannten Schwierigkeiten beseitigt, indem mittels Spaltung von Target-DNA an spezifischen Stellen außerhalb bekannter Sequenzen gewonnene Fragmente amplifiziert werden, wodurch die Notwendigkeit der Klonierung überwunden wird.

Erfindungsgemäß wird eine Methode für die Amplifikation eines eine unbekannte Sequenz enthaltenden Nucleinsäurefragments

durch Primerextension zur Verfügung gestellt, welche das Spalten einer Targetnucleinsäure zur Gewinnung von Targetnucleinsäurefragmenten, von denen ein Fragment eine Initiierungsstartregion mit bekannter Nucleotidsequenz für die Hybridisierung mit einem Initiierungsprimer enthält, das Herstellen von Targetnucleinsäurefragment/Vektoretteneinheiten aus den Targetnucleinsäurefragmenten durch Ligation jeder Einheit, die eine Vektorettestartregion bekannter Sequenz für die Hybridisierung mit einem Vektoretteprimer aufweist, und das Behandeln der Targetnucleinsäurefragment/Vektoretteneinheiten, zusammen oder schrittweise, mit geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen umfaßt, so daß ein Extensionsprodukt eines Initiierungsprimers synthetisiert wird, das komplementär ist zu einer einzelsträngigen Targetnucleinsäure/Vektoretteneinheit mit einer Initiierungsstartregion, an die ein Initiierungsprimer hybridisiert ist, der so ausgewählt ist, daß er im wesentlichen zu der Initiierungsstartregion komplementär ist, während kein solches Extensionsprodukt synthetisiert wird, welches zu einzelsträngigen Targetnucleinsäurefragment/Vektoretteneinheiten, die keine solche Initiierungsstartregion aufwei-

sen, komplementär ist.

Auf Wunsch kann das Extensionsprodukt der Amplifikation in Anwesenheit eines Vektoretteprimers unterzogen werden, der so ausgewählt wird, daß er im wesentlichen zu der Vektorettestartregion komplementär ist.

Die Targetnucleinsäurefragment/Vektoretteeinheiten werden also mit Initiierungsprimer und, wenn das Initiierungsprimer-Extensionsprodukt amplifiziert werden soll, wie es z. B. durch R.K. Saiki et al., Science 239, 487 - 491 (1987) beschrieben wird, zusätzlich mit Vektoretteprimer behandelt. Wo kein Vektoretteprimer verwendet wird, kann arithmetische oder lineare Amplifikation (nachstehend als lineare Amplifikation angegeben) durch Hybridisierung des Initiierungsprimers an die Initiierungsstartregion und anschließende Primerextension in Anwesenheit von geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen und Denaturierung erzielt werden. Dieser Prozeß des Startens, der Primerextension und Denaturierung kann so oft wiederholt werden, wie es zur Erreichung des gewünschten Amplifikationsgrades angemessen ist. Vorzugsweise wird die Amplifikation jedoch in Anwesenheit von sowohl Initiierungs- als auch Vektoretteprimern unter Anwendung des PCR-Verfahrens, auf das oben Bezug genommen wurde und das nachstehend definiert wird, bewirkt.

Nach einem weiteren Merkmal der Erfindung wird ein Kit für die Amplifikation eines Nucleinsäurefragments unbekannter Sequenz durch Primerextension zur Verfügung gestellt, der umfaßt:

1. Mittel zur Spaltung einer Targetnucleinsäure an einer spezifischen Stelle zur Gewinnung eines Targetnucleinsäurefragments;
2. eine Vektorette, die für die Ligation an ein Targetnucleinsäurefragment angepaßt ist, das durch Einsatz der Mittel zum

Spalten einer Targetnucleinsäure gemäß Definition in (1) gewonnen wurde, um dadurch im Einsatz eine Targetnucleinsäurefragment/Vektoretteeinheit zu bilden, wobei die Vektorette eine Vektorettestartregion bekannter Sequenz für die Hybridisierung mit einem Vektoretteprimer aufweist;

3. jedes der vier verschiedenen Nucleosidtriphosphate; und
4. ein Agens für die Polymerisation der Nucleosidtriphosphate von (3).

Eine Vektorettestartregion kann im Vektoretteile einer Targetnucleinsäurefragment/Vektoretteeinheit, wie nachstehend definiert, vorhanden sein oder nicht vorhanden sein. So kann die Vektorette (2) im Kit der Erfindung selbst keine Vektorettestartregion enthalten, vorausgesetzt, daß bei Verwendung eine Targetnucleinsäurefragment/Vektoretteeinheit gebildet wird, in der der Vektoretteile eine Vektorettestartregion enthält. So können solche Einheiten zum Beispiel entweder eine Vektorettestartregion im Vektoretteile der Targetnucleinsäurefragment/Vektoretteeinheit aufweisen, die durch Ligation gebildet wurde, oder sie können eine Vektorettestartregion haben, die nur als Resultat von Primerextension eines Initiierungsprimers gemäß nachstehender Beschreibung entsteht. Bezüge in der gesamten Beschreibung auf "Targetnucleinsäurefragment/Vektoretteeinheit" (auch als "Vektoretteeinheit" bezeichnet), wie nachstehend definiert, sind in diesem Sinne zu verstehen.

Vorzugsweise enthält der Kit zusätzlich einen Vektoretteprimer mit einer Nucleotidsequenz, die im wesentlichen zu der Vektorettestartregion der Targetnucleinsäurefragment/Vektoretteeinheit komplementär ist. Die Anwesenheit eines solchen Vektoretteprimers in dem erfindungsgemäßen Kit gestattet die Ausführung der Amplifikation der Targetnucleinsäurefragment/Vektoretteeinheit durch PCR-Techniken (gemäß nachstehender Definition), wenn das gewünscht wird.

Der Kit kann noch besser eine Reihe ineinandergeschachtelter ("nested") Vektoretteprimer (gemäß nachstehender Definition) enthalten, die zum Beispiel, wo notwendig oder erwünscht, für sekundäre Amplifikationsreaktionen und/oder für die direkte Sequenzierung der Amplifikationsprodukte gemäß Beschreibung von U.B. Gyllenstein und H. Ehrlich in Proc. Natl. Acad. Sci. USA 85, 7652 - 7656 (1988) verwendet werden können. Es versteht sich, daß alle diese Vektoretteprimer gut als direkte Sequenzierungsprimer für die durch lineare Amplifikation unter alleiniger Verwendung eines Initiierungsprimers gewonnenen distalen Enden von Fragmenten nützlich sein können.

Da die zu untersuchende Targetnucleinsäure normalerweise dem Nutzer des Kits zur Verfügung steht, werden Initiierungsprimer normalerweise nicht im Kit vorhanden sein, sondern können vom Nutzer des Kits hergestellt werden.

Wenn es allerdings gewünscht wird, kann der erfindungsgemäße Kit zusätzlich Initiierungsprimer und auf Wunsch auch ineinandergeschachtelte Initiierungsprimer enthalten.

Der erfindungsgemäße Kit kann vorteilhafterweise auch Puffer für die Ausführung der erfindungsgemäßen Methode enthalten, wobei ein besonderes Merkmal des Kits beispielsweise die Anwesenheit von Puffern zum Variieren der Kalium-, Magnesium- und Nucleosidtriphosphatkonzentrationen des Reaktionsgemisches ist. Diese letztgenannten Puffer können für die Bestimmung der optimalen Bedingungen für nachfolgende Zyklen der erfindungsgemäßen Methode wünschenswert sein.

Vorteilhafterweise kann der Kit mehr als ein Mittel, d. h. eine Vielzahl von Mitteln, zum Spalten von je einer Targetnucleinsäure an spezifischen Stellen umfassen. Wo eine Vielzahl solcher Mittel anwesend ist, umfaßt der Kit normalerweise für jedes dieser anwesenden Mittel eine andere Vektorette, um die Bildung von Targetnucleinsäurefragment/Vektoretteeinheiten in bezug auf jede

Gruppe von Targetnucleinsäurefragmenten, die gewonnen werden können, zu gestatten. Obwohl die Erfindung nicht auf diese Anwendung beschränkt ist, wird die Mehrheit der verschiedenen Vektoretten vorteilhafterweise zu einem Vektoretteprimer komplementäre Sequenzen teilen. In solchen Fällen wird ein einzelner Vektoretteprimer Amplifikation aus einer Vielzahl von Targetnucleinsäurespaltstellen ermöglichen.

Bei einer bevorzugten Ausführung enthält der erfindungsgemäße Kit also zusätzlich einen oder mehrere der unter den folgenden ausgewählten Bestandteile: einen oder mehrere Initiierungsprimer, ineinandergeschachtelte(n) Vektoretteprimer, ineinandergeschachtelte(n) Initiierungsprimer, Sequenzierungsprimer, Puffer für die Durchführung der erfindungsgemäßen Methode und Puffer zum Variieren der Magnesium-, Kalium- und Nucleosidtriphosphatkonzentrationen.

Nach einem weiteren Merkmal der Erfindung wird ein Vektorettebibliothekskit für die Amplifikation eines Nucleinsäurefragments unbekannter Sequenz durch Primerextension zur Verfügung gestellt, der umfaßt:

1. mindestens eine Vektorettebibliothek, wobei jede Bibliothek eine Gruppe von Targetnucleinsäurefragment/Vektoretteeinheiten umfaßt, die von Nucleotidsequenzen eines einzelnen Vertreters einer Tier-, Pflanzen- oder Mikroorganismenspecies gewonnen wurde; und

2. einen oder mehrere Initiierungsprimer für die Hybridisierung an die Initiierungsstartregion der Targetnucleinsäurefragment/Vektoretteeinheiten.

Der Vektorettebibliothekskit umfaßt vorzugsweise eine Vielzahl von Vektorettebibliotheken.

Die Targetnucleinsäurefragment/Vektoretteeinheiten können zum Beispiel entweder direkt aus den gewünschten Species oder indi-

rekt aus einer solchen Species nach anfänglicher Klonierung in Plasmid-, Phagen-, Cosmid- oder YAC-(yeast artificial chromosome)-Vektoren hergestellt werden. Bei der verwendeten Tier-, Pflanzen- oder Mikroorganismenspecies handelt es sich vorzugsweise um die Species Mensch, aber es kann auch jede andere Tier-species sowie Pflanzen- oder Mikroorganismenspecies wie z. B. Bakterien, Viren, Hefe oder Parasiten verwendet werden. Die Nucleotidsequenzen stammen vorzugsweise von genomischer DNA, können aber auch von sortierten Chromosomen oder Klonen stammen.

Die verwendeten Vektoretteneinheiten können entweder durch eine Einfachspaltungsmethode und/oder multiple Vektoretteneinheiten können durch Mehrfachspaltungsmethoden abgeleitet werden; zusammen oder getrennt können sie eine Vektorettbibliothek oder eine Vielzahl von Vektorettbibliotheken (gemäß nachstehender Definition) bilden.

Die verwendeten Targetnucleinsäurefragmente können von einer Reihe unterschiedlicher Quellen abgeleitet werden. So können die Targetnucleinsäurefragmente zum Beispiel von einzelnen Individuen einer Tier-, Pflanzen- oder Mikroorganismenspecies gewonnen werden, beispielsweise von für ihre Species typischen Individuen. Die Targetnucleinsäurefragmente können auch von einzelnen Individuen gewonnen werden, die bekanntlich für einen gegebenen Genort heterozygot sind, beispielsweise einen Ort, der Zystofibrose oder eine andere Erbkrankheit verursacht. Die Targetnucleinsäurefragmente können auch von einzelnen Individuen gewonnen werden, von denen bekannt ist, daß sie für einen gegebenen Genort, zum Beispiel Zystofibrose oder eine andere Erbkrankheit homozygot sind. Die Targetnucleinsäurefragmente können auch von einzelnen Individuen gewonnen werden, von denen bekannt ist, daß sie für einen gegebenen Genort, zum Beispiel Zystofibrose oder eine andere Erbkrankheit, normale Homozygoten sind. Die Targetnucleinsäurefragmente können auch von Gruppen von Individuen (im Gegensatz zu Einzelindividuen) mit(einem) gemeinsamen Phänotyp(en) gewonnen werden. Die Nucleinsäure oder

Gewebe von jedem Vertreter einer Gruppe, die einen Phänotyp teilt, kann auf Wunsch gepoolt werden. Jede Gruppe von Individuen besteht aus mindestens zwei und vorteilhafterweise aus weniger als 1000, beispielsweise aus 50 bis 500. Vektoretteneinheiten können aus den Targetnucleinsäurefragmenten hergestellt werden, und die Vektoretteneinheiten können gepoolt oder separat zur Bildung von Vektorettbibliotheken verwendet werden. Der gemeinsame Phänotyp kann auf Wunsch eine Krankheit oder Krankheitsprädisposition, obligates Tragen einer Erbkrankheit oder ein normaler Zustand ohne ein Zeichen der Krankheit oder Krankheitsprädisposition sein.

Der Vergleich der mit der erfindungsgemäßen Methode gewonnenen Nucleotidsequenzen identifiziert alle gemeinsamen genetischen Varianten in der Population, die beispielsweise mit einer gegebenen Krankheit oder Krankheitsprädisposition "assoziiert" sind. Man wird erkennen, daß das beträchtlich über den Umfang der früher unter Einsatz der RFLP-Technologie versuchten detaillierten Analyse hinausgeht.

Der erfindungsgemäße Vektorettbibliothekskit enthält zusätzlich vorzugsweise Vektorettprimer und vorteilhafterweise einen oder mehrere unter ineinandergeschachtelten Initiierungsprimern, ineinandergeschachtelten Vektorettprimern und Sequenzierungsprimern ausgewählte Primer.

Der Vektorettbibliothekskit kann auch geeigneterweise jedes einzelne von vier unterschiedlichen Nucleosidtriphosphaten und ein Agens für die Polymerisation der Nucleosidtriphosphate enthalten. Auf Wunsch kann der Vektorettbibliothekskit ferner Puffer für die Ausführung der Erfindung und/oder Puffer zum Variieren der Magnesium-, Kalium- und Nucleosidtriphosphatkonzentrationen enthalten. Diese letzteren Puffer können zum Bestimmen der optimalen Bedingungen für nachfolgende Zyklen des erfindungsgemäßen Verfahrens erwünscht sein.

Die Erfindung ist breit anwendbar. So kann die Erfindung zum Bei-

spiel angewendet werden, um Nucleotidsequenzen zu identifizieren, die für einen speziellen Mikroorganismus charakteristisch sind, wie z. B. einen Mikroorganismus, der für eine Krankheit bei Pflanzen oder Tieren, insbesondere bei Menschen, verantwortlich ist. Solche Mikroorganismen sind zum Beispiel ein Pilz, eine Hefe, ein Bakterium oder ein Virus sowie ein Parasit (wie z. B. Plasmodium (Malaria) oder Trypanosom (Schlafkrankheit)). Die erfindungsgemäße Methode kann auch angewendet werden, um Nucleotidsequenzen zu identifizieren, die für Arzneimittelresistenz, zum Beispiel Antibiotikumresistenz bei einem gegebenen Organismus verantwortlich sind. So gestattet die Erfindung die Erzeugung von Sonden für die Diagnose von Krankheiten bei Tieren oder Pflanzen und ermöglicht außerdem die Produktion von Sonden für die Diagnose von Arzneimittelresistenz.

Die Erfindung kann beispielsweise auch angewendet werden, um nachzuweisen: 1. Nucleotidsequenzvariationen zum Beispiel Punktmutationen, die für Erbkrankheiten bei Pflanzen, aber speziell bei Tieren, insbesondere beim Menschen, verantwortlich sind; 2. Nucleotidsequenzvariationen, zum Beispiel Deletionen, die für neoplastische Krankheiten bei Tieren, insbesondere beim Menschen, verantwortlich sind; 3. Nucleotidsequenzvariationen, die für Prädispositionen gegenüber Krankheiten oder Störungen bei Pflanzen, aber speziell bei Tieren, insbesondere beim Menschen, verantwortlich sind; und 4. Nucleotidsequenzvariationen, die für ein spezifisches Merkmal wie z. B. ein erwünschtes Merkmal, zum Beispiel bei Pflanzen Blütenfarbe, Ernteertrag oder Herbizidresistenz, verantwortlich sind.

Die Erfindung kann auch bei Tieruntersuchungen Anwendung finden, beispielsweise bei der Überwachung transgener Tiere.

Außerdem ist die Erfindung von besonderem Interesse bei der Sequenzierung eines Tiergenoms, insbesondere des Humangenoms, beispielsweise zur Identifizierung bislang unbekannter Polypeptide, die in vivo, insbesondere im Menschen, produziert werden können und die beispielsweise von therapeutischem Interesse sein können.

Die Erfindung kann auch bei der Erzeugung pharmazeutischer Proteine von Interesse sein, beispielsweise für die Sequenzierung von Regulationsregionen von Genen oder Expressionssystemen.

In Zusammenhang mit der Sequenzierung großer Genome, zum Beispiel des Humangenoms, wird weithin akzeptiert, daß die allgemein wahrgenommene Hauptschranke die Anfertigung physischer Karten darstellt, die überlappende "contigs" isolierter Cosmidklone umfassen. Die gegenwärtigen Strategien für solche Kartierungsprojekte wurden in "Mapping our Genes: Genome Projects: How Big, How Fast?" (Kartierung unserer Gene: Genomprojekte: Wie groß, wie schnell?) Congress of the United States, Office of Technology Assessment, The John Hopkins University Press, Baltimore and London (1988) im Überblick dargestellt. Weitere Details sind enthalten in: "Mapping the Human Genome: Experimental Approaches for Cloning and Ordering DNA Fragments" (Kartierung des Humangenoms: Experimentelle Methoden für die Klonierung und Ordnung von DNA-Fragmenten", R. M. Myer, Department of Physiology, University of California, San Francisco, USA, in: "Mapping our Genes Contractor Reports, Bd. 2, Bestell-Nr. PB 88-162 805/AS, verfügbar unter dem Schutz der US OTA vom National Technical Information Service (NTIS) 5285 Port Royal Road, Springfield, VA22161, USA. Eine Methode zur Kartierung des viel kleineren E.Coli-Genoms wird dargestellt bei C. L. Smith et al. (Science, 236, 1448 - 1453. 1987) und Y. Kohara et al. (Cell, 50, 495 - 508, 1987).

Die Erfindung wird zum Beispiel auch angewendet, um große Nucleotidfragmente zu sequenzieren, die durch Klonierung in Cosmiden und YACs (yeast artificial chromosomes) gereinigt wurden. Cosmide können verwendet werden, um Nucleotidfragmente von z. B. bis zu etwa 45 kb zu klonieren, und YACs können eingesetzt werden, um Fragmente mit einer durchschnittlichen Größe von über 300 kb zu klonieren (Nucleic Acids Research 17, 3425 - 3433 (1989) R. Anand et al.). Die Anwendbarkeit der Erfindung für die Sequenzierung solcher großen Nucleotidfragmente gestattet die

schnelle Sequenzierung sehr großer Nucleotidsequenzen.

Es ist ein besonders vorteilhafter Aspekt der Erfindung, daß die unter Anwendung der erfindungsgemäßen Methode ausgeführte Sequenzanalyse keine Klonierung der Targetnucleinsäure in sich schließt. In der Praxis bedeutet das, daß alle in der Targetnucleinsäureprobe vorhandenen Allele gleichzeitig analysiert werden können. Die Vorteile dieser Methode für die Identifizierung von Heterozygoten sowie homozygot Normalen und Mutanten ist früher beschrieben worden, wobei die direkte Sequenzierung von PCR-Produkten zur Analyse von Mutationen (z. B. Punktmutationen), die Erbkrankheiten verursachen, angewendet wurde. (Nucleic Acids Res., 16, 8233 - 8243 (1988) von C.R. Noewton et al.). Eine Punktmutation wird im Heterozygoten leicht sichtbar gemacht, weil zwei unterschiedliche Basen entsprechende komigrierende Bande in zwei Spuren von "Didesoxy"-Sequenzierungsgel erscheinen. Die Homozygoten (normal oder mutant) bei diesem Basenpaar zeigen eine einzelne Bande, die der einen oder anderen Base in dieser Position auf einem Sequenzierungsgel entspricht.

Früher bestand eine Voraussetzung für diese Art von Sequenzanalyse auf PCR-Basis darin, daß die Sequenz der Region von Interesse im voraus durch an sich bekannte Methoden bestimmt wurde. Das gestattet die Konstruktion von PCR-Primern für die Amplifikation und die nachfolgende sich wiederholende Sequenzierung der Region von Interesse. Die Erfindung gestattet es, eine solche Analyse mit Targetnucleinsäurefragmenten zu unternehmen, deren Sequenz nicht bereits vollkommen bestimmt ist. Sequenzierung unter Anwendung der Methode der Erfindung, beispielsweise an menschlicher genomischer DNA von einem Individuum zeigt das volle Ausmaß polymorpher Differenzen zwischen den diploiden Chromosomen. Ist von dem Individuum bekannt, daß es obligat heterozygot für eine unbekannte Mutation an einem gegebenen Genort ist, dann enthüllt die Sequenzierung beider Allele (vorzugsweise gleichzeitige Sequenzierung) nach der erfindungsgemäßen Methode unumgänglich alle vorhandenen polymorphen Basen, die daher Ursachen eines beobach-

teten Phänotyps sein könnten. Die Bedeutung individueller polymorpher Basenveränderungen kann durch nachfolgende Untersuchung der gleichen Targetnucleinsäuresequenzen in bekannten homozygoten Normalen oder Mutanten (sowie in anderen obligat Heterozygoten) bestätigt oder widerlegt werden. Man wird ferner erkennen, daß die Erfindung einen beträchtlichen Fortschritt gegenüber jetzigen Methoden auf der Basis von RFLP (restriction fragment length polymorphism - Restriktionsfragmentlängenpolymorphismus) für die Untersuchung von Variationen zwischen Chromosomen darstellt. In jenem Fall werden nur die Basenveränderungen nachgewiesen, die eine Restriktionsendonucleaseerkennungsstelle selten schaffen oder zerstören. Es sind Techniken für die leichte Analyse von Punktmutationen, die nicht in einem RFLP resultieren, verfügbar (Nucleic Acids Res., 17 Nr. 7, S. 2503 - 2516 (1989) von C. R. Newton et al.).

Ein weiteres Merkmal der Erfindung ist ihre Nützlichkeit bei der Analyse von polygenetischen oder viele Faktoren betreffenden Krankheiten. So wäre es zum Beispiel vorteilhaft festzustellen, ob bestimmte polymorphe Varianten an bestimmten Genorten eine Prädisposition für die Entwicklung von solchen viele Faktoren betreffenden polygenetischen Krankheiten wie Arteriosklerose, Hypertonie, Diabetes, Schizophrenie etc. anzeigen. Frühere Versuche zur Lösung dieses Problems unter Anwendung von RFLP-Techniken haben sich im allgemeinen als erfolglos erwiesen. Das ist wahrscheinlich nicht überraschend, da sich solche Studien auf die Untersuchung bestimmter Polymorphismen stützten, die auf einer großen Zahl arbeitsintensiver Southern-Blots mittels Sonden einzelner Kandidaten beruhten. So war der Bereich des Gesamtpolymorphismus im Humangenom, der tatsächlich auf Verknüpfung mit Krankheitsprädisposition untersucht wurde, in Wirklichkeit verschwindend klein. (Siehe "Molecular Approaches to Human Polygenic Disease" (Molekulare Methoden für die polygene Krankheit beim Menschen), Ciba Foundation Symposium 130, John Wiley, Chichester, 1987). Mit der vorliegenden Erfindung wird es möglich, Targetnucleinsäuren von einer Reihe von Individuen

mit einer gegebenen Krankheit, Krankheitsprädisposition oder einem anderen Phänotyp zu mischen. Aus solchen gepoolten Targetnucleinsäuren, zum Beispiel gepoolter menschlicher genomischer DNA, können Vektorettbibliotheken gemäß nachstehender Definition konstruiert werden. Die Erfindung gestattet die prozessive Sequenzierung dieser gepoolten genomischen DNA. Gemeinsame Polymorphismen werden ähnlich sichtbar wie bei den Beobachtungen mit einem einzelnen Heterozygoten. Wiederum bedeutet das Fehlen jeglicher Notwendigkeit der Klonierung rekombinanter DNA, daß alle im gepoolten genomischen Material vorhandenen Allele gleichzeitig analysiert werden können. Die Anzahl der im Pool vorhandenen Einzelproben ist nicht begrenzt. Die Anwendung der erfindungsgemäßen Methode zeigt einfach einen Consensus hinsichtlich Genotyp plus Polymorphismus. Ein Vergleich der auf diese Weise erzielten Ergebnisse mit ähnlichen Ergebnissen von der Analyse gepoolter genomischer DNA von Individuen, die die gegebene Krankheit, Krankheitsprädisposition oder einen anderen Phänotyp nicht zeigen, gestattet die Identifizierung von Polymorphismen, die mit der Krankheit oder dem Phänotyp (oder umgekehrt mit einem Fehlen der Krankheit oder des Phänotyps) assoziiert sind. Die Analyse ist nicht auf den kleinen Bereich des Humangenoms dicht bei Kandidatengenomen begrenzt. Diese Art der Analyse unter Anwendung der erfindungsgemäßen Methode läßt sich weiter erleichtern, indem die amplifizierte Produkte zusammengekommen werden, die mit einem gegebenen Initiierungsprimer und einer besonderen Vektorettbibliothek, die mit Targetnucleinsäure von entweder einer "normalen" oder einer "befallenen" gepoolten Population von Individuen hergestellt wurde, gewonnen wurden. Bestehen zwischen den Populationen tatsächlich offene kundige polymorphe Unterschiede, dann würde Mischen der amplifizierte Produkte, Denaturierung und Rehybridisierung (re-annealing) doppelsträngige Amplifikationsprodukte mit Basenpaaren mit fehlender Übereinstimmung erzeugen. Solche fehlenden Übereinstimmungen (mismatches) lassen sich unter Einsatz von Reagenzien wie OsO₄ oder Hydroxylamin gemäß R.G.H. Cotton, N.R. Rodrigues und R.D. Campbell PNAS, 4397 - 4401, 1988, aufdecken.

Als Hilfe für den Leser folgt ein Glossar bestimmter, in dieser Beschreibung verwendeter Termini.

Der Terminus "Targetnucleinsäure" bezieht sich auf eine Nucleotidsequenz, im allgemeinen genomische DNA, zum Beispiel pflanzliche oder tierische genomische DNA wie menschliche DNA oder bakterielle DNA. Eine solche "Targetnucleinsäure" zum Einsatz bei der erfindungsgemäßen Methode umfaßt normalerweise einen Abschnitt (im allgemeinen einen kleinen Abschnitt) von bekannter Sequenz und im allgemeinen einen viel größeren Abschnitt von unbekannter Sequenz.

Der Terminus "Targetnucleinsäurefragment" bedeutet in dieser Beschreibung ein Fragment einer Targetnucleinsäure, das durch Spaltung (wie nachstehend definiert) von Targetnucleinsäure gewonnen wurde. Der Terminus "Targetnucleinsäurefragment" ist deshalb nicht auf ein solches Fragment beschränkt, das durch die Verwendung einer Restriktionsendonuclease gewonnen wurde. Außerdem kann ein solches Fragment zum Beispiel einen Abschnitt mit bekannter Sequenz und im allgemeinen einen größeren Abschnitt mit unbekannter Sequenz umfassen, oder das Fragment kann von unbekannter Sequenz sein. Wo das Fragment von unbekannter Sequenz ist, können Spaltung und Erzeugung von Targetnucleinsäurefragment/Vektoretteneinheiten gemäß nachstehender Beschreibung ausgeführt werden. Unter solchen Umständen können Zufalls-DNA-Sonden aus dem viel größeren genomischen Fragment durch lineare Amplifikation mit einem "Zufalls"-Initiierungsprimer oder Amplifikation mit einem Zufalls-Initiierungsprimer plus Vektorettensprimer erzeugt werden. Durch spezifische Amplifikation entstehen Fragmente in zufälliger Weise, die gereinigt und bei der genetischen Kartierung verwendet werden können, vgl.: Cell 51, 319 - 337 (1987) H. Donis-Keller et al.

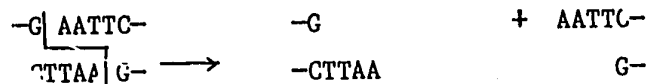
Der Terminus "Amplifikation" bezieht sich hierin auf die Replikation einer Nucleotidsequenz und/oder ihre komplementäre Sequenz durch nichtbiologische Mittel und schließt somit Amplifi-

kation unter Verwendung eines Initiierungsprimers allein ein, wobei ein solcher Primer an die Initiierungsstartregion einer Targetnucleinsäurefragment/Vektoretteneinheit hybridisiert wird, Primerextension in Anwesenheit von geeigneten Nucleotidtriphosphaten unter Hybridisierungsbedingungen und einem Agens für die Polymerisation von Nucleosidtriphosphaten bewirkt wird, sich Denaturierung an eine solche Primerextension anschließt, und dieser Prozeß von Starten, Primerextension und Denaturierung so oft wiederholt wird, wie es zur Erzielung eines gewünschten Amplifikationsgrades angemessen ist. Der Terminus "Amplifikation" umfaßt auch die Replikation durch Polymerasekettenreaktionstechniken (PCR-Techniken), wie sie beispielsweise von R.K. Saiki et al. in Science 239, 487 - 491 (1987) und in den US-Patenten Nr. 4,683,195 und 4,683,202 beschrieben wurden, wobei ein Initiierungsprimer und ein Vektorettprimer gemäß nachstehender Definition verwendet werden und die Expressionspolymerasekettenreaktionstechnik oder PCR-Technik, wie sie hier angewendet werden, sich auf solche Techniken beziehen, wie sie in diesen Referenzen beschrieben werden.

Der Terminus "nicht-biologisch" schließt hier nur Amplifikation durch direkte Klonierung und Vermehrung von Bakterienkolonien aus. Man wird also erkennen, daß der Terminus Amplifikation hierin solche Amplifikationsprozesse einschließt, wie sie in der PCT-Patentveröffentlichung WO 87/06270 (oder Biotechnology Bd. 6, Oktober 1988), PCT-Patentveröffentlichung WO 88/10315 oder PCT-Patentveröffentlichung WO 89/01050 beschrieben werden.

Der Terminus "Spalten" betrifft hierin die Spaltung einer Nucleinsäure an einer spezifischen Stelle. Eine solche Spaltung wird geeigneterweise unter Verwendung einer Restriktionsendonuclease, vorzugsweise eines 6bp-Cutters, der eine bekannte Erkennungssequenz und ein bekanntes Spaltungsmuster und folglich das Merkmal aufweist, DNA an spezifischen Stellen zu spalten, bewirkt.

In dieser Hinsicht wird die Position dieser "spezifischen Stellen" in einer gegebenen Targetnucleinsäure im Verhältnis zur Position eines anderen bekannten Elementes in der Targetnucleinsäure normalerweise nicht bekannt sein, aber die Sequenz der spezifischen Stellen wird ebenso bekannt sein wie ihre Spaltungsmuster. Folglich werden die terminalen Sequenzen der gewonnenen Restriktionsfragmente bekannt sein. So erkennt zum Beispiel die Restriktionsendonuclease EcoRI die Sequenz:



und spaltet eine solche Sequenz wie angegeben, so daß Restriktionsfragmente mit den angegebenen kohäsiven Enden entstehen. Da die Sequenz des kohäsiven Endes bekannt ist, ist es möglich, auf der Basis dieser Kenntnis Targetnucleinsäurefragment/Vektoreteeinheiten gemäß nachstehender Beschreibung in bezug auf den Ausdruck "Targetnucleinsäurefragment/Vektoreteeinheit" zu erzeugen. Mittel zur Spaltung einer Targetnucleinsäure an einer anderen spezifischen Stelle als einer Restriktionsendonuclease können eingesetzt werden, aber die Verwendung einer Restriktionsendonuclease wird bevorzugt. Jede passende Restriktionsendonuclease kann verwendet werden. Beispiele für einzelne Restriktionsendonucleasen schließen jene ein, die im einzelnen angegeben sind in:

Nucleic Acids Research, Sequences Supplement, Band 16, 1988, S. r271-r313, und in: Current Protocols in Molecular Biology 1987-1988" (Laufende Protokolle der Molekularbiologie 1987-1988), herausgegeben von Ausubel, F.M., Brent, R., Kingston, R.E., Moor, D.D., Smith, J.A., Seidman, J.G und Struhl, K., Wiley Interscience, Section 3, Tabelle 3.1.1. Restriktionsendonucleasen, die in der Lage sind, Fragmente mit kohäsiven Enden zu erzeugen, wie z. B. EcoRI, Hind III und XbaI, sind geeignet, egal, ob die kohäsiven Enden 5'- oder 3'-Überhänge haben. Es versteht sich auch, daß ebenso wie durch die Verwendung von Restriktionsendo-

nucleasen, die in der Lage sind, Fragmente mit entweder einem überhängenden kohäsiven 3'- oder 5'-Ende zu erzeugen, die Produktion von Targetnucleinsäurefragment/Vektoretteneinheiten auch unter Einsatz von Restriktionsendonucleasen möglich ist, die stumpfendige Fragmente in Verbindung mit geeigneten stumpfendigen Vektoretten gemäß nachstehender Definition erzeugen. Andere Mittel zur Spaltung einer Targetnucleinsäure bei (einer) spezifischen Stelle(n) als durch Standardrestriktionsendonuclease-digestion sind im Fachgebiet bekannt und schließen beispielsweise die Verwendung von Adapterprimern und Restriktionsenzymen der Klasse II-S (S.C. Kim et al., Science, 240, 504-506 (1988), W. Szybalski, Gene, 40, 169 (1985); A.J. Podhajska und W. Szybalski, Gene, 40, 175 (1985)) sowie verschiedene chemische Methoden ein (B.L. Iverson und P.B. Dervan, J. Amer. Chem. Soc., 109, 1241-1243 (1987); G.B. Dreyer und P.B. Dervan, Proc. Natl. Acad. Sci. USA, 82, 968 (1985); V.V. Vlassov et al., Nucleic Acids Res., 14, 4065 (1986); H.E. Moser und P.B. Dervan, Science, 238, 645 (1987); D.R. Corey und P.G. Schultz, Science, 238, 1401 (1987); J.P. Sluka et al., Science, 238, 1129 (1987)).

Außerdem ist es möglich, Targetnucleinsäurefragmente stumpfendig zu machen, und zwar entweder durch DNA-Polymerase vermittelte Einfüllung/Reparatur irgendwelcher überhängenden kohäsiven 5'-Enden oder durch Entfernung von 3'- und/oder 5'-Überhängen unter Verwendung von durch S1-Nuclease vermittelter Einzelstrangdigestion, unabhängig davon, ob sie durch Restriktionsenzymdigestion oder andere (z. B. chemische) Mittel erzeugt werden. Alle diese stumpfendigen Fragmente können durch Anheftung der entsprechenden stumpfendigen Vektorette in Targetnucleinsäurefragment/Vektoretteneinheiten umgewandelt werden.

Der Ausdruck "Initiierungsstartregion" bedeutet, wie er hier verwendet wird, den Teil eines gespaltenen, zum Beispiel Restriktionsenzym-digierten Targetnucleinsäurefragments, der eine bekannte Nucleotidsequenz aufweist und an den bei Verwendung ein Initiierungsprimer und, auf Wunsch, ineinandergeschach-

telte(r) Initiierungsprimer (wie nachstehend definiert), zum Beispiel überlappende(r) ineinandergeschachtelte(r) Initiierungsprimer hybridisieren kann (können). Im allgemeinen wird die erfindungsgemäße Methode so ausgeführt, daß nur eines der vorhandenen Targetnucleinsäurefragmente eine Initiierungsstartregion hat. Eine Ausnahme wäre die Verwendung eines Gemisches aus einer Vielzahl von verschiedenen Vektorettbibliotheken gemäß nachstehender Definition.

Der Ausdruck "Vektorettstartregion" bedeutet hierin jenen Teil der Targetnucleinsäurefragment/Vektoretteneinheit, der eine bekannte Nucleotidsequenz aufweist, wie es durch die Vektorette selbst definiert ist, und an den bei Verwendung der Vektorettprimer und, auf Wunsch, ineinandergeschachtelte(r) Vektorettprimer (gemäß nachfolgender Definition), zum Beispiel überlappende(r) ineinandergeschachtelte(r) Vektorettprimer hybridisieren kann (können). Die Vektorettstartregion ist in dem Strang komplementär zu dem Strang vorhanden, der die Initiierungsstartregion enthält. In dieser Hinsicht kann die "Vektorettstartregion", an die der Vektorettprimer und, auf Wunsch, ineinandergeschachtelte(r) Vektorettprimer hybridisieren, entweder in der Targetnucleinsäurefragment/Vektoretteneinheit vorhanden sein, die durch Ligation hergestellt wurde, oder in einer Targetnucleinsäurefragment/Vektoretteneinheit, die durch Primerextension eines Initiierungsprimers hergestellt wurde, oder in beiden. Somit wird man erkennen, daß der Vektorettprimer und, auf Wunsch, der (die) ineinandergeschachtelte(r) Vektorettprimer zum Einsatz bei der erfindungsgemäßen Methode für die Hybridisierung mit einer Vektorettstartregion ausgewählt werden können, die zum Beispiel erst im Anschluß an die Primerextension eines Vektorettprimers und, auf Wunsch, ineinandergeschachtelten(r) Initiierungsprimers generiert werden kann. Eine Folge davon ist, daß beispielsweise die Vektorette selbst kein vollkommen selbstkomplementäres doppelsträngiges DNA-Fragment zu sein braucht.

Der Terminus "Primer" bezieht sich hierin auf ein Oligonucleotid, und zwar entweder ein natürlich vorkommendes wie in einem gereinigten Restriktionsdigest oder ein synthetisch erzeugtes, das in der Lage ist, als Initiierungspunkt der Synthese zu wirken, wenn es unter Bedingungen gebracht wird, unter denen die Synthese eines Primerextensionsproduktes, das zu einem Nucleinsäurestrang komplementär ist, induziert wird, d. h. in Anwesenheit geeigneter Nucleosidtriphosphate und eines Agens für die Polymerisation wie z. B. DNA-Polymerase in einem geeigneten Puffer ("Puffer" schließt pH, Ionenstärke, Kofaktoren etc. ein) und bei einer geeigneten Temperatur.

Der Primer ist für maximale Wirksamkeit bei der Extension vorzugsweise einzelsträngig, kann aber wahlweise auch doppelsträngig sein. Ist der Primer doppelsträngig, wird er vor seiner Verwendung zur Herstellung von Extensionsprodukten zuerst behandelt, um seine Stränge zu trennen. Vorzugsweise ist der Primer ein Oligodesoxyribonucleotid. Der Primer muß zum Starten der Synthese von Extensionsprodukten in Anwesenheit des Agens für die Polymerisation ausreichend lang sein. Die exakten Primerlängen sind von vielen Faktoren einschließlich Temperatur und Primerquelle und Verwendungsmethode abhängig. In Abhängigkeit von der Komplexität der Targetsequenz werden die Initiierungs- und Vektoretteprimer typischerweise 15 bis 35 Nucleotide enthalten, obwohl sie auch mehr oder weniger Nucleotide enthalten können. Kurze Primermoleküle erfordern im allgemeinen niedrigere Temperaturen zur Bildung ausreichend stabiler Hybridkomplexe mit der Matrize.

Der Terminus "Initiierungsprimer" bezieht sich hierin auf einen Primer, der zur Hybridisierung mit der Initiierungsstartregion gemäß vorstehender Definition fähig ist. Bei Verwendung beispielsweise mit Vektoretteneinheiten, die aus totaler menschlicher genomischer DNA hergestellt werden, kann es besser sein, wenn der "Initiierungsprimer" länger als 15 bis 17 Nucleotide ist, um zu vermeiden, daß zufälliges Hybridisieren oder Starten an im Human-genom vorhandenen Sequenzen erfolgt, die zufällig mit denen der

Initiierungsstartregion übereinstimmen.

Der Terminus "Vektoretteprimer" bezieht sich hierin auf einen Primer, der zur Hybridisierung an die Vektorettestartregion der Targetnucleinsäurefragment/Vektoretteeinheit fähig ist. Der "Vektoretteprimer" wird eine derartige Nucleotidsequenz aufweisen, daß er in der Lage ist, an ein Initiierungsprimerextensionsprodukt, nach Trennung von seinem Komplement, zu hybridisieren, wodurch das Initiierungsprimerextensionsprodukt als Matrize für die Synthese eines Extensionsproduktes des Vektoretteprimers dient und dabei die Amplifikation erleichtert. Da die erfindungsgemäße Methode im allgemeinen so ausgeführt wird, daß nur eine der Targetnucleinsäurefragment/Vektoretteeinheiten eine Initiierungsstartregion aufweist, wird nur diese Einheit der Amplifikation unterzogen. Jene Targetnucleinsäurefragment/Vektoretteeinheiten, die keine Initiierungsstartregion aufweisen, werden nicht zur PCR-Amplifikation in der Lage sein, da, wenn auch die Bildung eines Vektoretteprimerextensionsproduktes möglich sein kann, kein Initiierungsprimer in der Lage sein wird, an das Vektoretteprimerextensionsprodukt zu hybridisieren, weil keine Initiierungsstartregion vorhanden ist und somit keine PCR-Amplifikation möglich ist.

Im Einsatz wird es vorgezogen, wenn die Synthese von Vektoretteprimeramplifikationsprodukten von der anfänglichen Synthese eines Extensionsproduktes des Initiierungsprimers abhängig ist. Das vermeidet die Bildung großer Mengen nichtamplifizierbarer Vektoretteprimerextensionsprodukte, von denen erwartet werden kann, daß sie die verfügbaren Nucleosidtriphosphate oder andere Kofaktoren im Reaktionsgemisch nachteilig erschöpfen.

Der Terminus "ineinandergeschachtelter Primer" bedeutet hier einen Primer, der durch einen oder mehrere Basenpaare in der 3'-Richtung weg vom 5'-Terminus des Initiierungsprimers oder in der 3'-Richtung weg vom 5'-Terminus des Vektoretteprimers oder in der 3'-Richtung weg vom 5'-Terminus beider solcher Initiierungs- und

Vektoretteprimer verschoben ist. Es versteht sich, daß die Sequenz des oder der ineinandergeschachtelten Primer(s) notwendigerweise aus der Sequenz, die zu den bekannten Initiierungs- oder Vektorettestartregionen komplementär ist, oder aus diesen beiden Regionen ausgewählt wird.

Der Terminus "Targetnucleinsäurefragment/Vektoretteeinheit" (hierin auch als "Vektoretteeinheit" bezeichnet) bezieht sich in dieser Beschreibung auf eine Nucleotidsequenz, zum Beispiel eine DNA-Sequenz, die ein Targetnucleinsäurefragment und einen Abschnitt mit bekannter Nucleotidsequenz, beispielsweise eine DNA-Sequenz, umfaßt, wobei der Abschnitt in einzelsträngiger Form zur Hybridisierung an einen Vektoretteprimer gemäß vorstehender Definition fähig ist. In dieser Hinsicht wird man verstehen, daß die "Targetnucleinsäurefragment/Vektoretteeinheit" zur Hybridisierung an einen Vektoretteprimer in der Lage sein kann, entweder aufgrund des Vorhandenseins eines Abschnitts mit bekannter Nucleotidsequenz in der obengenannten Einheit, wobei die Sequenz im wesentlichen zu der Sequenz des Vektoretteprimers komplementär ist, oder aufgrund der Fähigkeit eines Initiierungsprimerextensionsproduktes - auf der Basis eines Stranges der obengenannten Einheit als Matrize - eine mit der Sequenz des Vektoretteprimers im wesentlichen komplementäre Nucleotidsequenz zu umfassen. Es versteht sich in diesem Zusammenhang, daß die "Targetnucleinsäurefragment/Vektoretteeinheit" so aussieht, daß sie in einem Strang mindestens teilweise im wesentlichen die gleiche Sequenz aufweist wie der Vektoretteprimer. Dieser Strang enthält auch die Initiierungsstartregion.

Der Abschnitt mit bekannter Nucleotidsequenz, zum Beispiel DNA-Sequenz, kann von jeder geeigneten Quelle abgeleitet werden, vorausgesetzt, daß er die obengenannte Forderung erfüllt, daß er in einzelsträngiger Form an den Vektoretteprimer hybridisierungsfähig ist. So kann die Vektorette zum Beispiel separat unter Verwendung eines DNA-Synthesizers hergestellt und die gewonnene Vektorette an das (die) Targetnucleinsäurefragment(e) ligiert

werden, um die Targetnucleinsäurefragment/Vektoreteeinheit(en) zu gewinnen. Insofern wird die Vektorette in geeigneter Weise für die Ligation an das (die) Nucleinsäurefragment(e) angepaßt, beispielsweise hybridisiert ein kohäsives Ende an dem (den) Targetnucleinsäurefragment(en) mit einem kohäsiven Ende an der Vektorette zur Bildung einer Vektoreteeinheit gemäß vorstehender Definition, oder ein stumpfendiges Targetnucleinsäurefragment kann zur Bildung einer Vektoreteeinheit gemäß vorstehender Definition an eine stumpfendige Vektorette ligiert werden. Es ist jedoch nicht notwendig, daß die Vektorette vor der Ligation mit dem Targetnucleinsäurefragment vorgebildet wird, wengleich das bevorzugt sein kann. So kann die obengenannte Einheit zum Beispiel in einem entsprechenden Fall durch Ligation einer einzelsträngigen DNA an das Targetnucleinsäurefragment hergestellt werden, indem beispielsweise der Überhang des kohäsiven Endes benutzt wird, um eine erste einzelsträngige DNA daran zu knüpfen, und dann gestattet wird, daß eine zweite einzelsträngige DNA so ligiert wird, daß die erwünschte Einheit entsteht.

Bei einem Ausführungsbeispiel der Erfindung enthält die Targetnucleinsäurefragment/Vektoreteeinheit einen blockierenden Vektoretteabschnitt. Der Terminus "blockierende Vektorette" bezieht sich hierin auf eine Vektorette oder den Vektoretteabschnitt einer Targetnucleinsäurefragment/Vektoreteeinheit, in der eine oder beide freie terminale Basen in modifizierter Form vorliegen, um Ligation von Nucleotiden daran zu verhindern oder um Primerextension zu verhindern, zum Beispiel in Anwesenheit geeigneter Nucleosidtriphosphate und eines Mittels für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen. Solche Modifikationen sind an sich bekannt und können zum Beispiel in der Anwesenheit eines Didesoxynucleosids bestehen. So kann zum Beispiel die doppelsträngige blockierende Vektorette ein 3'-terminales Didesoxynucleosid wie Didesoxyadenosin (ddA) aufweisen. Eine solche Modifikation kann auch Ribonucleoside einschließen, bei denen das Diol der Ribose gespalten ist, z. B. Periodat. Wahlweise kann ein 3'-Desoxynucleosid, beispielsweise ein 3'-De-

so Nucleosid, beispielsweise ein 3'-Desoxyadenosinrest am 3'-Terminus angefügt werden, wobei z. B. Cordycepintriphosphat und terminale Transferase verwendet werden. Als weitere Alternative können, durch an sich bekannte Methoden, chemisch eine 3'-Amino- oder 3'-Thiofunktionalität am 3'-Ende eingeführt werden.

Bei einem weiteren Ausführungsbeispiel der Erfindung umfaßt eine Vektorette oder ein Vektoretteabschnitt zwei teilweise hybridisierte einzelsträngige Sequenzen, die einen solchen Grad an Nichtkomplementarität besitzen, daß die Vektoretteprimerextension nicht unter Einsatz einer solchen Vektorette bewirkt werden kann.

Der Terminus "Nucleosidtriphosphat" bezieht sich hierin auf die Triphosphate von Nucleosiden, die entweder in DNA oder RNA vorhanden sind und daher Nucleoside einschließen, die Adenin, Cytosin, Guanin, Thymin und Uracil als Base umfassen, wobei die Zuckerkomponente Desoxyribose oder Ribose ist. Im allgemeinen werden die Desoxyribonucleoside in Kombination mit einer DNA-Polymerase verwendet. Es versteht sich jedoch, daß andere modifizierte Basen, die mit einer der herkömmlichen Basen Adenin, Cytosin, Guanin, Thymin und Uracil zur Basenpaarung fähig sind, eingesetzt werden können. Solche modifizierten Basen schließen beispielsweise 7-Deazaguanin und Hypoxanthin ein.

Der Terminus "Nucleotid", wie er hierin verwendet wird, kann sich auf Nucleotide beziehen, die entweder in DNA oder RNA vorhanden sind und daher Nucleotide einschließen, die Adenin, Cytosin, Guanin, Thymin und Uracil als Base umfassen, wobei die Zuckerkomponente Desoxyribose oder Ribose ist. Es versteht sich jedoch, daß andere modifizierte Basen, die zur Basenpaarung mit einer der herkömmlichen Basen Adenin, Cytosin, Guanin, Thymin und Uracil fähig sind, in dem bei der Erfindung eingesetzten Initiierungsprimer und Vektoretteprimer verwendet werden können. Solche modifizierten Basen schließen beispielsweise 7-Deazaguanin und Hypoxanthin ein.

Das Agens für die Polymerisation der Nucleosidtriphosphate kann jede Verbindung oder jedes System sein, welches die Synthese von Primerextensionsprodukten einschließlich Enzymen bewirkt. Zu geeigneten Enzymen für diesen Zweck zählen zum Beispiel: E.coli-DNA-Polymerase I (Richardson C.C. et al., J. Biol. Chem. 239, 222 (1964)), Klenow-Fragment von E.coli-DNA-Polymerase I (Jacobsen, H. et al., Eur. J. Biochem. 45, 623-627 (1974)), T4-DNA-Polymerase (Panet, A. et al., Biochemistry 12, 5045-5050 (1973)), T7-DNA-Polymerase (Tabor, S. und Richardson C.C., Proc. Natl. Acad. Sci. USA 84, 4767-4771 (1987)), andere verfügbare DNA-Polymerasen, Reverse Transkriptase und andere Enzyme einschließlich thermostabile Enzyme. Der Terminus "thermostabiles Enzym" bezieht sich entsprechend seiner Anwendung in dieser Beschreibung auf ein Enzym, das gegenüber Wärme relativ stabil ist und wärmebeständig ist und die Kombination der Nucleotide in der richtigen Weise katalysiert (erleichtert), damit die Primerextensionsprodukte, die zu jedem Nucleinsäurestrang komplementär sind, gebildet werden. Im allgemeinen wird die Synthese am 3'-Ende jedes Primers initiiert und verläuft in der 5'-Richtung den Matrizenstrang entlang, bis die Synthese aufhört, wobei im allgemeinen Moleküle unterschiedlicher Länge erzeugt werden. In Zusammenhang mit der vorliegenden Erfindung endet die Synthese im allgemeinen an einer Position, die durch die Targetnucleinsäurespaltstelle bestimmt wird, und es entstehen Moleküle der gleichen Länge. Es können jedoch Enzyme, einschließlich thermostabile Enzyme auftreten, die die Synthese am 5'-Ende initiieren und in der anderen Richtung fortfahren, wobei ein ähnliches Verfahren angewendet wird, wie es oben beschrieben ist. Ein bevorzugtes thermostabiles Enzym, welches bei dem erfindungsgemäßen Verfahren eingesetzt werden kann, wird aus Thermus aquaticus extrahiert und gereinigt und hat eine relative Molekülmasse von etwa 86 000 bis 90 000 Dalton, wie in der EP-Veröffentlichung Nr. 237,362 beschrieben (siehe auch EP-Veröffentlichung Nr. 258,017). Thermus-aquaticus-Stamm YT1 ist ohne Beschränkung von der American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland, USA als ATCC 25,104 erhältlich.

Es leuchtet ein, daß DNA-Polymerasen mit besseren oder vorteilhafteren Eigenschaften durch Zufallsmutation von diese Proteine exprimierenden Klonen gewonnen werden können. Es wäre beispielsweise offensichtlich günstiger, eine mutierte Version doppelsträngiger DNA zu erhalten, die Taq-DNA-Polymerase codiert, was die Expression eines Proteins mit besseren Eigenschaften, wie z. B. Fehlen von 5'-Exonucleaseaktivität, zur Folge hätte. Techniken zur Gewinnung solcher erwünschten mutierten Versionen sind bekannt und dem durchschnittlichen Molekularbiologen geläufig.

Der Terminus "komplementär" wird hier in bezug auf Nucleotide verwendet und bedeutet ein Nucleotid, das bei Einschluß in DNA oder RNA mit einem anderen spezifischen Nucleotid ein Basenpaar bildet. So ist Desoxyadenosintriphosphat komplementär zu Thymintriphosphat und Desoxyguanosintriphosphat ist komplementär zu Desoxycytidintriphosphat, während Desoxyguanosintriphosphat nicht zu Thymintriphosphat komplementär ist. Es versteht sich in diesem Zusammenhang, daß Thymintriphosphat und Desoxyguanosintriphosphat zwar unter bestimmten Umständen Basenpaare bilden können, aber für die Zwecke dieser Beschreibung nicht als komplementär angesehen werden.

Die Primer werden hierin so ausgewählt, daß sie "im wesentlichen" komplementär zu den verschiedenen Strängen jeder spezifischen zu extendierenden oder amplifizierenden Sequenz sind. Das bedeutet, daß die Primer ausreichend komplementär sein müssen, um mit ihren entsprechenden Strängen zu hybridisieren. Deshalb brauchen die Primersequenzen nicht die exakte Sequenz ihrer Matrizen widerzuspiegeln, obgleich das normalerweise am besten wäre.

Der Terminus "Vektorettbibliothek" wird hier verwendet, um auf die Vielzahl von Targetnucleinsäurefragment/Vektoretteneinheiten Bezug zu nehmen, die nach Spalten einer Targetnucleinsäure an allen potentiellen Spaltstellen in bezug auf eine gegebene Re-

striktionsendonuclease, die die Targetnucleinsäure enthält, und herstellen von Targetnucleinsäurefragment/Vektoretteneinheiten aus dem Gesamtgemisch von Targetnucleinsäurefragmenten durch Zyklen von Ligation an einen in geeigneter Weise angepaßten Vektoretteabschnitt (und, falls erforderlich, wiederholte Spaltung) gewonnen werden. Im allgemeinen wird nur eine einzelne Vektoretteneinheit in einer gegebenen Vektorettebibliothek die Initiierungsstartregion von Interesse enthalten. Im Falle von menschlicher genomischer DNA, die mit einer eine spezifische 6-bp-Sequenz (ein 6-bp-Cutter) erkennenden Restriktionsendonuclease gespalten wird, beträgt die durchschnittliche Größe der entstehenden Targetnucleinsäurefragmente 4096 bp, und die Targetnucleinsäure erzeugt ungefähr 10^6 solcher Fragmente. Folglich wird eine etwa 10^6 Targetnucleinsäurefragment/Vektoretteneinheiten enthaltende Vektorettebibliothek aus humaner genomischer, mit einer 6-bp-Cutter-Restriktionsendonuclease gespaltener DNA gewonnen, und nur eine solche Vektoretteneinheit in der gesamten Human-Vektorettebibliothek wird eine gegebene Initiierungsstartregion enthalten, die zur Initiierung der Amplifikation in Anwesenheit eines Initiierungsprimers, auf Wunsch eines Vektoretteprimers, geeigneter Nucleosidtriphosphate und eines Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen in der Lage ist.

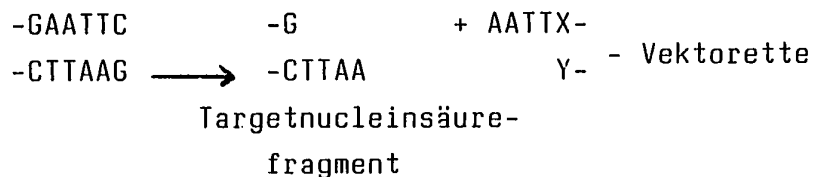
Verschiedene Vektorettebibliotheken können aus der gleichen Targetnucleinsäure durch Spaltung mit unterschiedlichen Restriktionsendonucleasen und Ligation von entsprechend angepaßten Vektoretteabschnitten zur Erzeugung von Targetnucleinsäurefragment/Vektoretteneinheiten hergestellt werden. Alle verfügbaren Restriktionsendonucleasen können, wenn gewünscht, bei diesem Prozeß eingesetzt werden, und in Grenzen kann ein Vektoretteabschnitt an jeder Restriktionsenzymkennungsstelle in der Targetnucleinsäure an Targetnucleinsäurefragmente ligiert werden. Dieses Merkmal ist nicht immer erwünscht, da im Idealfall die Initiierungsstartregion von Interesse in einer gegebenen Vektorettebibliothek durch 100 Basenpaare oder mehr von der Anheftungsstelle des

Vektoretteabschnitts getrennt sein wird.

Das kommt daher, weil Initiierungsprimerextensionsprodukte oder Initiierungsprimer/Vektoretteprimeramplifikationsprodukte unterhalb dieser Größe in der erfindungsgemäßen Praxis so geringe Sequenzinformationen erzeugen, daß sie von wenig Wert für die effiziente Sequenzierung langer Nucleotidsequenzen sind. Außerdem wird die Nucleotidsequenz solcher kleinen Produkte in den Produkten enthalten sein, die unter Verwendung einer Vektorettebibliothek gewonnen wurden, in der der Initiierungsprimer weiter von der Vektoretteabschnittanheftungsstelle entfernt ist. Die Verwendung einer Vielzahl unterschiedlicher Vektorettebibliotheken mit einem besonderen Initiierungsprimer gestattet die Identifizierung jener Bibliotheken, bei denen Extensions- oder Amplifikationsprodukte von angemessener Größe für die Sequenzierung sind. Beispielsweise kann es besonders günstig sein, Initiierungsprimerextensions- oder -amplifikationsprodukte von ungefähr 200 bp, 400 bp, 600 bp, 800 bp, 1000 bp usw. auszuwählen, die mit einem gegebenen Initiierungsprimer von speziellen Vektorettebibliotheken gewonnen wurden. Die Sequenzierung solcher Produkte aus den Vektorettebibliotheken, in denen sie für einen gegebenen Initiierungsprimer vorkommen, unter Verwendung eines Vektorette- oder ineinandergeschachtelten Vektorettesequenzierungsprimers und an sich bekannten Methoden erzeugt wahrscheinlich überlappende Sequenzdaten für eine große Region an der 3'-Seite des Initiierungsprimers. Der Umfang der Sequenzdaten, die in einer Analyserunde von einer Vielzahl von Vektorettebibliotheken mit einem gegebenen Initiierungsprimer erzeugt werden, wird nur durch die Größe der Initiierungsprimerextensions- oder -amplifikationsprodukte, die in der Praxis gewonnen werden können, und/oder durch den Abstand (von der Initiierungsprimerregion) zur entferntesten, in der Vielzahl von Vektorettebibliotheken repräsentierten Restriktionsendonuclease-stelle begrenzt.

In einem vorteilhaften Ausführungsbeispiel des erfindungsgemäßen

Verfahrens werden die Targetnucleinsäurefragment/Vektoretteneinheiten aus Targetnucleinsäurefragmenten durch Ligation so hergestellt, daß die Vektorette nicht von der Targetnucleinsäurefragment/Vektoretteneinheit abgespalten werden kann, die durch das gleiche Agens gebildet wurde, das zur Spaltung der Targetnucleinsäure zur Gewinnung von Targetnucleinsäurefragmenten verwendet wurde. Es ist besonders vorteilhaft, wenn die Targetnucleinsäure mit einer Restriktionsendonuclease gespalten wird, um ein Targetnucleinsäurefragment zur Ligation an eine Vektorette zu erhalten, wobei die Sequenz der Vektorette so ausgewählt wird, daß die Restriktionsendonucleaseerkennungssequenz der Restriktionsendonuclease in der Targetnucleinsäurefragment/Vektoretteneinheit nicht vorhanden ist, zum Beispiel im Fall des oben beschriebenen EcoR1,



wo X jedes Nucleosid außer C ist und Y sein komplementäres Nucleosid ist.

Die Erfindung kann so ausgeführt werden, daß die Synthese von Primerextensionsprodukten oder vorzugsweise Vektoretteprimersamplifikationsprodukten von der anfänglichen Synthese eines Extensionsproduktes des Initiierungsprimers abhängig ist.

Das kann zum Beispiel durch lineare Amplifikation oder vorzugsweise durch den Einsatz eines Vektoretteprimers erzielt werden, der nur in der Lage ist, an das Extensionsprodukt des Initiierungsprimers zu hybridisieren. So umfaßt der Vektoretteteil der Targetnucleinsäurefragment/Vektoretteneinheit einen doppelsträngigen Abschnitt mit einem ersten und einem zweiten Strang, wobei der erste Strang eine terminale Polymerisationsblockierkomponente aufweist und der zweite Strang, der an jenen Strang des Tar-

getnucleinsäurefragments mit der Initiierungsstartregion ligiert ist, einen einzelsträngigen Abschnitt trägt, wobei die terminale Polymerisationsblockierungskomponente so wirkt, daß Extension des ersten Stranges verhindert wird, um ein Komplement zu dem einzelsträngigen Abschnitt des zweiten Stranges in Anwesenheit geeigneter Nucleosidtriphosphate und eines Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen zu bilden. Der Vektoretteprimer ist somit unfähig zur Hybridisierung an die Targetnucleinsäurefragment/Vektoretteeinheit, aber fähig zur Hybridisierung an das Extensionsprodukt des Initiierungsprimers. Also werden bis zur Erzeugung eines Extensionsproduktes des Initiierungsprimers keine Vektoretteprimerextensionsprodukte gewonnen, wonach die gewonnenen Extensionsprodukte normalerweise einen Abschnitt einschließen, der komplementär zu der (den) bekannten Nucleotidsequenz(en) des einzelsträngigen Abschnitts des zweiten Stranges ist. Dieses vorteilhafte Ausführungsbeispiel wird in nachstehender Figur 6 veranschaulicht.

Die Polymerisationsblockierkomponente verhindert erfolgreich die Polymerisation einer Nucleotidsequenz aus einem Primer zur Bildung des Komplements einer Matrizennucleotidsequenz in Anwesenheit eines Agens für die Polymerisation von Nucleosidtriphosphaten wie einer DNA-Polymerase, zum Beispiel Klenow-Fragment von E. coli-DNA-Polymerase I, T7-DNA-Polymerase oder Taq-DNA-Polymerase. Die Polymerisationsblockierkomponente kann jede geeignete Gruppe sein, die für diesen Zweck bekannt ist, wie zum Beispiel ein geeignetes modifiziertes Nucleosid, beispielsweise ein Didessoxynucleosid oder ein 3'-Desoxynucleosid wie Cordycepin oder eine 3'-Amino- oder eine 3'-Thiofunktionalität.

Bei einem weiteren, besonders bevorzugten Ausführungsbeispiel der Erfindung umfaßt der Vektorette teil der Targetnucleinsäurefragment/Vektoretteeinheit einen doppelsträngigen Abschnitt mit einem ersten und zweiten Strang, wobei der zweite Strang des Vektorette teils an jenen Strang des Targetnucleinsäurefragments ligiert

ist, der die Initiierungsstartregion enthält, und die Nucleotidsequenz des ersten Stranges, zweiten Stranges und Vektoretteprimers so ausgewählt werden, daß der Vektoretteprimer zur Hybridisierung an das Komplement des zweiten Stranges, aber nicht an den ersten Strang unter den gleichen Hybridisierungsbedingungen in der Lage ist. Es versteht sich, daß die Anwesenheit einer Polymerisationsblockierkomponente bei diesem Ausführungsbeispiel nicht notwendig ist. Es versteht sich ferner, daß in diesem Fall die Sequenz des Vektoretteprimers im wesentlichen die gleiche sein wird wie die Sequenz mindestens eines Abschnitts des zweiten Stranges der Vektorette. Dieses besonders bevorzugte Ausführungsbeispiel der Erfindung wird nachstehend unter Bezugnahme auf Figur 6 weiter diskutiert.

Ein weiteres bevorzugtes Ausführungsbeispiel der Erfindung umfaßt die Herstellung einer Vielzahl verschiedener Vektorettebibliotheken zum Einsatz mit dem gleichen einzelnen Initiierungsprimer, wobei jede Vektorettebibliothek hergestellt wird durch Spalten von Targetnucleinsäure an unterschiedlichen Spaltstellen und Gewinnen von Targetnucleinsäurefragment/Vektoretteeinheiten aus den Targetnucleinsäurefragmenten durch Ligation, um dadurch die Vektorettebibliothek zu bilden; Behandeln jeder Vektorettebibliothek entweder separat oder zusammen mit geeigneten Nucleosidtriphosphaten und einem Agens zur Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen, um auf diese Weise eine Vielzahl Initiierungsprimerextensionsprodukte auf der Basis des Einsatzes des gleichen einzelnen Initiierungsprimers zu gewinnen.

Die Größe solcher Extensionsprodukte wird durch den Abstand vom Initiierungsprimer zur nächsten 3'-Stelle für die speziellen Spaltungsmittel, zum Beispiel das zur Konstruktion der speziellen Bibliothek verwendete Restriktionsenzym, bestimmt.

Auf Wunsch können eines oder mehrere der Initiierungsprimerextensionsprodukte isoliert und/oder sequenziert werden, oder zumindest kann ein Teil des Extensionsproduktes sequenziert werden.

So kann dieses Ausführungsbeispiel günstig zur Identifizierung eines gewünschten, normalerweise des längsten Targetnucleinsäurefragments, das eine Initiierungsstartregion enthält, angewendet werden, so daß das 3'-terminale Ende in geeigneter Weise mit einem ineinandergeschachtelten Vektoretteprimer gemäß vorstehender Beschreibung sequenziert werden kann, um einen neuen Startpunkt zum weiteren Einsatz der erfindungsgemäßen Methode wie dieses bevorzugten Ausführungsbeispiels zur Verfügung zu stellen.

Die Sequenz des 3'-terminalen Endes des obengenannten längsten Targetnucleinsäurefragments kann somit die Initiierungsstartregion eines neuen Targetnucleinsäurefragments für eine weitere Runde der Bildung von Vektorettebibliothek-Mehrfachinitiierungsprimerextensionsprodukten, Identifizierung des längsten Targetnucleinsäurefragments und Sequenzierung sein.

Bei der Wahl einer neuen Initiierungsstartregion auf der Basis neuer, unter Anwendung der erfindungsgemäßen Methode generierter Sequenzdaten, und zwar am 3'-terminalen Ende eines Targetnucleinsäurefragments, können solche Sequenzdaten in der Regel mit den öffentlich verfügbaren Datensammlungen bekannter Nucleinsäuresequenz (zum Beispiel Genbank, EMBL) verglichen werden, um zu gewährleisten, daß eine vorgeschlagene neue Initiierungsstartregion nicht zufällig genau mit einer bekannten Nucleinsäuresequenz irgendwo in beispielsweise der genomischen DNA von Interesse übereinstimmt. Das tritt offensichtlich höchstwahrscheinlich in jenen Fällen ein, wo das 3'-terminale Ende eines bestimmten Targetnucleinsäurefragments sich wiederholende Elemente wie zum Beispiel Alu-Sequenzen umfaßt. In solchen Fällen ist es von Vorteil, die erfindungsgemäße Methode an einer Vielzahl von Vektorettebibliotheken mit einem gegebenen Initiierungsprimer auszuführen, um zu garantieren, daß mindestens eines der entstehenden Extensionsprodukte ein nichtrepetitives/einzigartiges 3'-terminales Ende für die Auswahl einer weiteren Initiierungsstartregion hat.

Das schrittweise Voranschreiten von einer vorher unbekanntem Initiierungsstartregion zu einer anderen längs einer Targetnucleinsäure, zum Beispiel menschlicher genomischer DNA, kann in geeigneter Form überwacht werden, indem Proben der Targetnucleinsäure verwendet werden, die separat bis zur Vollständigkeit mit den gleichen Restriktionsendonucleasen gespalten werden, wie sie bei der Herstellung von Targetnucleinsäurefragment/Vektoreteeinheiten ("Vektoretebibliotheken gemäß vorstehender Definition) eingesetzt wurden, und der Agarosegelelektrophorese und Southern-Blotting unterzogen werden. Die Untersuchung der so hergestellten Filter mit einem ersten Initiierungsprimer zeigt ein Muster von Banden in Übereinstimmung mit den verschiedenen Restriktionsenzymerkennungsstellen, die diese erste Initiierungsstartregion in der Targetnucleinsäure umgeben. Die Anwendung der erfindungsgemäßen Methode mit einer Vielzahl von Vektoretebibliotheken und diesem ersten Initiierungsprimer erzeugt eine Reihe von Extensionsprodukten, bei denen alle 3'-terminalen Enden jeweils durch die Position im Verhältnis zur Initiierungsstartregion der nächsten Erkennungsstelle für das Restriktionsenzym, das zur Erzeugung der betreffenden Vektoretebibliothek verwendet wurde, definiert werden. So entsteht effektiv eine Karte der Restriktionsstellen zur 3'-Seite eines ersten Initiierungsprimers. Nach anschließender Auswahl einer zweiten neuen Initiierungsstartregion mit vorher unbekannter Sequenz erfolgt die Verbindung mit der ersten Initiierungsstartregion durch erneute Untersuchung der obigen Southern-Blot-Filter mit dem zweiten neuartigen Initiierungsprimer. Das erhaltene Bandenmuster wird identisch sein mit demjenigen, welches mit dem ersten Initiierungsprimer in jenen Fällen gewonnen wurde, wo keine Erkennungsstelle für das betreffende Restriktionsenzym zwischen der ersten und zweiten Initiierungsstartregion liegt. In jenen Fällen, wo eine Erkennungsstelle für das betreffende Restriktionsenzym zwischen den Initiierungsstartregionen, gemäß Beurteilung nach dem Erscheinen von kleineren Extensionsprodukten in der entsprechenden Vektoretebibliothek, auftritt, wird normalerweise

se bei erneuter Untersuchung der Southern-Blot-Filter mit dem zweiten Initiierungsprimer ein Fragment mit unterschiedlicher Größe beobachtet werden. Durch Wiederholung dieses Verfahrens werden Folgerichtigkeit, Genauigkeit und Zuverlässigkeit des schrittweisen Voranschreitens von einer Initiierungsstartregion zu einer anderen längs der Targetnucleinsäure aufrechterhalten und gewährleistet.

Es versteht sich, daß sich die Sequenz der 3'-terminalen Enden all der vielen Initiierungsprimerextensionsprodukte leicht ermitteln läßt, wenn der gleiche Vektoretteprimer oder ineinandergeschachtelte Vektoretteprimer für die Sequenzierung nach an sich bekannten Methoden verwendet werden. Auf diese Weise kann die gesamte Sequenz eines unbekanntem Abschnitts von Target-DNA-Nucleinsäure leicht und systematisch und mit viel größerer Angemessenheit als beispielsweise mit der M13 "Shot-gun-Klonierung" bestimmt werden. Das ist darauf zurückzuführen, daß die Initiierungsprimerextensionsprodukte nach ihrer Größe geordnet werden können und daher die Anordnung ihrer Sequenzen in der ursprünglichen Targetnucleinsäure sichtbar wird. Jedes Initiierungsprimerextensionsprodukt hat eine 5'-Extremität, die durch den Initiierungsprimer bestimmt wird, und eine 3'-Extremität, die durch die nächste 3'-Stelle für das spezielle Spaltungsmittel, zum Beispiel Restriktionsenzym, das bei der Synthese der bestimmten Vektorettebibliothek verwendet wurde, bestimmt wird.

Bei einem bevorzugten Ausführungsbeispiel der Erfindung wird (werden) irgendeines oder alle der gewonnenen Initiierungsprimerextensionsprodukte zumindest an dem (den) Ende(n), die distal zu einem gegebenen Initiierungsprimer liegen, sequenziert (wie nachstehend definiert), um die Sequenz eines weiteren Initiierungsprimers zu bestimmen, um dadurch weitere Initiierungsprimerextensionsprodukte auf der Basis von Primerextension des weiteren Initiierungsprimers zu gewinnen.

Bei einem weiteren bevorzugten Ausführungsbeispiel der Erfindung wird ein Initiierungsprimerextensionsprodukt oder ein Teil davon sequenziert (wie nachstehend definiert), um dadurch das Extensionsprodukt oder einen Teil davon zu charakterisieren.

Wie oben beschrieben, ist eine überaus wichtige Anwendung der Erfindung die Identifizierung eines vorher nichtidentifizierten Genotyps, beispielsweise eines Genschadens, der für einen Phänotyp, z. B. eine Erbkrankheit oder eine Störung, verantwortlich ist, oder die Identifizierung eines vorher nichtidentifizierten Genotyps, zum Beispiel eines Genschadens, der für die Prädisposition für einen Phänotyp, zum Beispiel eine Krankheit, verantwortlich oder ein dazu beitragender Faktor ist.

So kann die erfindungsgemäße Methode zum Beispiel in bezug auf einen Genotyp wie eine Erbkrankheit oder Störung für Nucleinsäure angewendet werden, die den Genotyp nicht enthält (z. B. Genschäden) und für Nucleinsäure, die den Genotyp enthält, z. B. zu untersuchende Genschäden, wobei die Identifizierung des Genotyps, z. B. Genschaden, durch Vergleich der durch Sequenzierung der zwei Nucleinsäureproben erzeugten Informationen erfolgt. Ein solcher Vergleich könnte zum Beispiel auf einfache Weise durch Vergleichen der Sequenzierungsgele, geeigneterweise durch automatisches Scanning, erfolgen. Es versteht sich in dieser Hinsicht, daß die spezifischen Sequenzen nicht an sich bestimmt zu werden brauchen, vorausgesetzt, daß genügend Daten erzeugt werden, um den Nachweis oder die Identifizierung eines Unterschiedes oder von Unterschieden zwischen den Targetnucleinsäureproben zu ermöglichen, und die Termini "Sequenzierung" und "sequenziert" sollen demzufolge hierin nicht nur die spezifische Nucleotidsequenzbestimmung einschließen, sondern auch den Nachweis und die Identifizierung von Sequenzunterschieden ohne spezifische Nucleotidsequenzbestimmung. Es ist angemessen, die erfindungsgemäße Methode für die Targetnucleinsäure eines obligaten Heterozygoten, beispielsweise für die zu untersuchende

Erbkrankheit oder Störung anzuwenden. Notwendigerweise werden sowohl ein normales als auch ein mutiertes Allel für den betreffenden Genort in einem solchen Individuum vorhanden sein, und jene mit dem erfindungsgemäßen Verfahren identifizierten Stellen, wo mehr als ein einzelnes Nucleotid beim Sequenzieren vorhanden ist, sind Kandidaten für den Phänotyp, z. B. Krankheit oder Störung, die Mutation verursacht.

Außerdem wird vermutet, daß bestimmte Genotypen, z. B. Genschäden, Individuen für Phänotypen, z. B. Krankheiten wie prämatüre Arteriosklerose, Hypertonie, Diabetes und Krebs prädisponieren können. Wenn solche Genschäden zum Beispiel identifiziert werden könnten, dann könnten solche Risikopatienten überwacht und jeder Ansatz der Krankheit in einem frühen Stadium behandelt werden. Die erfindungsgemäße Methode kann für die Identifizierung solcher prädisponierenden Genotypen angewendet werden. So kann die erfindungsgemäße Methode auf der einen Seite für die Nucleinsäure einer Vielzahl von Individuen angewendet werden, die mit einem zu untersuchenden Phänotyp behaftet sind, und auf der anderen Seite für die Nucleinsäure einer Vielzahl von Individuen, die kein Anzeichen für den Phänotyp zeigen, wobei die Identifizierung eines Genotyps durch Vergleich der Sequenzen der Nucleinsäureproben erfolgt. Geeigneterweise wird Nucleinsäure von der Vielzahl von Individuen, die mit dem zu untersuchenden Phänotyp behaftet sind, gepoolt und dem erfindungsgemäßen Verfahren unterzogen, und ähnlich wird Nucleinsäure von den Individuen, die kein Anzeichen des Phänotyps zeigen, gepoolt und dem erfindungsgemäßen Verfahren unterzogen. Der Vergleich der Sequenzunterschiede zwischen den beiden Pools identifiziert das Vorhandensein eines prädisponierenden Genotyps, sofern einer oder mehrere anwesend sind. Der Vorteil dieser Technik besteht darin, daß sie die Identifizierung des individuellen prädisponierenden Genotyps unabhängig von der Häufigkeit seines Vorkommens und unabhängig von der Gesamtkompliziertheit und der Anzahl der unterschiedlichen, zu dem Gesamtgenotyp beitragenden Genfaktoren ermöglicht. Wenn also die Anwesenheit einer Kombination

von offensichtlich nicht in Beziehung stehenden Genschäden für die Prädisposition für eine zu untersuchende Krankheit verantwortlich ist oder einen dazu beitragenden Faktor bildet, dann ist die erfindungsgemäße Methode dazu in der Lage, das zu identifizieren.

Ein Ausführungsbeispiel der Erfindung umfaßt die Zirkularisierung eines Targetnucleinsäurefragments mit Termini, die zur Ligation aneinander fähig sind, wobei das Targetnucleinsäurefragment einen Abschnitt mit bekannter Nucleotidsequenz enthält, wobei die bekannte Nucleotidsequenz oder der Abschnitt davon als Initiierungsstartregion für die Hybridisierung mit einem Initiierungsprimer dienen kann, einen Initiierungsprimer an die Initiierungsstartregion hybridisieren kann und die so gebildete Hybride der Primerextension in Anwesenheit von angemessenen Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen unterziehen kann. Zum Beispiel können zwei Initiierungsprimer, die im entgegengesetzten Sinn zueinander orientiert sind als es für PCR normal ist, verwendet werden, um Amplifikation zu bewirken.

In dieser Hinsicht kann Zirkularisierung des Targetnucleinsäurefragments als die Herstellung der Targetnucleinsäurefragment/Vektoretteneinheit durch Ligation angesehen werden, und daher kann zum Beispiel Primerextension auf der Basis des zirkularisierten Targetnucleinsäurefragments als Matrize erfolgen. Auf Wunsch kann das zirkularisierte Targetnucleinsäurefragment jedoch gespalten werden, und Primerextension kann auf der Basis des Spaltprodukts als Matrize erfolgen, wie es beschrieben wird in: Nucleic Acids Research, Bd. 16, S. 8186, 1988, veröffentlicht nach der UK-Patentanmeldung Nr. 8818020.3, für die Verbandspriorität beansprucht wird. Ferner kann das Spaltprodukt auf Wunsch verwendet werden, um eine Targetnucleinsäurefragment/Vektoretteneinheit durch Ligation herzustellen, und Initiierungsprimerextension kann auf der Basis der so gewonnenen Targetnucleinsäurefragment/Vektoretteneinheit als Matrize erfolgen.

Alle diese Ausführungsbeispiele liegen im Geltungsbereich der Erfindung.

So kann zum Beispiel das gebildete zirkularisierte Targetnucleinsäurefragment gespalten werden, um ein Spaltprodukt zu gewinnen, das mindestens einen Abschnitt mit bekannter Nucleotidsequenz enthält, wobei dieser Abschnitt als Initiierungsstartregion für die Hybridisierung mit einem Initiierungsprimer dienen kann, in welchem Fall beispielsweise das zirkularisierte Targetnucleinsäurefragment innerhalb der Region mit bekannter Nucleotidsequenz gespalten werden kann, um ein lineares Molekül mit bekannter Sequenz zu bilden, das an seinen Termini unbekannte Sequenz flankiert; es kann an einem Terminus mit bekannter Nucleotidsequenz gespalten werden, um ein lineares Molekül mit einer bekannten Nucleotidsequenz an einem Terminus und einem bekannten Spaltstellenmuster am anderen Terminus zu bilden, um es für die Ligation an eine Vektorette fähig zu machen; oder noch besser kann es außerhalb der bekannten Nucleotidsequenz gespalten werden, um ein lineares Molekül zu bilden, welches die bekannte Nucleotidsequenz enthält und ein bekanntes Spaltstellenmuster an mindestens einem Terminus, normalerweise beiden Termini aufweist, wodurch es möglich wird, den Terminus oder die Termini zu ligieren, um eine Targetnucleinsäurefragment/Vektoretteeinheit zu bilden. So umfaßt ein bevorzugtes Ausführungsbeispiel der Erfindung die Zirkularisierung eines Targetnucleinsäurefragments mit Termini, die zur Ligation aneinander fähig sind, wobei das Targetnucleinsäurefragment einen Abschnitt aufweist, der als Initiierungsstartregion für die Hybridisierung mit einem Initiierungsprimer dienen kann; die Spaltung des zirkularisierten Targetnucleinsäurefragments außerhalb der bekannten Nucleotidsequenz zur Bildung eines linearen Moleküls, das die bekannte Nucleotidsequenz enthält und ein bekanntes Spaltstellenmuster an mindestens einem Terminus für die Ligation zur Bildung einer Targetnucleinsäurefragment/Vektoretteeinheit aufweist; Bilden der Targetnucleinsäurefragment/Vektoretteeinheit durch Ligation und Behandeln der Targetnuclein-

säurefragment/Vektoreteeinheit zusammen oder nacheinander mit geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen.

Dieses bevorzugte Ausführungsbeispiel der Erfindung ist von Interesse in bezug auf zirkularisierte Targetnucleinsäurefragmente von beispielsweise mindestens 1 kb bis etwa 20 kb, aber es kann von besonderem Interesse in bezug auf größere zirkularisierte Targetnucleinsäurefragmente von zum Beispiel 20 kb bis etwa 120 kb wie zirkularisierte Targetnucleinsäurefragmente von etwa 100 kb sein. Die obere Größengrenze ist von praktischen Erwägungen abhängig und basiert daher auf der maximalen Größe von Targetnucleinsäurefragmenten, die zirkularisiert werden können. Eine solche Zirkularisierung kann durch bekannte Techniken erfolgen /F.S. Collins (1988) Genome Analysis - A practical approach (Genomanalyse - eine praktische Methode) (Editor K. Davies) p73-94, IRL Publishers, Oxford/, zum Beispiel durch Herbeiführung von Ligation bei niedriger Konzentration. Dieses bevorzugte Ausführungsbeispiel kann zum Beispiel ausgeführt werden, indem genomische DNA mit einem Restriktionsenzym digeriert wird, und zwar vorzugsweise mit einem Restriktionsenzym, das relativ selten schneidet (z. B. durchschnittliche Fragmentgröße 10 - 20 kb) wie XbaI, KpnI oder BamHI. Die Fragmente können dann zur Bildung von Ringen selbstligiert werden. Die Ringe können danach mit einem Restriktionsenzym geschnitten werden, vorzugsweise mit einem Restriktionsenzym, das häufig schneidet (zum Beispiel HinfI), und die Fragmente können an eine entsprechende (zum Beispiel HinfI) Vektorette ligiert werden. So kann beispielsweise, wenn die einer bekannten XbaI-Stelle benachbarte Sequenz zur Verfügung steht, die Sequenz, die der nächsten, möglicherweise entfernten XbaI-Stelle benachbart ist, durch Analyse des Vektoretteproduktes gewonnen werden. Auf diese Weise kann eine Reihe von Sprüngen ausgeführt werden, vorausgesetzt, daß zusätzliche Startpunkte für die Methode vorhanden sind. Es versteht sich, daß die Fähigkeit zur Ausführung solcher Sprünge einen beträchtlichen

Vorteil gegenüber den PCR-Techniken, wie vorstehend beschrieben, darstellt, da es möglich sein kann, amplifizierte Vektoretteprodukte in einem beträchtlichen Abstand von der Stelle des Initiierungsprimers zu erhalten. Somit ist die Methode nicht durch die Schwierigkeit begrenzt, die gewöhnlich bei der Herstellung von Standard-PCR-Produkten von mehr als etwa 5 kb auftritt.

Ausführungsbeispiele

Die Erfindung soll nachstehend zu ihrem noch besseren Verständnis anhand von Beispielen mit Bezug auf die dazugehörigen Zeichnungen beschrieben werden. In den Zeichnungen zeigen:

Figur 1(a) eine doppelsträngige Targetnucleinsäure, Figur 1(b) die restringierte Targetnucleinsäure, Figur 1(c) die durch Ligation der Targetnucleinsäurerestriktionsfragmente und der Vektorette gewonnenen Targetnucleinsäurefragment/Vektoretteeinheiten, und Figur 1(d) die Hybridisierung von (i) Initiierungsprimer und (ii) Vektoretteprimer mit den Targetnucleinsäurefragment/Vektoretteeinheiten.

Figur 2 ein Ausführungsbeispiel der Erfindung; Figur 2(a) die bildliche Darstellung von Restriktionsfragmenten, die durch teilweise Digerierung von genomischer DNA gewonnen wurden und der Gelelektrophorese unterzogen wurden. Figur 2(b) das Gemisch unterschiedlicher Typen von Fragmenten mit hoher relativer Molekülmasse, die für die weitere Verarbeitung entsprechend der Größe ausgewählt wurden. Figur 2(c) das Gemisch von Produkten, die nach Ligation zur Bildung von Targetnucleinsäurefragment/Vektoretteeinheiten gewonnen wurden.

Figur 3 ein Ausführungsbeispiel der Erfindung, bei dem versucht wird, die relative Konzentration von Restriktionsfragmenten, die zur Hybridisierung an den Initiierungsprimer fähig sind, zu erhöhen.

Figur 4 ein weiteres Ausführungsbeispiel der Erfindung, bei dem der Einsatz von S1-Nuclease vermieden und versucht wird, die re-

relative Konzentration von Restriktionsfragmenten, die zur Hybridisierung sowohl an den Initiierungsprimer als auch den Vektoretteprimer fähig sind, zu erhöhen, Figur 4(a) das Gemisch von Produkten, die nach der Behandlung der Restriktionsfragmente, die durch komplette Digerierung der Target-DNA-Sequenz von Interesse erzeugt wurden, mit blockierender Vektorette in Anwesenheit von DNA-Ligase gewonnen werden könnten. Figur 4(b) eine blockierende Vektorette zum Einsatz bei der Erfindung nach Digerierung mit dem Restriktionsenzym Eco R1 und Figur 4(c) die Produkte, die durch komplette Digerierung des in Figur 4(a) gezeigten Produktgemisches gewonnen wurden.

Figur 5 zwei unterschiedliche Vektoretteabschnitte zum Einsatz bei der erfindungsgemäßen Methode, wobei Vektoretteabschnitt (i) eine blockierende doppelsträngige Vektorette ist, die einen kurzen oberen (5') Strang mit einer zur blockierenden Polymerisation fähigen 3'-terminalen Gruppe und einen unteren (3') Strang aufweist, der den Schwanz bildet, wobei die 3'-terminale Gruppe des oberen (5') Stranges zur blockierenden Polymerisation in Anwesenheit von Nucleosidtriphosphaten und eines Agens für die Polymerisation der Nucleosidtriphosphate unter Hybridisierungsbedingungen in der Lage ist; Vektoretteabschnitt (ii) ist doppelsträngig, weist aber einen Grad von Nichtkomplementarität zwischen den Strängen auf.

Figur 6 eine schematische Darstellung der Anwendung der Erfindung auf Vektoretteabschnitte von Figur 5.

Figur 7 eine schematische Darstellung der Verwendung von Vektorettebibliotheken.

Figur 8 die schematische Anwendung der Erfindung auf die Amplifikation von zirkularisierten Targetnucleinsäurefragmenten.

Figur 9 die Amplifikation eines Teils des α -1 Antitrypsingens unter Einsatz der Oligonucleotide 58 und 61 (Schritt 6 in Me-

thode I) und einer XbaI-Vektoretteneinheit/Bibliothek.

Figur 10 die Restriktionskarte des Plasmids B3 (pUC8 mit einem 440bp-Insert an der EcoRI-Stelle).

Figur 11 (i) und (ii) Photographien eines Agarosegels, welche die in den nachstehenden Beispielen 1, 5 und 7 gewonnenen amplifizierten Fragmente zeigen.

Figur 12 eine Sequenz aus dem Phenylalaninhydroxylasegen, das Exon 9 flankiert, und die entsprechenden Bindepitionen für die Oligonucleotide 58, 63, 66 und 59.

Figur 13 einen Teil einer Photographie eines Agarosegels, die das nach Beispiel 8 gewonnene amplifizierte Produkt zeigt.

Figur 14 eine Photographie eines Agarosegels, die die nach Beispiel 9 gewonnenen amplifizierten Produkte zeigt.

Figur 15 eine Photographie eines Agarosegels, die die nach Beispiel 10 amplifizierten Produkte zeigt.

Figur 16 eine Photographie eines Agarosegels, die die nach Beispiel 11 gewonnenen amplifizierten Produkte zeigt.

Figur 17 eine Photographie eines Agarosegels, die die nach Beispiel 12 gewonnenen amplifizierten Produkte zeigt.

Figur 18 eine Photographie eines Agarosegels, die die nach Beispiel 13 gewonnenen amplifizierten Produkte zeigt.

Figur 19 eine Photographie eines Agarosegels, die die nach Beispiel 14 gewonnenen amplifizierten Produkte zeigt.

Figur 20 eine Photographie eines Agarosegels, die die nach Beispiel 15 gewonnenen amplifizierten Produkte zeigt.

Figur 21 ein Autoradiogramm eines Sequenzierungsgels, das die Sequenz der in Beispiel 16 erzeugten Produkte zeigt.

Figur 22 die Nucleotidsequenzdaten; die aus den in Beispiel 16 beschriebenen Produkten generiert wurden.

Figur 23 eine Sektion des Phenylalaninhydroxylasegens, das Exon 1 flankiert, und die entsprechenden Positionen der Oligonucleotide 58, 62 und 67 sowie die bei der Konstruktion der Vektoreteinheit verwendete EcoRI-Restriktionsstelle.

Figur 24

(a) eine Photographie eines Agarosegels, welche das nach Beispiel 17 gewonnene primäre Amplifikationsprodukt zeigt;

(b) eine Photographie eines Agarosegels, welche die nach Beispiel 17 gewonnenen sekundären Amplifikationsprodukte zeigt.

Figur 25 die Nucleotidsequenzdaten, die von den in Beispiel 17 beschriebenen Produkten generiert wurden.

25(a) die mit Oligonucleotid 62 gelesene Sequenz (s. nachstehende Definition), Figur 25(b) die mit Oligonucleotid 67 gelesene Sequenz (s. nachstehende Definition) und Figur 25(c) die mit den Oligonucleotiden 62 und 67 gelesene Gesamtsequenz.

Figur 26 eine Sektion des α -1 Antitrypsingens, das Exon 5 flankiert, und die entsprechenden Positionen der Oligonucleotide 58, 60, 68, 69 und 70 sowie die EcoRI-Restriktionsstelle für Hybridisierung (annealing) und Ligation der Vektorette.

Figur 27

(a) eine Photographie eines Agarosegels, welche das nach Beispiel 18 gewonnene primäre Amplifikationsprodukt zeigt;

(b) eine Photographie eines Agarosegels, die das nach Beispiel 18 gewonnene sekundäre Amplifikationsprodukt zeigt;

(c) ein Autoradiogramm eines Nylonfilters, von dem in Figur 27(b) gezeigten Gel geblottet und mit einer α -1 Antitrypsin-

sonde untersucht.

Figur 28 die Nucleotidsequenz von D5, ein anonymes Fragment an Chromosom 7 des Humangenoms, wobei die relativen Positionen der Oligonucleotide 71 und 72 angegeben werden.

Figur 29 eine Restriktionskarte der Region, die D5 enthält. Die entsprechenden Positionen der Oligonucleotide 71 und 72 werden angegeben. Die Positionen der gemäß Beispiel 19 synthetisierten Produkte werden ebenfalls gezeigt.

Figur 30 eine Photographie eines Agarosegels, die die gemäß Beispiel 19 gewonnenen amplifizierten Produkte zeigt.

Figur 31 die Nucleotidsequenz, die von dem Bcl-1-Produkt generiert wurde, das entsprechend der Beschreibung in Beispiel 19 gewonnen wurde. Die Position der aus den Restriktionsenzymdigests vorhergesagten Hind-III- und Hae-III-Stellen wird ebenfalls gezeigt. Die Position von Oligonucleotid 73, das als Initiierungsprimer in nachfolgenden Zyklen von Vektoreinheitamplifikation verwendet würde, wird ebenfalls angegeben.

Figur 32 eine Photographie eines Agarosegels, die die nach Beispiel 20 gewonnenen amplifizierten Produkte zeigt.

Figur 33 die Nucleotidsequenz, die aus dem EcoRI-Produkt generiert wurde, das gemäß Beschreibung in Beispiel 20 gewonnen wurde.

Figur 34 eine Photographie eines Agarosegels, die die Fragmente zeigt, die nach Digestion mit Restriktionsenzymen des EcoRI-Produktes, das gemäß der Beschreibung in Beispiel 20 gewonnen wurde, hergestellt wurden.

Figur 35 eine Photographie eines Agarosegels, die die gemäß Beispiel 21 gewonnenen amplifizierten Produkte zeigt.

Figur 36

- (a) eine Restriktionskarte einer Sektion einer Hind III/Cosmid-Vektorettbibliothek, die die Positionen der Oligonucleotide 58, 76 und 77 anzeigt;
- (b) eine Photographie eines Agarosegels, die die gemäß Beispiel 22 gewonnenen amplifizierten Produkte zeigt.

Figur 37

- (a) eine Restriktionskarte eines Segments von an eine EcoRI-Vektorette ligierter Target-genomischer DNA. Die Positionen der Oligonucleotide 27, 76, 78 und 79 sind angegeben.
- (b) eine Restriktionskarte eines Segments von an eine EcoRI-Vektorette ligierter Target-genomischer DNA. Die Positionen der Oligonucleotide 78, 79 und 84 sind angegeben.
- (c) ein Segment des Phenylalaninhydroxylasegen-Exons 9, das an eine modifizierte EcoRI-Vektorette ligiert ist. Die Positionen der Oligonucleotide 63 und 79 sind angegeben.
- (d) eine Restriktionskarte des Segments von Target-genomischer DNA (Beschreibung in 37(c) oben), die an eine modifizierte EcoRI-Vektorette ligiert ist. Die Positionen der Oligonucleotide 76, 79 und 80 sowie 81 sind angegeben.
- (e) eine Restriktionskarte eines Segments von Target-genomischer DNA (Beschreibung in 37(b) oben), die an eine modifizierte EcoRI-Vektorette ligiert ist. Die Positionen der Oligonucleotide 79, 80, 81 und 84 sind angegeben.
- (f) ein Segment des Phenylalaninhydroxylasegen-Exons 9, das an eine modifizierte EcoRI-Vektorette ligiert ist. Die Positionen der Oligonucleotide 63, 79, 80 und 81 sind angegeben.
- (g) eine Restriktionskarte eines Segments von Target-genomischer DNA, die an eine modifizierte Hind-III-Vektorette ligiert ist. Die Positionen der Oligonucleotide 76, 78 und 79 sind angegeben.
- (h) eine Restriktionskarte eines Segments von Target-genomischer DNA, die an eine modifizierte Hind-III-Vektorette ligiert ist. Die Positionen der Oligonucleotide 76, 79, 81 und 82 sind angegeben.

(i) eine Restriktionskarte eines Segments von Target-genomischer DNA, die an eine modifizierte Hind-III-Vektorette ligiert ist. Die Positionen der Oligonucleotide 76, 79, 81 und 83 sind angegeben.

Figur 38 eine Photographie eines Agarosegels, welche die amplifizierten Produkte zeigt, die nach den Beispielen 23 (a), (b), (c), (d), (e), (f), (g), (h) und (i) gewonnen wurden.

Figur 39

(a) eine Restriktionskarte der Target-genomischen DNA-Sequenz, die die durch Sonde KM19 nachgewiesene polymorphe Pst-I-Stelle enthält. Die polymorphe Pst-I-Stelle ist durch ein Sternchen gekennzeichnet.

(b) die Bildung ringförmiger Produkte durch Ligation von Fragmenten von Pst-I-gespaltenen Target-genomischer DNA.

(c) die Konstruktion von Hind-III-Vektoretteneinheit-Bibliotheken durch Ligation von Oligonucleotiden an durch Hind III gespaltene Pst-I-Ringe.

Figur 40 eine Photographie eines Agarosegels, die die gemäß Beispiel 24 gewonnenen amplifizierten Produkte zeigt.

Figur 1(a) zeigt schematisch eine doppelsträngige Targetnucleinsäure von beispielsweise 100 kb oder mehr, deren Sequenz mit Ausnahme der anfänglichen 5'-Nucleotide (schraffiert dargestellt) z. B. etwa 20 Nucleotide, unbekannt ist.

Ein Ort von Interesse, X, mit unbekannter Sequenz wird in einem unbekanntem Abstand vom 5'-Ende gezeigt. Die Targetnucleinsäure wird der Restriktion (siehe Figur 1(b)) unterzogen, wobei ein 6-bp-Cutter wie EcoRI eingesetzt wird, von dem zu erwarten ist, daß er eine Nucleinsäure im Durchschnitt aller 4096 bp spaltet. Die Erkennungssequenz für EcoRI ist GAATTC, die Spaltstelle wird durch den Pfeil angegeben.

EcoRI spaltet also die Targetnucleinsäure, wobei kohäsive Enden

entstehen. Das kohäsive Ende der Targetnucleinsäurerestriktionsfragmente wird dann an eine entsprechende Vektorette ligiert, wie es in Figur 1(c) dargestellt ist, wobei die Vektorette schraffiert ist. Die Vektorette kann von jeder geeigneten Nucleotidsequenz sein, wie oben beschrieben, wird aber an ihrem 5'-Ende ein Nucleotidsequenzmuster aufweisen, das mit der Gewährleistung der Ligation mit den kohäsiven Enden der oben aufgeführten Restriktionsfragmente vereinbar ist. Wenn also EcoRI als die Restriktionsendonuclease verwendet wird, werden das 3'-Ende der Restriktionsfragmente und das 5'-Ende der Vektorette die folgenden Sequenzen haben:

Restriktions-	5' - G	AATTC - 3'	Vektorette
fragmente	3' - CTAA	G - 5'	

Vorzugsweise wird der Teil des Restriktionsendonucleaseerkennungsmusters, der nicht an der Ligation beteiligt ist, in der Vektorette verändert, um das Erkennungsmuster zu zerstören. Somit ist jede zum Spalten der Targetnucleinsäure anwesende Restriktionsendonuclease unfähig, die einmal gebildeten Restriktionsfragment/Vektoretteeinheiten zu spalten. Wo es sich bei der verwendeten Restriktionsendonuclease um EcoRI handelt, ist daher das 5'-terminale Ende der Vektorette vorzugsweise

5' AATTA 3'	oder 5'AATTT 3'	oder 5' AATTG 3'
T	A	C

wobei der CG-Abschnitt des Erkennungsmusters in AT verändert wird oder CG zu dem im EcoRI-Erkennungsmuster transponiert wird. Die Restriktionsfragment/Vektoretteeinheiten werden dann zusammen oder vorzugsweise nacheinander mit Initiierungsprimer und Vektoretteprimer (wobei der Initiierungsprimer zuerst zugefügt wird) unter Hybridisierungsbedingungen behandelt (siehe Figur 1(d)). Der Initiierungsprimer kann nur an den Teil des einen Targetrestriktionsfragmentes hybridisieren, der eine bekannte Sequenz enthält (siehe Figur 1(d) (i)). Der Vektoretteprimer ist so kon-

struiert, daß er an den Nucleinsäurestrang hybridisiert, der durch Extension des Initiierungsprimers gebildet wird (und, wie in Figur 1(d) gezeigt, den Nucleinsäurestrang, der zu dem Nucleinsäurestrang komplementär ist, an den der Initiierungsprimer hybridisieren kann). Der Vektoretteprimer hybridisiert an die Vektorettestartregion jeder vorhandenen Restriktionsfragment/Vektoretteeinheit (siehe Figur 1(d) (ii), aber in Anwesenheit von geeigneten Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate erfolgt die Amplifikation nur hinsichtlich der Restriktionsfragment/Vektoretteeinheit, an die der Initiierungsprimer hybridisiert, dann erweitert wurde. Es versteht sich, daß vorzugsweise die Hybridisierung (ii) nur bei Einsatz nach der Hybridisierung (i), Extension in Anwesenheit von Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate und nachfolgende Denaturierung erfolgt. Das verhindert eine aus Figur 1(d) hervorgehende Situation, wo zu sehen ist, daß der Vektoretteprimer (ii) hybridisiert an und daher in der Lage ist, ein Extensionsprodukt zu bilden mit jeder Targetnucleinsäurefragment/Vektoretteeinheit, die in dem komplexen Gemisch vorhanden ist. Es ist daher vorzuziehen, zuerst die Vektoretteeinheit mit dem Initiierungsprimer zu behandeln und dann, nach den oben im Detail aufgeführten Schritten, den Vektoretteprimer hinzufügen.

Es wird daher möglich sein, das amplifizierte Targetfragment, und auf Wunsch das ganze Fragment, beispielsweise durch oben beschriebene Techniken zu identifizieren. Wenn gewünscht, braucht jedoch nur das 3'-Ende des amplifizierten Targetrestriktionsfragments sequenziert zu werden, so daß ein neuer Initiierungsprimer konstruiert werden kann. Dieser neue Initiierungsprimer kann dann eingesetzt werden, um Amplifikation eines neuen Targetrestriktionsfragments zu erzielen, das durch Digestion der Targetnucleinsäure mit einer anderen Restriktionsendonuclease, vorzugsweise auch einem 6-bp-Cutter, gewonnen wurde, und das erfindungsgemäße Verfahren kann wiederholt werden. Auf diese Weise kann eine Targetnucleinsäure vom 5'-Ende sequenziert werden,

ohne daß die Notwendigkeit zur Klonierung besteht und wobei die Sequenzierungsrichtung konstant und konsequent in die 5'- bis 3'-Richtung verläuft. Durch Wiederholung der erfindungsgemäßen Methode ist es daher möglich, den obengenannten Locus X von Interesse schnell zu erreichen und zu sequenzieren.

Figur 2 veranschaulicht ein Ausführungsbeispiel der Erfindung. Die interessierende DNA-Sequenz wird partieller Digestion beispielsweise mit EcoRI unterzogen, und die gewonnenen Restriktionsfragmente werden Gelelektrophorese, wie in Figur 2(a) dargestellt, unterzogen. Die Elektrophorese verläuft in die durch den Pfeil angegebene Richtung, und folglich erscheinen Fragmente mit hoher relativer Molekülmasse zum oberen Ende des Gels zu, während die Restriktionsfragmente mit niedrigerer relativer Molekülmasse zum unteren Ende des Gels zu erscheinen..

Die X-Achse stellt eine steigende Konzentration von Restriktionsenzym oder verlängerte Inkubation dar. Restriktionsfragmente von weniger als einer vorbestimmten Größe, zum Beispiel 10 kb, werden als unbrauchbar ausgesondert, und die Restriktionsfragmente mit der höheren relativen Molekülmasse werden zum nächsten Stadium überführt. Diese Restriktionsfragmente mit der höheren relativen Molekülmasse stellen ein Gemisch überlappender partieller Digestionsfragmente, wie in Figur 2(b) gezeigt, dar, wobei die Restriktionsstellen an einem Restriktionsfragmentterminus mit (1-) oder innerhalb des Fragments mit (y) gekennzeichnet sind. Alle Restriktionsfragmente sind dann mit der gleichen einzelnen Restriktionsendonuclease, zum Beispiel EcoRI, geschnitten, und folglich hat dann jedes einzelne Restriktionsfragment kohäsive Enden, die für die zur Herbeiführung der Digestion eingesetzte Restriktionsendonuclease charakteristisch sind. Nur die extremen 5'- und 3'-Fragmente werden ein einzelnes derartiges kohäsives Ende besitzen, während die verbleibenden Fragmente je zwei solche kohäsiven Enden haben werden. Eine Vektorette gemäß vorstehender Definition, bei der die Restriktionsendonucleaseerkennungsequenz zerstört worden ist, kann dann mit

den obengenannten Targetrestriktionsfragmenten in Anwesenheit einer Ligase gemischt werden, um ein Produktgemisch, wie in Figur 2(c) als Beispiel dargestellt, zu gewinnen, wobei die Vektorette als terminale Box dargestellt ist. Es versteht sich, daß die Anwesenheit von Ligase nicht nur die Herbeiführung von Ligation der Targetrestriktionsfragmente mit der Vektorette ermöglicht, sondern auch Ligation der Targetrestriktionsfragmente selbst. Man wird deshalb einsehen, daß es unter den Bedingungen von Ligation in Figur 2 nicht unwahrscheinlich ist, daß die Segmente A und B in den anfänglichen Targetnucleinsäurefragment/Vektoretteeinheiten in der ursprünglichen Targetgenomischen Nucleinsäure nicht benachbart sind, sondern während der Ligation anomal zusammengekommen sind. Die Ligationsprodukte werden dann kompletter Digerierung mit der Restriktionsendonuclease unterzogen, die verwendet wurde, um die partielle Digestion zu bewirken, zum Beispiel EcoRI. Die Produkte vollständiger Digestion werden im allgemeinen doppelsträngige Targetrestriktionsfragmente sein, die eine Vektorette aufweisen, die an jedem Ende des Fragments ligiert ist, und Targetrestriktionsfragmente mit einer Vektorette an nur einem Ende des Fragments. Es wird eine Hintergrundpopulation von Targetrestriktionsfragmenten, beispielsweise Fragmente A und B in Figur 2(c), geben, die keine Vektorette an einem der Enden haben. Die erstgenannten Produkte werden einen geringeren Anteil des gesamten Produktgemisches darstellen. Bei Denaturierung wird nur eine Spur von Produktvektoretteeinheiten in der Lage sein, mit dem Initiierungsprimer zu hybridisieren, aber ein sehr wesentlicher Prozentsatz des Produktgemisches wird zur Hybridisierung mit dem Vektoretteprimer in der Lage sein, wenn nicht eine wesentliche Modifikation der Vektorette erfolgt, wie nachstehend beschrieben. Die Amplifikation des gewünschten Restriktionsfragments wird daher möglich sein, aber der wesentliche Prozentsatz der Vektoretteeinheitsbibliothek, der zur Hybridisierung mit dem Vektoretteprimer fähig ist, kann die Notwendigkeit zur Folge haben, wesentliche Mengen des Vektoretteprimers, des Agens für die Polymerisation der Nucleosidtriphosphate (z. B. Taq-DNA-Polymerase)

und der Nucleosidtriphosphate selbst zu verwenden. Diese Schwierigkeit kann in gewissem Maße durch die Verwendung kleiner Mengen Ausgangsmaterial aus dem Wege geräumt werden, aber es wäre von Vorteil, den relativen Anteil der Targetrestriktionsfragment/Vektoretteeinheiten, die zur Hybridisierung an den Initiierungsprimer in dem Produktgemisch fähig sind, selektiv zu erhöhen.

Das Prinzip für das Herangehen an den Aufbau einer Vektoretteeinheitbibliothek gemäß Darstellung in Figur 2 ist folgendermaßen. Die vollständige Digestion mit einem 6-bp-Schneidenzym erzeugt eine Population von Fragmenten mit einer mittleren Größe von 4096 bp. Es wäre zu erwarten, daß die Ligation von Vektoretteeinheiten an diese Population (ungefähr 10^6 Fragmente von totaler menschlicher genomischer DNA) viele Fragmente mit einer Vektoretteeinheit an jedem Ende erzeugt. Viele davon werden erheblich kleiner sein als 4096 bp, geht man von der Gauss-Verteilung von Restriktionsstellen aus. Im Einsatz könnte ein Vektoretteprimer unter diesen Umständen wie beide Primer in einem Standard-PCR wirken. Große Mengen unechter Amplifikationsprodukte würden entstehen sowie ein von einem Initiierungsprimer herführendes Produkt. Bei der vorstehenden Diskussion wurde auf Methoden Bezug genommen, bei denen die Synthese von Vektoretteprimeramplifikationsprodukten von der anfänglichen Synthese eines Extensionsproduktes des Initiierungsprimers abhängig ist. Bei dem Ausführungsbeispiel von Figur 2 kann die Vektoretteeinheit selbst mit dem Vektoretteprimer hybridisieren und Kopieren von dem Vektoretteprimer initiieren. Der obige Nachteil von PCR-Produktbildung aus Targetnucleinsäurefragmenten mit Vektoretteabschnitten an jedem Ende wird durch einen Größenfraktionierungsschritt überwunden, so daß bei allen solchen Vektoretteeinheitskonstrukten 2 identische Vektoretten vorhanden sind, die durch einen derartigen Abstand (z. B. größer als 10 kb) getrennt sind, daß Standard-PCR so unwirksam ist, daß kein signifikantes Produkt entsteht.

Figur 3 veranschaulicht ein Ausführungsbeispiel, bei dem versucht wird, die relative Konzentration von Restriktionsfragmenten zu erhöhen, die zur Hybridisierung an den Initiierungsprimer in dem Vektoretteneinheitsproduktgemisch, das gemäß der obigen Beschreibung in bezug auf Figur 2 gewonnen wurde, fähig sind. In Figur 3 (a) ist (i) ein Restriktionsfragment, das eine Initiierungsstartregion (1) enthält und einen Vektorettenteil (schraffiert) (2) aufweist, (ii) ein Restriktionsfragment mit einem Vektorettenteil (schraffiert) (2), der an jedem Ende ligiert ist, aber ohne Initiierungsstartregion und (iii) ein Restriktionsfragment ohne Initiierungsstartregion, aber mit einem einzelnen Vektorettenteil (schraffiert) (2). In dieser Hinsicht wird das Produktgemisch nach Denaturierung (siehe Figur 3 b) nur mit Initiierungsprimer (3) behandelt, und zwar in Anwesenheit eines Agens für die Polymerisation von Nucleosidtriphosphaten wie Taq-DNA-Polymerase und Nucleosidtriphosphaten. In diesem Stadium wird kein Vektorettenteil verwendet. Die zur Hybridisierung an den Initiierungsprimer fähigen Targetrestriktionsfragment/Vektoretteneinheiten werden somit selektiv repliziert (siehe Figur 3 (c)); dabei ist ohne die notwendige Initiierungsstartregion keine Replikation von Fragmenten möglich. Das so gewonnene Produktgemisch wird dann der Behandlung mit einer einzelsträngigen spezifischen Endonuclease wie S1-Nuclease (siehe Figur 3 (d)) unterzogen. Diese Endonuclease spaltet die internen Phosphodiesterbindungen in einzelsträngiger DNA auf, was die Produktion von stumpfendiger doppelsträngiger DNA zur Folge hat. Dieser Prozeß erhöht die relative Konzentration der Targetrestriktionsfragment/Vektoretteneinheiten, die zur Hybridisierung an den Initiierungsprimer fähig sind, über die verbleibenden Hintergrundfragmente. Außerdem kann auf Wunsch dieser Prozeßschritte von Denaturierung, Behandlung mit Initiierungsprimer nur in Anwesenheit eines Agens für die Polymerisation von Nucleosidtriphosphaten und Nucleosidtriphosphaten mit anschließender Behandlung des gewonnenen Produktgemisches mit einer einzelsträngigen spezifischen Endonuclease wie S1-Nuclease so oft wiederholt werden, wie es für wünschenswert oder notwendig

erachtet wird, um die relative Konzentration der Targetrestriktionsfragment/Vektoretteneinheiten, die zur Hybridisierung an den Initiierungsprimer fähig sind, zu erhöhen. Wenn die Konzentration solcher Vektoretteneinheiten im Verhältnis zu den verbleibenden Hintergrundfragmenten als ausreichend hoch angesehen wird, dann kann das Produktgemisch zusätzlich mit Vektorettprimer behandelt werden, um Amplifikation zu bewirken, wie es in den US-Patenten Nr. 4,683,195 und 4,683,202 beschrieben ist.

Ein Nachteil der oben beschriebenen Technik besteht darin, daß S1-Nuclease nicht dazu tendiert, so einzelstrangspezifisch zu sein, wie es erwünscht sein könnte und somit auch etwa doppelsträngige DNA abgebaut werden kann.

Außerdem wäre lineare Amplifikation in bezug auf die Technik in Figur 3 nicht angemessen, weil nur eine einzelne Runde von Initiierungsprimerextension S-1-Nuclease-resistente doppelsträngige DNA liefern würde.

Figur 4 veranschaulicht ein weiteres erfindungsgemäßes Ausführungsbeispiel, bei dem die Verwendung von S1-Nuclease vermieden wird und versucht wird, die relative Konzentration von Restriktionsfragmenten zu erhöhen, die zur Hybridisierung an den Initiierungsprimer in dem Produktgemisch in der Lage sind, das durch komplette Restriktionsendonucleasedigestion der interessierenden DNA-Sequenz gewonnen wurde. Der Vorteil dieses Ausführungsbeispiels besteht darin, daß keine Fraktionierung der gewonnenen Fragmente notwendig ist und auch keine Gelelektrophorese als Anfangsschritt erforderlich ist. Nach vollständiger Digestion der interessierenden DNA-Sequenz mit einer Restriktionsendonuclease wie EcoRI werden die gewonnenen Fragmente mit einer blockierenden Vektorette (wie v_orstehend definiert) in Anwesenheit einer DNA-Ligase behandelt, um ein komplettes Gemisch von Produkten zu gewinnen, vgl. Figur 4(a). Es versteht sich, daß die Fragmente A und B in der ursprünglichen DNA-Sequenz nicht unbedingt benachbart sind, sondern in der DNA-Ligasereaktion anomal verbun-

den worden sein können. Die blockierende Vektorette besteht aus einer doppelsträngigen Oligonucleotidsequenz, wie in Figur 4(b) gezeigt, mit einem ersten kohäsiven Ende an einem ihrer Enden, das für die Ligation an die entsprechenden kohäsiven Enden der Restriktionsfragmente angepaßt ist, worin aber der Teil der Restriktionsendonucleaseerkennungsstelle, der nicht an der Ligation teilnimmt, in der Vektorette verändert wird, um die Erkennungsstelle zu zerstören, so daß die einmal gebildete Restriktionsfragment/Vektoretteinheit nicht durch die Restriktionsendonuclease gespalten werden kann, was zu Restriktionsfragmenten führen würde. Das andere Ende der blockierenden Vektorette umfaßt einen Terminus mit einem blockierenden Rest, der vorzugsweise ein 3'-Didesoxynucleosid wie Didesoxyadenosin (ddA) sein könnte, oder irgendein anderer 3'-Rest, so daß Primerextension dort nicht voranschreiten kann, zum Beispiel ein 3'-Desoxynucleosid oder eine andere an sich bekannte chemische Modifikation. Der untere Strang überlappt den oberen Strang, wie in Figur 4(b) gezeigt. Dieser Strang endet vorzugsweise in einem nichtphosphorylierten Rest, um Probleme von Selbstligation zu vermeiden. Die Region (4) besitzt die gleiche Sequenz wie der Vektoretteprimer, so daß der Vektoretteprimer in der Lage ist, mit dem Extensionsprodukt eines Initiierungsprimers zu hybridisieren, aber nicht mit dem nichttextendierbaren oberen Strang der 3'-modifizierten blockierenden Vektorette. Die blockierende Vektorette wird geeigneterweise als zwei separate Stränge hergestellt, die nacheinander hybridisiert werden. Der erste und der zweite Strang werden jeweils separat hergestellt, beispielsweise manuell, oder geeigneterweise unter Einsatz eines DNA-Synthesizers, wobei der erste Strang durch chemische Mittel 5'-phosphoryliert wird, geeigneterweise enzymatisch, z. B. durch die Verwendung von Polynucleotidkinase, und eine terminale Transferase geeigneterweise verwendet wird, um ein 3'-terminales Didesoxynucleotid oder andere modifizierte Nucleotide einzuführen. Wahlweise können andere blockierende Komponenten (zum Beispiel Amino) chemisch am 3'-Ende durch an sich bekannte Methoden eingeführt werden.

Den so hergestellten ersten und zweiten Strang läßt man dann hybridisieren, um die doppelsträngige blockierende Vektorette zu bilden.

Das in Figur 4(a) gezeigte und im Anschluß an die Behandlung der Restriktionsfragmente mit blockierender Vektorette in Anwesenheit von DNA-Ligase gewonnene Produktgemisch wird vollständiger Digestion mit der gleichen Restriktionsendonuclease unterzogen, die zur Schaffung der Restriktionsfragmente vor der Bildung der Restriktionsfragment/Vektoretteeinheiten verwendet wurde, wodurch die Produktarten (i), (ii) und (iii), wie in Figur 4(c) gezeigt, gewonnen werden. Zur Reduzierung der relativen Konzentration von Restriktionsfragmenten (iii), die in dem Gemisch von Restriktionsfragment/Vektoretteeinheiten (i) und (ii) vorhanden sind, kann der oben aufgeführte Zyklus so viele Male wiederholt werden, wie es zur Bildung eines Pools von Restriktionsfragment/Vektoretteeinheiten, zum Beispiel auf der Basis der Verwendung der Restriktionsendonuclease EcoRI, als notwendig oder erwünscht erachtet wird.

Figur 5 zeigt in (i) eine blockierende Vektorette, die aus einem 14-Basen-Topoligonucleotid und einem 42-Basen-Bottomoligonucleotid besteht, die zusammen mit den vier Basenüberhängen (ein 5'-Überhang) hybridisiert (annealed) sind, die zu den in der gespaltenen Target-DNA erzeugten Enden komplementär sind.

N^1 , N^2 , N^3 , N^4 , N^5 und $N^{5'}$ entsprechen der Definition in der nachstehenden Tabelle 1. Das 3'-Ende des 14-Basen-Oligonucleotids wird so modifiziert (X), daß 5' \rightarrow 3'-Extension von jenem Ende nicht unter Verwendung einer der bekannten DNA-Polymerasen, z. B. Klenow-Fragment, T7-DNA-Polymerase (Sequenase) oder Taq-DNA-Polymerase durchgeführt werden kann. Diese modifizierte Base kann ein Didesoxynucleosid oder ein 3'-Desoxynucleosid (z. B. Cordycepin) sein. Der 3'-Zuckerrest kann auf Wunsch durch alle an sich bekannten chemischen Methoden modifiziert werden.

Die 3'-terminale Base des Bottom-42-mers und die komplementäre Base an Position 5 (vom 5'-Ende) im 14-Basen-Topoligonucleotid werden so gewählt, daß nach Ligation der blockierenden Vektorette an die gespaltene Target-DNA das entstehende Hexanucleotid (an der Ligationsstelle) durch die ursprünglich zur Spaltung der genomischen DNA verwendete Restriktionsendonuclease nicht erkannt wird. Das ist von Bedeutung, wenn gewährleistet wird, daß jedes Fragment gespaltener genomischer DNA eine blockierende Vektorette aufweist, die an seine Enden ligiert ist.

Es ist klar, daß die Ligationsprodukte aneinander religierte Fragmente gespaltener genomischer DNA enthalten, und diese werden wieder digeriert, um in einer zweiten Ligrationsreaktion empfangende Enden für die blockierende Vektorette zu erzeugen. Dieser Restriktions- und Ligrationszyklus muß möglicherweise mehr als einmal wiederholt werden, beispielsweise drei- bis viermal, um zu gewährleisten, daß jedes Fragment gespaltener genomischer DNA an jedem Ende eine blockierende Vektorette hat. Es wird bevorzugt, daß jedes Fragment gespaltener genomischer DNA an jedem Ende eine blockierende Vektorette aufweist, um zu verhindern, daß zufälliges Starten von irgendwelchen nichtblockierten 3'-Enden von Genom-DNA-Fragmenten im Reaktionsgemisch nach Denaturierung und Behandlung mit einem Initiierungsprimer erfolgt.

Die in Figur 5(i) gezeigte blockierende Vektorette ist für Target-genomische DNA geeignet, die mit Restriktionsendonucleasen gespalten wurde, die 5'-Überhänge erzeugen. Für Restriktionsendonucleasen, die einen 3'-Überhang erzeugen, müßte die blockierende Vektorette leicht modifiziert werden. So könnte zum Beispiel das 14-Basen-Topoligonucleotid durch einen 10-Basen-Topstrang (dem der 4-Basen-Überhang am 5'-Ende fehlt) ersetzt werden, und der 42-Basen-Bottomoligonucleotidstrang könnte durch einen 46-Basen-Bottomstrang ersetzt werden, wobei sich die 4 Extrabasen am 3'-Ende befinden und komplementär zu den Enden sind, die durch die Restriktionsendonuclease erzeugt wurden, die zum Spalten der genomischen Target-DNA verwendet wurde.

Figur 5 (ii) zeigt einen nichtkomplementären Vektoretteabschnitt, der aus zwei hybridisierten (annealed) einzelsträngigen Sequenzen (einem 57-Basen- und einem 53-Basen-Oligonucleotid) besteht, die einen derartigen Grad an Nichtkomplementarität besitzen, daß die Vektoretteprimerextension nicht sofort stattfinden kann. Die obige Diskussion in bezug auf die blockierende Vektorette von Figur 5 (i) gilt exakt für den nichtkomplementären Vektoretteabschnitt, mit der Ausnahme, daß der nichtkomplementäre Vektoretteabschnitt keine modifizierten Nucleoside zu besitzen braucht.

Figur 6 zeigt eine schematische Darstellung der Anwendung der Erfindung auf die Vektoretteabschnitte von Figur 5 (i). Genomische Target-DNA wird bis zur Vollständigkeit mit einer einzelnen Restriktionsendonuclease digeriert, um Fragmente, wie in (i) gezeigt, zu erzeugen. Die Blockiervektoretteabschnitte /5 in Figur 5(i)/ werden in Anwesenheit von DNA-Ligase an die Enden jedes Fragments gespaltener genomischer Target-DNA ligiert. Amplifikation erfolgt dann nach Denaturierung unter Verwendung der Primer mit bekannter Nucleotidsequenz x und y in Anwesenheit von z. B. Taq-Polymerase.

So stellt x in Figur 6 (iii) den Initiierungsprimer dar und besitzt die gleiche oder zumindest im wesentlichen die gleiche Sequenz wie die in Strang 1 mit IP markierte Region (IP = Initiierungsprimer): x ist in der Lage, in Strang 2 an die Initiierungsstartregion (IPR = initiating priming region) zu hybridisieren. Strang 2, der die Initiierungsstartregion (IPR) enthält, hat auch einen Nucleotidsequenzabschnitt, der gleich oder im wesentlichen gleich ist wie die Nucleotidsequenz des Vektoretteprimers (VP). Im Einsatz liefert die Primerextension von x einen Strang, der eine Vektorettestartregion (VPR = vectorette priming region) enthält, die zur Hybridisierung mit dem durch den Primer y dargestellten Vektoretteprimer (VP) geeignet ist. Da Strang 1 selbst keine Vektorettestartregion enthält und aufgrund des Vorhandenseins der Polymerisationsblockierkomponente nicht der Primerex-

tension unterzogen werden kann und Strang 2 auch keine Vektorettestartregion aufweist, kann keine Primerextension eines Vektoretteprimers erfolgen, bis durch Primerextension des Initiierungsprimers eine Vektorettestartregion geschaffen worden ist.

Es versteht sich daher, daß beim ersten Amplifikationszyklus nur der Primer x ein Extensionsprodukt erzeugen kann, welches sich bis zum Ende des 42-Basen-Oligonucleotids der blockierenden Vektorette (iii) erstreckt. Der Primer y ist im ersten Zyklus redundant, weil für ihn keine komplementäre Sequenz zur Hybridisierung und zur Erzeugung von Extensionsprodukten vorhanden ist. Die Sequenz von Primer y ist genau die gleiche wie Basen 2 bis 31 vom 5'-Ende des 42-Basen-Bottomoligonucleotids in Figur 5 (i) und Basen 13 bis einschließlich 42. des 53-Basen-Bottomoligonucleotids in Figur 5 (ii). Im zweiten Zyklus und danach (v und vi) kann der Primer y an das Extensionsprodukt von Primer x hybridisieren, und die komplementäre Synthese dieses extendierten Produktes (iv) kann weitergehen. Somit kann Primer y nicht an irgendeines der in (iii) gezeigten Ligationenprodukte hybridisieren. Er kann nur an ein im wesentlichen vollständig extendiertes Produkt von Primer x hybridisieren, wobei es sich um die bekannte Nucleotidsequenz an dem Genort von Interesse handelt. So wird nur ein Amplifikationsprodukt, das Primer x, die unbekannte Nucleotidsequenz z und Primer y enthält, erzeugt. Die Nucleotidsequenz von z kann von beiden Enden unter Verwendung der Primer x und y als Sequenzierungsprimer bestimmt werden. Wahlweise können ineinandergeschachtelte (nested) Sequenzierungsprimer x' und y' hergestellt werden. Diese enthalten Sequenz 3' von Primer x oder 3' von Primer y, zum Beispiel die Basen 25 bis 42 vom 5'-Ende des 42-mer-Oligonucleotids in Figur 5(i).

Figur 6 veranschaulicht die erfindungsgemäße Methode durch Bezugnahme auf eine Targenucleinsäurefragment/Vektoretteeinheit, bei der der Vektoretteabschnitt eine Polymerisationsblockierkomponente enthält. Ähnliche Erwägungen gelten dort, wo der Vektoretteabschnitt mindestens eine Region Nichtkomplementarität gemäß vorstehender Beschreibung und Darstellung als Beispiel in Figur 5 (ii) umfaßt. Somit wird Strang 1 die gleiche oder mindestens im wesentlichen die gleiche Sequenz besitzen, wie die mit IP markierte Region. Ein Initiierungsprimer (IP) wird in der Lage sein, an die Initiierungsstartregion (IPR) in Strang 2 zu hybridisieren. Strang 2, der IPR enthält, hat auch einen Nucleotidsequenzabschnitt, der der Nucleotidsequenz des Vektoretteprimers (VP) gleicht oder mindestens im wesentlichen gleicht, aber dieser Abschnitt von Strang 2 hybridisiert nicht an den entsprechenden Abschnitt von Strang 1, und zwar aufgrund des Grades an Nichtkomplementarität, der absichtlich in Strang 1 eingeführt wurde. Der Vektoretteprimer (VP) kann also nicht an Strang 1 hybridisieren, weil Strang 1 keine Vektorettestartregion (VPR) enthält, und er kann auch nicht an Strang 2 hybridisieren, weil Strang 2 einen Sequenzabschnitt enthält, der VP gleicht, aber keine komplementäre Sequenz oder VPR. Somit können keine Vektoretteprimerextensionsprodukte gebildet werden, ehe nicht eine Vektorettestartregion (VPR) durch Primerextension des Initiierungsprimers (IP) geschaffen worden ist.

Der oben beschriebene Prozeß wird dann vorzugsweise unter Verwendung einer anderen Restriktionsendonuclease und anderer entsprechend konstruierter blockierender Vektorette wiederholt, um einen anderen Pool von Restriktionsfragment/Vektoretteeinheiten zu bilden, der eine andere Vektorettebibliothek darstellt.

Dieser Prozeß kann separat für jede gewünschte Zahl verschiedener Restriktionsendonucleasen (z. B. 10 bis 30, geeigneterweise 15 bis 25, vorteilhafterweise etwa 20) und entsprechend konstruierte blockierende Vektoretten wiederholt werden, um

eine Vielzahl von Bibliotheken solcher Hybriden zu bilden. Die Vielzahl von Bibliotheken kann dann unter Hybridisierungsbedingungen in Anwesenheit von Nucleosidtriphosphaten und einem Agens für die Polymerisation der Nucleosidtriphosphate wie Taq-DNA-Polymerase mit einem Initiierungsprimer behandelt werden. Der Zusatz von Vektoretteprimer unter den für Polymerasekettenreaktion (PCR = polymerase chain reaction) geeigneten Bedingungen gemäß Beschreibung in den US-Patenten Nr. 4,683,195 und 4,683,202 hat Amplifikation lediglich der interessierenden Loci zur Folge, und solche amplifizierte Fragmente können der Sequenzierung unterzogen werden, wodurch die Konstruktion eines neuen Initiierungsprimers und die Wiederholung des obengenannten Verfahrens ermöglicht wird, wobei der neue Initiierungsprimer als Ausgangspunkt für die Amplifikation und, auf Wunsch, Sequenzierung einer weiteren Region von Interesse verwendet wird, wobei die Verfahrensweise auf Wunsch wiederholt wird.

Figur 7 veranschaulicht, was oben gesagt wurde, indem eine schematische Darstellung der Verwendung von Vektorettebibliotheken gegeben wird. Genomische DNA wird bis zur Vollständigkeit mit der Restriktionsendonuclease EcoRI digeriert. Vektoretteabschnitte mit 4 Basenüberhängen, die zu den 4 durch EcoRI-Spaltung erzeugten Basenüberhängen komplementär sind, werden an die Enden jedes Fragments gespaltener genomischer DNA ligiert. Das nennt man eine EcoRI-Vektorettebibliothek. Ähnliche Bibliotheken werden unter Verwendung von beispielsweise etwa 15 bis 20 Restriktionsendonucleasen, die Hexanucleotidsequenzen erkennen, konstruiert. Wenn diese Bibliotheken gepoolt und der Amplifikation unterzogen werden, kann eine Reihe von PCR-Produkten gewonnen werden, oder sie können noch besser individuell und separat verwendet werden, um von jeder Bibliothek ein einzelnes PCR-Produkt zu erzeugen.

Figur 7 (a) zeigt eine hypothetische Restriktionskarte für die 3'-Seite eines interessierenden Locus mit bekannter Nucleotidsequenz (w). Bei Amplifikation von gepoolten Vektorettebiblio-

theiken unter Einsatz der Primer \underline{x} und \underline{y} werden die in (b) gezeigten PCR-Produkte Z, Z1, Z2, Z3 und Z4 produziert. Elektrophorese dieser PCR-Produkte auf einem Agarosegel würde das in (c) gezeigte Bandenmuster erzeugen, wo die PCR-Reaktionen eher separat mit jeder Vektorettbibliothek als mit einer gepoolten Vielzahl von Bibliotheken ablaufen. Das größte auf dem Gel sichtbar gemachte PCR-Produkt (in diesem Fall Z4) kann gel-gereinigt werden, und seine 3'-Ende-Nucleotidsequenz kann unter Verwendung von Oligomer \underline{y} oder von einem ineinandergeschachtelten Oligomer \underline{y}' als Sequenzierungsprimer bestimmt werden. Diese Nucleotidsequenz am 3'-Ende von Z4 könnte dann, bei weiteren "waling" Schritten unter Anwendung dieser Methode, dienen als oder verwendet werden für die Synthese eines neuen Initiierungsprimers \underline{x} .

Figur 8 zeigt die schematische Anwendung der Erfindung für die inverse PCR. Die inverse PCR wurde letztlich beschrieben bei Triglia, T. et al., *Nucleic Acids Research*, Band 16, 1988, S. 8186, und von Ochmann, H. et al., *Genetics* 120; 621-623 (November 1988).

Genomische DNA wird mit einem Restriktionsenzym geschnitten, das vorzugsweise relativ selten schneidet (z. B. durchschnittliche Fragmentgröße 10 bis 20 kb) wie XbaI, KpnI oder BamHI. Die Fragmente mit niedriger Konzentration werden zur Bildung von Ringen selbstligiert. Die Ringe werden dann mit einem Restriktionsenzym geschnitten, das vorzugsweise häufig schneidet (wie HinfI), und die Fragmente werden an eine entsprechende Vektorette ligiert, z. B. eine HinfI-Vektorette. Wenn also beispielsweise die einer bekannten Xba-Stelle benachbarte Sequenz zur Verfügung steht, kann die der nächsten Xba-Stelle benachbarte Sequenz durch Analyse des Vektorettteeinheitsproduktes ermittelt werden. Auf diese Weise könnte eine Reihe von Sprüngen erfolgen, die zusätzliche Startpunkte für das erfindungsgemäße Verfahren liefern würden. Zwar wurde oben als Beispiel eine durchschnittliche Fragmentgröße von 10 bis 20 kb angeführt, es wäre aber noch besser,

die Targetnucleinsäure so zu spalten, daß noch größere Fragmente von beispielsweise etwa 100 kb entstehen. Solche Fragmente können z. B. durch den Einsatz von Restriktionsenzymen gewonnen werden, die genomische DNA sehr selten schneiden wie NotI, Bss-HII und Sal I.

Figur 11(i) ist eine Photographie eines Agarosegels, die in den Spuren 1 und 11 den mit HaeIII gespaltenen Marker \emptyset x174, in Spur 12 den Marker \mathcal{K} /HindIII, in Spur 2 eine PCR-Kontrolle und in Spur 9 das amplifizierte Fragment zeigt, das unter Verwendung der Oligonucleotide 58 und 61 in Beispiel 1 und gemäß Darstellung in Figur 9 gewonnen wurde.

Figur 11(ii) ist der obere Abschnitt einer Photographie eines Agarosegels, der den Marker \mathcal{K} /HindIII in den Spuren 1 und 13, den mit HaeIII gespaltenen Marker \emptyset x174 in den Spuren 2 und 12, eine PCR-Kontrolle in Spur 3 und das unter Verwendung der Oligonucleotide 58 und 61 in Beispiel 1 (siehe Figur 9) gewonnene amplifizierte Fragment in Spur 10 zeigt.

Figur 11(iii) ist der untere Abschnitt einer Photographie eines Agarosegels, der den Marker \mathcal{K} /HindIII in den Spuren 1 und 11, den mit HaeIII gespaltenen Marker \emptyset x174 in Spur 6, das gemäß Beispiel 5 gewonnene amplifizierte Fragment in den Spuren 2, 3, 4 und 5 und das gemäß Beispiel 7 gewonnene amplifizierte Fragment in den Spuren 7, 8, 9 und 10 zeigt.

Die verbleibenden Figuren werden in den nachstehenden Beispielen im Detail behandelt.

Die Erfindung soll nachstehend unter Bezugnahme auf die folgenden Methoden und Beispiele näher veranschaulicht, aber nicht eingeschränkt werden, wobei die im folgenden einzeln aufgeführten Oligodesoxynucleotide verwendet wurden, wobei jede angegebene Nucleotidsequenz im herkömmlichen

5' \rightarrow 3'-Sinn zu lesen ist:-

Oligonucleotid Typ A (eine Gruppe von 14 Basen langen Oligonucleotiden mit 4 oder 5 Endbasen am 5'-Ende mit variierende Sequenz und daher geeignet zur Verwendung mit 5'-Überhängen. Die 4 Endbasen stellen den 5'-Überhang dar, während die fünfte Endbase vorhanden ist, um die Restriktionsenzymkennungsstelle zu zerstören).

Speziell wurden verwendet:

Oligonucleotid 1

CTAGGAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease Xba I oder NheI oder SpeI digeriert wurde)

Oligonucleotid 2

AATTGAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease EcoRI digeriert wurde)

Oligonucleotid 3

GATCGAAGGAGAGG

(zur Verwendung mit DNA, die mit den Restriktionsendonucleasen BamHI oder Bgl II oder Xho II oder Bcl I digeriert wurde)

Oligonucleotid 4

AGCTGAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease Hind III gespalten wurde)

Oligonucleotid 5

TCGAGAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease Sal I gespalten wurde)

Oligonucleotid 6

CGCGGAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease MluI oder Bss HII gespalten wurde)

Oligonucleotid 7

GTACGAAGGAGAGG

(zur Verwendung mit Restriktionsendonuclease-Asp718-gespaltener DNA)

Oligonucleotid 8

TCGACAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease XhoI gespalten wurde)

Oligonucleotid 9

CATGCAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease NcoI gespalten wurde)

Oligonucleotid 10

GGCCCAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease Not I oder Eag I gespalten wurde)

Oligonucleotid 11

CCGGTAAGGAGAGG

(zur Verwendung mit DNA, die mit den Restriktionsendonucleasen Bspm II oder Xma I oder Acc III gespalten wurde)

Oligonucleotid 12

CATGTAAGGAGAGG

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease

BspHI oder Nco I digeriert wurde)

Oligonucleotid 13

CTAGTAAGGAGAGG

(zur Verwendung mit DNA, die mit den Restriktionsendonucleasen
Avr II, Nhe I oder Xba I digeriert wurde)

Oligonucleotid Typ B (eine Gruppe von 42 Basen langen Oligonucleotiden, bei denen die 3'-Endbase eine variierende Sequenz hat; die Oligonucleotide eignen sich zur Verwendung mit 5'-Überhängen).

Speziell wurden verwendet:

Oligonucleotid 14

CGAATCGTAACCGTTTCGTACGAGAATCGCTGTCCTCTCCTTC

Dieses Oligonucleotid ist mit den Oligonucleotiden 1, 2, 3, 4, 5, 6 und 7 zu verwenden.

Oligonucleotid 15

CGAATCGTAACCGTTTCGTACGAGAATCGCTGTCCTCTCCTTG

(Die Nucleotidsequenz ist die gleiche wie bei 14, mit dem Unterschied, daß der 3'-terminale Rest C durch G ersetzt wird).
Dieses Oligonucleotid ist mit den Oligonucleotiden 8, 9 und 10 zu verwenden.

Oligonucleotid 16

CGAATCGTAACCGTTTCGTACGAGAATCGCTGTCCTCTCCTTT

(Die Nucleotidsequenz ist die gleiche wie bei 14, mit dem Unterschied, daß der 3'-terminale Rest C durch T ersetzt wird).

Oligonucleotid 17

CGAATCGTAACCGTTTCGTACGAGAATCGCTGTCCTCTCCTTA

(Die Nucleotidsequenz ist die gleiche wie bei 14, mit dem Unter-

schied, daß der 3'-terminale Rest C durch A ersetzt wird). Dieses Oligonucleotid ist mit dem Oligonucleotiden 11, 12, 13 und 18 (nachstehend definiert) zu verwenden.

Oligonucleotid 18 (10mer)

TAAGGAGAGG

Oligonucleotid 18 wird an Oligonucleotid 17 hybridisiert. Sie sind zur Verwendung als Vektoretteneinheit mit DNA konstruiert, die mit einer zur Schaffung stumpfer Enden fähigen Restriktionsendonuclease digeriert wurde).

Die Oligonucleotide 19 und Oligonucleotide der Serie Typ C sind zum Einsatz mit 3'-Überhängen geeignet:

Oligonucleotid 19 (10mer)

5'- AAA GGA GAG G-3'

Dieses Oligonucleotid ist zur Verwendung mit den Oligonucleotiden 20 bis 25 gemäß nachstehender Definition konstruiert.

Oligonucleotid Typ C (eine Gruppe von 46 Basen langen Oligonucleotiden)

Oligonucleotid 20

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTT
TGC A-3'

(zur Verwendung mit DNA, die mit den Restriktionsendonucleasen Nsi I oder Pst I digeriert wurde)

Oligonucleotid 21

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTT
AGC T-3'

(zur Verwendung mit DNA, die mit den Restriktionsendonucleasen Sst I oder Sac I digeriert wurde)

Oligonucleotid 22

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTT
CAT G-3'

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease
Sph I digeriert wurde)

Oligonucleotid 23

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTT
GTA C-3'

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease
Kpn I digeriert wurde)

Oligonucleotid 24

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTT
ACG T-3'

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease
Aat II digeriert wurde)

Oligonucleotid 25

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTT
GGC C-3'

(zur Verwendung mit DNA, die mit der Restriktionsendonuclease
Apa I digeriert wurde)

Oligonucleotid Typ D

Das ist eine Gruppe von 57 Basen langen Oligonucleotiden mit
nur 4 bis 5 Basen am 5'-Ende, deren Sequenz variiert.

Oligonucleotid 26

CTAGGAAGGAGAGGACGCTGTCTGTCTGAAGGTAAGGAACGGAGGAGAGAAGGGAGAG

Zur Verwendung mit Oligonucleotid 40 (nachstehend definiert)
und mit XbaI, Nhe I oder Spe I digerierter DNA.

Oligonucleotid 27

5'- AAT TGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 40 und mit EcoRI gespaltener
DNA.

Oligonucleotid 28

5'-TCGAGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 40 und mit Sal I gespaltener
DNA.

Oligonucleotid 29

5'-CGC GGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 40 und mit Bss HII oder Mlu I
gespaltener DNA.

Oligonucleotid 30

5'-AGC TGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 40 und mit Hind III gespal-
tener DNA.

Oligonucleotid 31

5'-GAT CGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 40 und mit Bam HI, Bc II oder
Bgl II gespaltener DNA.

Oligonucleotid 32

5'-CCG GGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG

Zur Verwendung mit Oligonucleotid 40 und mit Apa LI gespaltener

DNA.

Oligonucleotid 33

5'-TGC AGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 40 und mit Apa LI gespaltener
DNA.

Oligonucleotid 34

5'-TCG ATA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 41 und mit Xho I gespaltener
DNA.

Oligonucleotid 35

5'-GGC CTA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 41 und mit Eag I oder Not I
oder Xma III gespaltener DNA.

Oligonucleotid 36

5'-CCG GTA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 41 und mit Bspm II oder Xma I
oder Acc III gespaltener DNA.

Oligonucleotid 37

5'-CAT GTA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 41 und mit Bsp HI oder Nco I
gespaltener DNA.

Oligonucleotid 38

5'-CTA GTA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 41 und mit Avr II, NheI oder
Xba I gespaltener DNA.

Oligonucleotid 39

5'-TA AGG AGA GGA CGC TGT CTG TCG AAG GTA ACG AAC GGA GGA GAG
AAG GGA GAG-3'

Zur Verwendung mit Oligonucleotid 41 und DNA, die mit einer zur
Schaffung stumpfer Enden fähigen Restriktionsendonuclease ge-
spalten wurde.

Oligonucleotid Typ E

Das ist eine Gruppe von 53 Basen langen Oligonucleotiden mit
einem variierenden 3'-terminalen Rest.

Oligonucleotid 40

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TC-3'

Zur Verwendung mit den Oligonucleotiden 26 bis 33.

Oligonucleotid 41

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TA-3'

Zur Verwendung mit den Oligonucleotiden 34 bis 39.

Oligonucleotid Typ F

Hierbei handelt es sich um ein 53 Basen langes Oligonucleotid
zur Verwendung mit der Gruppe von Oligonucleotiden von Typ G
(43 bis 48) gemäß nachstehender Definition und zur Verwendung
mit 3'-Überhängen.

Oligonucleotid 42

5'-AAAGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA GAG
AAG GGA GAG-3'

Oligonucleotid Typ G

Das ist eine Gruppe von 57 Basen langen Oligonucleotiden mit 4
Basen am 3'-Ende mit variierender Sequenz.

Oligonucleotid 43

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TTT GCA-3'

Zur Verwendung mit Oligonucleotid 42 und mit Pst I oder Nsi I
gespaltener DNA.

Oligonucleotid 44

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TTA GCT-3'

Zur Verwendung mit Oligonucleotid 42 und mit Sst I oder Sac I
gespaltener DNA.

Oligonucleotid 45

5'- CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TTC ATG-3'

Zur Verwendung mit Oligonucleotid 42 und mit Sph I gespaltener
DNA.

Oligonucleotid 46

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TTG TAC-3'

Zur Verwendung mit Oligonucleotid 42 und mit Kpn I gespaltener
DNA.

Oligonucleotid 47

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TTA CGT-3'

Zur Verwendung mit Oligonucleotid 42 und mit Aat II gespaltener DNA.

Oligonucleotid 48

5'-CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA CGA GAA TCG CTG TCC
TCT CCT TTG GCC-3'

Zur Verwendung mit Oligonucleotid 42 und mit Apa I gespaltener DNA.

Die angeführten Restriktionsendonucleasen sind zum Spalten der DNA zum Einsatz mit dem speziellen Oligonucleotid oder Oligonucleotidpaar zu verwenden. So ist Oligonucleotid 1 beispielsweise mit genomischer DNA zu verwenden, die mit Nhe I oder Spe I oder Xba I gespalten wurde.

Alle Oligonucleotide zum Einsatz mit 5'-Überhängen (d. h. Typ A, D und F (42)) und die Oligonucleotide 18, 19 und 39 werden vorzugsweise phosphoryliert.

Oligonucleotide 49 - 57

Sie werden bei der Konstruktion alternativer Vektoretteneinheiten eingesetzt.

Oligonucleotid 49

5'-AAT TGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA
GGA G-3'

Oligonucleotid 50

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT GTC CTC TCC TTC-3'

Oligonucleotid 51

5'-AAT TGA AGG AGA GGA CGC TGT CAG AGG ACG GTT ACG AAC GTA
GGA CAG AAG GGA GAG-3'

Oligonucleotid 52

AAT TGA AGG AGA GGA CGC TGA CTG TCG AAC GTA CGG ATA GGA GTC GAG
AAG GGA GTC GAG AAG GGA GAG-3'

Oligonucleotid 53

5'-AAT TGA AGG AGA GGA GAG AAG GGA CGC TGT CTG TCG AAG GTA AGG
AAC GGA GGA-3'

Oligonucleotid 54

5'-CGA ATC GTA ACC GTT CGT ACG AGA ATC GCT TCC CTT CTC TCC TCT
CCI TC-3'

Oligonucleotid 55

5'-AAT TGA AGG AGA GGA CGC TGT CTG TCG AAG GTA AGG AAC GGA GGA
GAG AAG GGA GAG AAA GAG GAA GGG AAG-3'

Oligonucleotid 56

5'-CTT CCC TTC CTC TTT CTC TCC CTT CTC GAA TCG TAA CCG TTC GTA
CGA GAA TCG CTG TCC TCT CCT TC-3'

Oligonucleotid 57

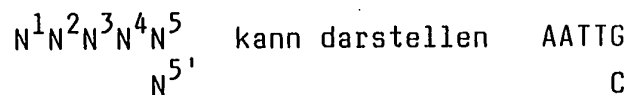
5'-AAT TGA AGG AGA GGC AGA AGG GAG AG-3'

Alternative Vektoretteneinheiten

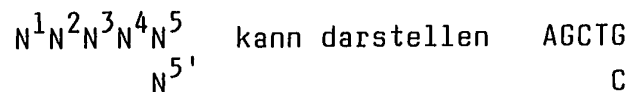
Oligonucleotid 49 sollte an Oligonucleotid 50 hybridisiert werden. Oligonucleotide 51, 52, und 57 sollten an Oligonucleotid 40 hybridisiert werden.

Oligonucleotid 53 sollte an Oligonucleotid 54 hybridisiert werden. Oligonucleotid 55 sollte an Oligonucleotid 56 hybridisiert werden. Diese Oligonucleotide würden mit genomischer DNA verwendet werden, die mit EcoR1 gespalten wurde.

In Tabelle 1 stellen N^1 , N^2 , N^3 und N^4 vier Nucleotide dar, die kohäsive Enden zur Verfügung stellen, die zur Ligation an eine entsprechende Restriktionsstelle auf einem Targetnucleinsäurefragment fähig sind. Die Nucleotide N^5 und $N^{5'}$ werden bewußt so ausgewählt, daß das Restriktionsstellen-Erkennungsmuster zerstört wird. Zum Beispiel:



in bezug auf die EcoRI-Restriktionsstelle, und



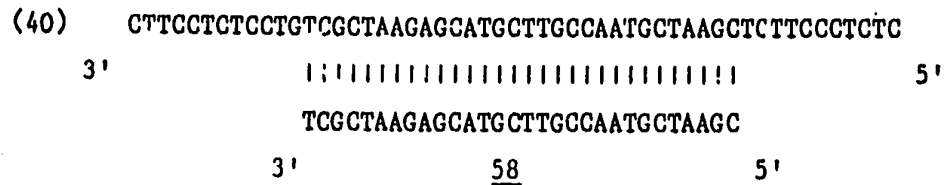
in bezug auf die Hind III-Restriktionsstelle.

Wo N^1 , N^2 , N^3 , N^4 , N^5 und $N^{5'}$ das obengenannte EcoRI-Restriktionsstellen-Erkennungsmuster darstellen, werden die spezifischen Vektoretteneinheiten 1 bis 10 nachstehend in Tabelle 2 einzeln aufgeführt, wobei jedes Oligonucleotid durch seine entsprechende Identifikationsnummer markiert wird. In den Vektoretteneinheiten 11 und 12 von Tabelle 2 stellen N^1 , N^2 , N^3 , N^4 , N^5 und $N^{5'}$ das obengenannte Hind III-Restriktionsstellen-Erkennungsmuster dar, und jedes Oligonucleotid wird durch seine entsprechende Identifikationsnummer markiert.

Die Oligonucleotide 58, 59 und 60 können mit allen Konstrukten verwendet werden.

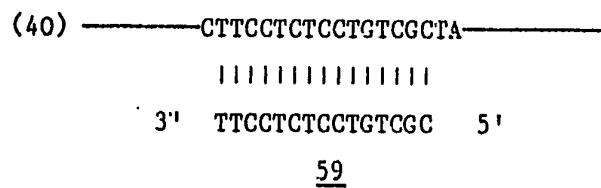
Oligonucleotid 58 ist ein universeller Vektoretteprimer.

z. B.:



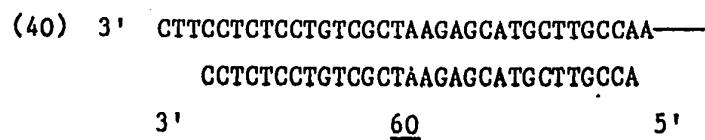
Oligonucleotid 59 ist ein universeller Sequenzierungsprimer und ein ineinandergeschachtelter Vektoretteprimer.

z. B.:



Oligonucleotid 60 ist ein ineinandergeschachtelter universeller Vektoretteprimer.

z. B.:



Kurze Oligonucleotide sind die 14mer + 42mer hybridisierten.

Lange Oligonucleotide sind die 57mer + 53mer hybridisierten.

Typ A und Typ D sind äquivalente kurze bzw. lange Oligonucleotide.

Deshalb haben 1 und 26 den gleichen Überhang und sind beide mit DNA zu verwenden, die mit Nhe I, Spe I oder Xba I gespalten wurde. Ähnlich sind 2 und 27 beide mit DNA zu verwenden, die mit EcoRI gespalten wurde.

Genau das gleiche gilt für Typ B und Typ E, 19 und Typ F (42) sowie Typ C und Typ G.

Jedes Oligonucleotid ist an DNA zu ligieren, die mit der Restriktionsendonuclease gespalten wurde, die am nächsten zu dem Oligonucleotid angegeben ist. Die Typen A und B (und Typ D und Typ E) sind zur Verwendung mit 5'-Überhängen vorgesehen.

Die Oligonucleotide 19 + Typ C (und Typ F + Typ G) sind zur Verwendung mit 3'-Überhängen vorgesehen.

Die Oligonucleotide 1 bis 8 werden mit Oligonucleotid 14 gepaart.

Die Oligonucleotide 9 bis 13 und 18 werden mit Oligonucleotid 17 gepaart.

Oligonucleotid 19 wird mit den Oligonucleotiden 20 bis 25 gepaart.

Die Oligonucleotide 26 bis 33 werden mit Oligonucleotid 40 gepaart.

Die Oligonucleotide 34 bis 39 werden mit Oligonucleotid 41 gepaart.

Typ F (42) wird mit Typ-G-Oligonucleotiden 43 bis 48 gepaart.

Oligonucleotid 7 (und 32) kann durch Oligonucleotid 11 (und 36) zur Verwendung mit Bspm II-restringierter DNA ersetzt werden.

Es versteht sich, daß die Oligonucleotide 1 bis 7 zum Hybridisieren an Oligonucleotid 14 vorhanden sind. Die Oligonucleotide 8 bis 10 sind zum Hybridisieren an Oligonucleotid 15 da. Die Oligonucleotide 11 bis 13 sind zum Hybridisieren an Oligonucleotid 17 vorhanden. Die Verwendung von Oligonucleotid 16 kann sich als unnötig erweisen.

Oligonucleotid 58

CGAATCGTAACCGTTCGTACGAGAATCGCT

(ein 30 Basen langes Oligonucleotid, das als "Universalvektoretteprimer" in der in Schritt 6 von nachstehender Methode I aufgeführten Polymerasekettenreaktion (PCR) vorgesehen ist).

In den nachstehenden Beispielen verwendete zusätzliche Oligonucleotide

Diese Oligonucleotide können mit jeder Vektorettebibliothek verwendet werden, und zwar unabhängig von der eingesetzten Restriktionsendonuclease und mit Ausnahme von Oligonucleotid 65, siehe nachstehende detaillierte Aufstellung:

Oligonucleotid 61

GAGACTGGTATTTTGTTC AATCATTAA G

(ein 3'-Oligonucleotid von Exon V des α -1 Antitrypsingens)

Oligonucleotid 62

AAAAGCCAGAGACCTCACTCCCGGGGAGCC

(ein 5'-Oligonucleotid von Exon 1 des menschlichen Phenylalaninhydroxylasegens, dessen Mutationen Phenylketonurie (PKU) verursachen).

Oligonucleotid 63

AGGGACTTACTGTGGCGAGCTTTTCAATGT

(ein 3'-Oligonucleotid von Exon 9 des Phenylalaninhydroxylasegens (PAH-Gens))

Oligonucleotid 64

GGGCCTCAGTCCCAACATGGCTAAGAGGTG

(ein 3'-Oligonucleotid von Exon III des α -1-Antitrypsingens)

Oligonucleotid 65

AATTCACACAGGAAACAGCTATGACCATG

(ein verlängerter M13-Reverse-Primer)

Dieser ist als einer der PCR-Primer für die Amplifikation von Inserts entweder in M13-DNA oder pUC-Plasmid-DNA zu verwenden.

Oligonucleotid 66

5'-ATC AGG TGC ACC CAG AGA GGC AAG GCC-3'

(ein 3'-Oligonucleotid von Exon 9 des Phenylalaninhydroxylasegens (PAH-Gens))

Methode I

Sie umfaßt die folgenden sechs Schritte:

Schritt 1 - Phosphorylierung von Oligonucleotid Typ A, anschließend Gelreinigung der phosphorylierten Species.

Schritt 2 - Addition eines ddA-Rests an das 3'-Ende des phosphorylierten Oligonucleotids Typ A von Schritt 1 mit anschließender Gelreinigung.

Schritt 3 - Hybridisierung des modifizierten Oligonucleotids Typ A an das entsprechende Oligonucleotid Typ B.

Schritt 4 - DNA-Herstellung und Restriktionsdigestion.

Schritt 5 - Ligation der hybridisierten Oligonucleotide Typ A und Typ B an die gespaltene dephosphorylierte DNA von Schritt 4.

Schritt 6 - Amplifikation der Ligationsprodukte von Schritt 5 und Nachweis und Analyse der PCR-Produkte.

Solche Schritte können zum Beispiel folgendermaßen ausgeführt werden:

Schritt 1

Folgende Puffer und andere Standardlösungen wurden verwendet:

I.1 10X Kinasepuffer

0,5 M Tris-HCl (pH 7,6)

0,1 M MgCl₂

50 mM Dithiothreitol

1 mM Spermidin

1 mM EDTA

aufbewahrt bei - 20 °C.

I.2 Adenosin-5'-triphosphat

Natriumsalz als 10-mM-Lösung (Pharmacia)

Verdünnt mit Wasser auf 20 pMol/µl,

aufbewahrt bei - 20 °C.

I.3 Radioisotope

a) Adenosin-5'-/ γ -³²P/triphosphat (Amersham)

Spezifische Aktivität etwa 6000 Ci/mmol.

b) Didesoxyadenosin-5'-/ α -³²P/triphosphat (Amersham)

Spezifische Aktivität etwa 3700 Ci/mmol.

I.4 Formamidprobenpuffer

80 % (Vol./Vol.) deionisiertes Formamid

50 mM Tris-borat, pH 8,3

1 mM EDTA

0,1 % (Masse/Vol.) Bromphen-1blau

0,1 % (Masse/Vol.) Xylencyanol

Das Formamid wurde durch 30minütiges Rühren mit 2 g Amberlit-MB-1-Harz (Bio-Rad) und Abfiltrieren unter Einsatz eines Whatman-Filters der Nummer 1 deionisiert. Der Puffer wurde im Dunkeln aufbewahrt.

I.5 40 %iges Vorrats-Acrylamid

pro 500 ml;

Acrylamid (elektrophoretische Qualität 190 g)

Bisacrylamid (elektrophoretische Qualität 10 g)

Das Acrylamid wurde 30 Minuten mit 5 g Amberlit-MB-1-Harz (Bio-Rad) gerührt und unter Verwendung von zwei Platten Whatman-Filter Nummer 1 filtriert. Die Vorratslösung wurde im Dunkeln bei 40 °C aufbewahrt.

I.6 10%iges Ammoniumpersulfat

1 g Ammoniumpersulfat (Sigma, Elektrophoresequalität) wurde gelöst und mit sterilem deionisiertem Wasser auf 10 ml aufgefüllt. Es wurde im Dunkeln aufbewahrt.

I.7 0,5X Gelmischung

(20 % Acrylamid, 7 M Harnstoff)

pro 50 ml:

21 g Harnstoff

25 ml 40 %iges Vorrats-Acrylamid (1,5)

2,5 ml 10 x TBE (1,8)

Unmittelbar vor der Verwendung sind polymerisierende Agenzien zuzusetzen:

300 µl 10 %iges Ammoniumpersulfat (1,6)

40 µl TEMED

I.8 10X TBE pro Liter:

Tris 109 g

EDTA 9,5 g

Borsäure 55 g

10X TBE-Puffer wurde, wenn erforderlich, in Wasser auf 1- oder 0,5fache Stärke verdünnt.

I.9 Puffer A

0,1 M Tris-HCl, pH 7,7

10 mM Triethylamin

1 mM EDTA

Der Puffer wurde bei 4 °C aufbewahrt und einmal im Monat frisch hergestellt.

I.10 Reagens B pro 100 ml:

50 ml Ethanol

50 ml steriles deionisiertes Wasser

Phosphorylierung des 14mer Oligonucleotids Typ A

Folgendes wurde in einem 1,5 ml Sarstedt-Röhrchen gemischt:

Oligonucleotid Typ A	100 pmol
10X Kinasepuffer (1,1)	5 µl
Adenosin-5'-triphosphat (1,2)	9,5 µl (190 pmol)
Adenosin-5'- γ ³² P/triphosphat (1.3)	5 µl (10 pmol)
T4-Polynucleotidkinase E.coli B (Amersham)	40 Einheiten
H ₂ O	auf 50 µl.

Das Verhältnis von Oligonucleotiden zu ATP-Molekülen wurde bei 1:2 gehalten.

Die Phosphorylierung der Oligonucleotide erfolgte durch 60 Minuten lange Inkubation bei 37 °C.

20 µl des Formamidprobenpuffers (1.4) wurden zugesetzt, und die Enzymprobe wurde 10 Minuten in einem 90 °C-Wasserbad denaturiert, bevor sie auf ein 20 % Acrylamid/7M Harnstoff-Denaturierungsgel gebracht wurde.

Es erfolgte Gelreinigung der kinasierten (kinased) Oligonucleotide. Phosphorylierte Oligonucleotide wurden durch Elektrophorese auf denaturierenden Polyacrylamidgelen gemäß der Beschreibung durch Wu R. et al. "Purification and sequence analysis of synthetic oligonucleotides" (Reinigung und Sequenzanalyse von synthetischen Oligonucleotiden) in Oligonucleotide Synthesis - a practical approach, herausgegeben von M. J. Gait, IRL Press, gereinigt.

c) Endreinigung durch eine Säule

NeNsorbTM-(DuPont)-Nucleinsäure-Reinigungspatronen wurden für diesen Endreinigungsschritt eingesetzt.

1) Prä-äquilibrierung der Säule

Die Säule wurde an einem Klemmständer befestigt. Das Innere der Patrone wurde mit 2 ml Methanol von HPLC-Qualität gewaschen. Dadurch wurde jegliches lose Packungsmaterial auf das Säulenbett zurückgespült.

Der Adapter wurde fest in das obere Ende der Patrone gepreßt, so daß ein luftdichter Verschuß entstand. Eine Einwegplast-spritze wurde mit Luft gefüllt und am Adapter befestigt. Mit konstantem sanftem Druck wurde das gesamte Methanol durch die Patrone gedrückt (Durchflußgeschwindigkeit: 1 Tropfen in 2 Sekunden). Auf die gleiche Weise wurde die Säule 'gestartet', indem 2 ml Puffer A (1.9) hindurchgeleitet wurden.

2) Probenaufgabe

2,8 µl Triethylamin wurden den 2 ml Wasser zugesetzt, die das Oligonucleotid enthielten. Das Ganze wurde dann unter Verwendung einer Spritze und einer Nadel direkt oben auf das Säulenbett gegeben. Die Probe wurde unter Einsatz einer mit Luft gefüllten Spritze, wie in C(1), durch die Säule geleitet.

3) Probenwäsche

3 ml Puffer A (1.9) wurden unter Einsatz einer Spritze und Nadel auf die Säule gegeben und wie in C(1) durch die Säule gedrückt.

4) 1 ml Reagens B(1.10) wurde oben auf die Säule gegeben. Die DNA wurde durch sanften Druck unter Verwendung einer mit Luft gefüllten Spritze, wie in C(1), von dem Säulenbett eluiert. Das ablaufende Medium wurde in einem 1,5-ml-Sarstedt-Röhrchen gesammelt und in einer Zentrifuge, an die ein Vakuum angelegt war (UNIVAP, Uniscience) zur Trocknung eingedampft. Das getrocknete, gereinigte Oligonucleotid wurde, wenn erforderlich, resuspendiert.

Schritt 2

Folgende Puffer und andere Standardlösungen wurden verwendet:

- 2.1 1,4 M Natriumcacodylat
 pH 7,5 mit HCl
 Aufbewahrung bei - 20 °C.

- 2.2 5 mM Cobaltdichlorid
Aufbewahrung bei - 20 °C.
- 2.3 1 mM Dithiothreitol
Aufbewahrung bei - 20 ° C.
- 2.4 Didesoxyadenosin-5'-triphosphat-lyophilisiertes
Natriumsalz.
(Boehringer Mannheim)
Verdünnt auf 10 mM Vorratslösung in H₂O
und 1 mM Arbeitslösung in H₂O
- 2.5 0,5 M EDTA, pH 8,0

Addition eines Didesoxyadenosinrests am 3'-Ende des Oligonucleotids
Typ A

Oligonucleotid Typ A wurde phosphoryliert, und die phosphorylierte Species wurde, wie in Schritt 1 beschrieben, gereinigt. Mit einer Rückgewinnung von 80 % werden 80 pmol Oligonucleotid kinasiert und stehen für die Terminale-Transferase-Reaktion zur Verfügung.

Die 80 pmol kinasiertes und gereinigtes Oligonucleotid I wurden in 48 µl sterilem deionisiertem Wasser resuspendiert.

Folgendes wurde zu dem kinasierten Oligonucleotid hinzugegeben:

- (48 µl Oligonucleotid (80 pmol)
- 1 µl 1 mM ddATP (1 nmol) (2.4)
- 10 µl 1,4 M Cacodylsäure (2.1)
- 10 µl 1 mM DTT (2.3)
- 20 µl 5 mM Cobaltdichlorid (2.2)
- 5 µl ddATP (³²P) (15 pmol) (1.3b)
- 6 µl Terminale Transferase
(150 Einheiten Boehringer TdT)

Das Reaktionsgemisch wurde 1 Stunde bei 31 °C inkubiert, und die Reaktion wurde dann durch den Zusatz von 4 µl 0,5 M EDTA beendet.

40 µl des Formamidprobenpuffers wurden zugesetzt (1.4), und die Probe wurde durch 10minütiges Inkubieren in einem Wasserbad von 90 °C vor dem Aufgeben auf ein 20 % Acrylamid/7 M Harnstoff-Denaturierungsgel denaturiert.

Die Gelelektrophorese entsprach genau der Beschreibung in Schritt 1.

Die nunmehr 15mer-Oligonucleotid-Typ A-Bande wurde ausgeschnitten und genau entsprechend der Beschreibung in Schritt 1 gereinigt. Die endgültige Rückgewinnung wird auf 50 pmol geschätzt. Diese werden in 250 µl sterilem deionisiertem Wasser resuspendiert, so daß sich eine Konzentration von 0,2 pmol/µl ergibt.

Schritt 3

Hybridisierung (annealing) des modifizierten Oligonucleotids Typ A an das entsprechende Oligonucleotid Typ B

Die folgende Lösung war erforderlich.

3.1 10 x Hybridisierungspuffer
 100 mM Tris-HCl (pH 8,0)
 100 mM MgCl₂

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen mit Schraubkappe gemischt:

25 pmol modifiziertes Oligonucleotid vom Typ A (gewonnen über Schritt 1 und 2)

25 pmol Oligonucleotid Typ B
5 µl 10 x Hybridisierungspuffer (3.1)
und H₂O auf 50 µl.

Wenn das Oligonucleotid von Typ A unter den Oligonucleotiden 1, 2, 3 oder 4 ausgewählt wird, dann ist ein geeignetes Oligonucleotid vom Typ B dafür (1-4) das Oligonucleotid 14.

Das Eppendorf-Röhrchen wurde in ein siedendes Wasserbad (passenderweise ein kleines Becherglas) gelegt. Nach 5 Minuten ließ man das Wasser in dem Becherglas langsam über einen Zeitraum von 2 Stunden auf Raumtemperatur abkühlen. Schließlich wurde der Inhalt

des Röhrchens durch 10 Sekunden langes Zentrifugieren am Boden gesammelt. Das hybridisierte Oligonucleotid, mit einer Konzentration von 0,5 pmol/ μ l ist fertig zum Einsatz bei Ligationen.

Schritt 4

DNA-Herstellung und Restriktion

Folgende DNA-Proben wurden verwendet:

1. Menschliche genomische DNA I
2. Menschliche genomische DNA II (Genomische DNA, hergestellt aus Lymphoblastoidzellen), bei der dem Kulturmedium 0,1 mg/ml 5-Azacytidin zugesetzt worden war, um Methylierung an den Cytidinresten während Replikation von DNA und Zellwachstum zu verhindern.

Folgende Puffer wurden verwendet:

4.1 10x salzarmer Puffer
100 mM Tris-HCl, pH 7,5
100 mM MgCl₂
10 mM Dithioerythritol

4.2 10x Puffer mit mittlerem Salzgehalt
100 mM Tris-HCl, pH 7,5
100 mM MgCl₂
500 mM NaCl
10 mM Dithioerythritol

4.3 10x salzreicher Puffer
500 mM Tris-HCl, pH 7,5
100 mM MgCl₂
1000 mM NaCl
10 mM Dithioerythritol

4.4 10x BamHI-Puffer
100 mM Tris-HCl, pH 8,0
50 mM MgCl₂
1000 mM NaCl
10 mM 2-Mercaptoethanol.

4.5 10x EcoRI-Puffer
100 mM Tris-HCl, pH 7,5
100 mM MgCl₂
500 mM NaCl

4.6 TE-Puffer
10 mM Tris-HCl, pH 7,5
1 mM EDTA

4.7 Agaroseprobenpuffer
15 % (Masse/Vol.) Ficoll 400
0,05 % (Masse/Vol.) Bromphenolblau
0,05 % (Masse/Vol.) Xylencyanol

Der Probenpuffer wurde in 1X TBE-Puffer (1.8) aufgefüllt.

4.8 10X CIP-Puffer
500 mM Tris-HCl, pH 8,0
1 mM EDTA

(i) Die folgenden Bedingungen wurden zum Spalten der genomischen DNA mit spezifischen Restriktionsendonucleasen angewendet:

50 µg DNA

20 µl des geeigneten 10X Restriktionspuffers.

50 Einheiten der gewünschten Restriktionsendonuclease.

Das Endvolumen des Reaktionsgemisches wurde auf 200µl eingestellt.

Digestion erfolgte 2 Stunden bei 37 °C. Die Reaktion wurde durch den Zusatz von 4 µl 0,5 M EDTA (2.5) und Erwärmen der Probe über einen Zeitraum von 10 Minuten auf 70 °C beendet.

(ii) 1 µg der Probe wurde durch Agarosegelelektrophorese kontrolliert, um komplette Digestion sicherzustellen.

Agarosegelelektrophorese

Die Gele wurden aus 0,7 % (Masse/Vol.) Agarose (Pharmacia) in 1X TBE-Puffer (1.8) hergestellt und zur Elektrophorese in TBE-Puffer der gleichen Stärke getaucht. Ein Fünftel des Agarosegel-

puffervolumens (4.7) wurde zu der Probe gegeben, bevor sie in die Vertiefungen des Gels gegeben wurde. Nach der Elektrophorese wurde das Gel 10 Minuten lang durch Eintauchen in deionisiertes Wasser mit einem Gehalt von 1 µg/ml Ethidiumbromid angefärbt. Die DNA wurde unter Einsatz eines Transilluminators (Macrovue-LKB) sichtbar gemacht.

(iii) (a) Die digerierte DNA wurde durch Phenolextraktion gereinigt:

Ein gleiches Volumen gepufferten redestillierten Phenols (BRL, Phenol mit Nucleinsäurequalität, mit TE gepuffert (4.6)) wurde zu der restringierten DNA hinzugegeben, und das Gemisch wurde 30 Sekunden wirbelbehandelt. Die entstehende weiße Emulsion wurde durch 5minütiges Zentrifugieren in zwei Phasen getrennt. Die wäßrige Phase wurde in ein frisches Röhrchen übertragen. Alles restliche Phenol wurde, nachdem die wäßrige Schicht zuerst mit Phenol : Chloroform : Isoamylalcohol (25 : 24 : 1) - ein gleiches Volumen - extrahiert worden war, mit einem gleichen Volumen von Chloroform : Isoamylalcohol (24 : 1) entfernt.

(b) Die DNA wurde schließlich durch Ethanolpräzipitation gereinigt: 1/10 Volumen von 3M Natriumacetat wurde zu der DNA-Lösung hinzugegeben, um eine Endkonzentration von 300 mM Natriumacetat zu erhalten. Die DNA wurde durch den Zusatz von 2,5 Volumen 95 %iges Ethanol und anschließendes 30minütiges Stehen bei - 80 °C und nachfolgende Zentrifugierung (Mikrozentrifuge, 5 Minuten lang, 12 000 U/min) präzipitiert. Die pelletierte DNA wurde mit 80 % (Vol./Vol.) kaltem Ethanol gewaschen und unter Vakuum getrocknet. Die DNA wurde bei Anforderung in Puffer gelöst.

(iv) Dephosphorylierung der gespaltenen DNA

Die gespaltene DNA wurde zur Verhinderung von Selbstligation dephosphoryliert.

Die restringierte und gereinigte DNA wurde in 175 µl sterilem deionisiertem Wasser gelöst.

Zu dieser DNA wurden hinzugegeben:

5 μ l Kalbsdarmphosphatase (CIP)
 (Boehringer Mannheim) (1 Einheit/ μ l)
 20 μ l 10X CIP-Puffer (4.8)

Die Probe wurde 60 Minuten lang bei 37 °C inkubiert. 4 μ l 0,5 M EDTA wurden zugesetzt, und die dephosphorylierte DNA wurde durch Phenolextraktion und Ethanolpräzipitation gemäß Beschreibung in (iii) (a) und (iii) (b) oben gereinigt. Das fertige DNA-Pellet wurde in 50 μ l TE-Puffer (4.6) gelöst, so daß eine Konzentration von 1 μ g/ μ l entstand.

Schritt 5

Ligation der hybridisierten (annealed) Oligonucleotide I & II an die gespaltene dephosphorylierte DNA

Folgende Puffer wurden verwendet:

5.1 10X Ligationspuffer
 200 mM Tris-HCl, pH 7,4
 100 mM MgCl₂
 100 mM DTT
 10 mM ATP

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen gemischt:

Gespaltene genomische DNA (7,4 pmol)	10 μ g
Hybridisierte Oligonucleotide von Schritt 3	11 pmol
10X Ligationspuffer	10 μ l
T4-DNA-Ligase (Boehringer Mannheim)	2 Einheiten
Sterilisiertes deionisiertes Wasser, um das Volumen auf 100 μ l zu bringen	

Das Ligationsgemisch wurde über Nacht bei 4 °C stehen gelassen. Die Effizienz der Ligation wurde durch Gelelektrophorese von 5 % der Ligationsprodukte eingeschätzt - siehe Schritt 4 (ii) bezüglich Einzelheiten der Verfahrensweise.

Die Ligationsprodukte wurden dann durch Phenolextraktion und Ethanolpräzipitation gemäß der Beschreibung in Schritt 4 (iii) gereinigt. Das fertige DNA-Pellet wurde in 9,5 µl TE-Puffer (4.6) gelöst, um eine Konzentration von 1 µg/µl zu erhalten.

Schritt 6

Amplifikation der Ligationsprodukte von Schritt 5

Folgende Puffer und andere Standardlösungen wurden verwendet.

6.1 10X Amplifikationspuffer

500 mM KCl
100 mM Tris-HCl, pH 8,3
15 mM MgCl₂
0,1 % Gelatine
Aufbewahrung bei - 20 °C

6.2 Nucleotide

10X Vorrat
2 mM dATP
2 mM dTTP
2 mM dGTP
2 mM dCTP
Aufbewahrung bei - 20 °C

6.3 Größenmarker

Mit HaeIII gespaltener Bakteriophage ØX174.

Die Größe der Bande in Basenpaaren ist:

1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118 und 72.

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen gemischt:

1 µg der Ligationsprodukte von Schritt 5.
100 pmol Oligonucleotid 58
100 pmol Oligonucleotid 61 oder 62 oder 63
10 µl 10X Amplifikationspuffer (6.1)
10 µl 10X Vorratsnucleotide.

Das Volumen wurde mit destilliertem deionisiertem Wasser auf 100 μ l aufgefüllt.

Das Röhrchen wurde für 5 Minuten in ein siedendes Wasserbad gelegt. Taq-Polymerase (Anglian) wurde in 1X Amplifikationspuffer (6.1) auf 1 Einheit/ μ l verdünnt.

2 μ l des verdünnten Enzyms wurden zu dem obigen gesiedeten Reaktionsgemisch hinzugegeben und gemischt. Der Inhalt wurde in einer Mikrozentrifuge 5 Sekunden in Richtung Boden geschleudert. 50 μ l leichtes Mineralöl (Sigma) wurde anschließend auf das Reaktionsgemisch gegeben (um während der folgenden Reaktionen Verdampfung zu verhindern).

Das Röhrchen wurde in ein Wasserbad gelegt von:

- 1) 60 $^{\circ}$ C für 2 Minuten
- 2) 72 $^{\circ}$ C für 3 Minuten
- 3) 91 $^{\circ}$ C für 2 Minuten

Die Schritte 1), 2) und 3) wurden dann 39mal wiederholt. Das Röhrchen wurde danach 2 Minuten lang in das Wasserbad von 60 $^{\circ}$ C gelegt, woran sich Inkubation bei 72 $^{\circ}$ C über einen Zeitraum von 8 Minuten anschloß. Am Ende der Amplifikation wurde der Inhalt Richtung Boden geschleudert und das Mineralöl abpipettiert.

15 μ l der amplifizierten Produkte wurden durch Agarosegelelektrophorese entsprechend der Beschreibung in Schritt 4 (ii) analysiert. Zur Einschätzung der Größe der amplifizierten Produkte wurde mit Hae III (6.3) gespaltener \emptyset X174 neben der amplifizierten Probe elektrophoretisiert.

Methode II

Methode II umfaßt folgende Schritte:

1. Addition eines ddA-Rests an das 3'-Ende von Oligonucleotid Typ A mit anschließender Gelreinigung der 3'-end-markierten Species.

2. Phosphorylierung des 3'-end-markierten Oligonucleotids Typ A mit anschließender Separation von nichteingebauten Nucleotiden.
3. Hybridisierung des modifizierten Oligonucleotids Typ A an das entsprechende Oligonucleotid Typ B.
4. Spaltung der genomischen DNA.
5. Ligation der hybridisierten Oligonucleotide A + B an die gespaltene DNA.
6. Amplifikation der Ligationsprodukte. Nachweis und Analyse der PCR-Produkte.

Die obigen Schritte können zum Beispiel wie folgt ausgeführt werden:

Schritt I

Die experimentellen Details sind genau die gleichen wie bei Schritt 2 von Methode I.

Schritt 2

Das Oligonucleotid wird gemäß der Beschreibung in Schritt I von Methode I phosphoryliert. Das phosphorylierte Oligonucleotid wird direkt durch eine Säule gereinigt, wie es in Schritt I (ii)C von Methode I im einzelnen beschrieben wurde. Bevor die Probe auf die Säule gebracht wird, werden dem Reaktionsgemisch anstelle von Triethylamin 200 µl Puffer A (1.9) zugesetzt.

Schritte 3 - 6 sind genau die gleichen wie bei Methode I.

Methode III

Sie umfaßt die folgenden Schritte:

1. Phosphorylierung der 14mer Oligonucleotids Typ A mit anschlie-

- gender Gelreinigung der phosphorylierten Species.
2. Addition eines ddA-Rests an das 3'-Ende des kinasierten Oligonucleotids mit anschließender Gelreinigung.
 3. Hybridisierung des Oligonucleotids Typ A an das entsprechende Oligonucleotid Typ B.
 4. Restriktion der DNA.
 5. Ligation der hybridisierten Oligonucleotide A + B an die restringierte DNA.
 6. Addition eines Didesoxynucleotids an freie 3'-Enden.
 7. Amplifikation der Ligationsprodukte. Nachweis und Analyse der Amplifikationsprodukte.

Solche Schritte können zum Beispiel wie folgt ausgeführt werden:
Die Schritte 1 bis 5 entsprechen genau den Details in Methode I

Schritt 6

Die gereinigten Ligationsprodukte von Schritt 5 werden in 30 µl destilliertem deionisiertem Wasser gelöst. Zu der DNA wird folgendes hinzugegeben:

- 5 µl 10x Amplifikationspuffer (6.1)
- 15 pmol des geeigneten Didesoxynucleotids
- 1 Einheit Taq-Polymerase (Anglian)

Wasserzusatz, um das Volumen auf 50 µl zu bringen.

Das Reaktionsgemisch wird 30 Minuten bei 37 °C inkubiert. Die DNA wird durch Phenolextraktion und Ethanolpräzipitation gemäß der Beschreibung in Methode I, Schritt 4 (iii), gereinigt. Das DNA-Pellet wird in 50 µl TE-Puffer (4.6) gelöst, um eine

Konzentration von 1 µg/µl zu erhalten.

Schritt 7

entspricht Schritt 6 von Methode I.

Methode IV

Genau wie Methode II, mit dem Unterschied, daß Schritt 6 von Methode III zwischen den Schritten 5 und 6 von Methode II eingeschoben wird.

Methode V

Methode V umfaßt die folgenden Schritte:

1. Phosphorylierung des 57mer Oligonucleotids vom Typ D mit anschließender Separation desselben von nichteingebautem Nucleotid.
2. Hybridisierung des phosphorylierten, zum Beispiel kinasierten Oligonucleotids vom Typ D, an das entsprechende Oligonucleotid vom Typ E. Es folgen die Schritte 4 ff. von den Methoden I-IV. Solche Schritte können beispielsweise wie folgt ausgeführt werden:
 1. Das Oligonucleotid wird phosphoryliert, wie es in Methode I, Schritt I, beschrieben ist. Die Reinigung erfolgt durch eine NeNsorbTM-Säule gemäß der Beschreibung in Methode I, Schritt I (ii)C.
 2. Die Hybridisierung erfolgt, wie im einzelnen in Methode I aufgeführt. Die Schritte von 4 an aufwärts erfolgen wie vorstehend aufgeführt (Methoden I und III).

Methode VI

Methode VI umfaßt die folgenden Schritte:

1. Phosphorylierung des 14mer Oligonucleotids vom Typ A (oder Oligonucleotide 18 oder 19 oder 39 oder Typ F (42) oder Typ D)

mit anschließender Gelreinigung der phosphorylierten Species.

2. Hybridisierung des phosphorylierten Oligonucleotids z. B. 18, 19 oder Typ D an das entsprechende geeignete Oligonucleotid Typ B (oder Typ C) oder Typ E oder Typ C.

3. Restriktionsenzymdigestion der genomischen DNA.

4. Ligation der hybridisierten Oligonucleotide an die restringierte DNA.

5. Restriktion der Ligationsprodukte vom obigen Schritt.

6. Ligation weiterer hybridisierter Oligonucleotide an Produkte von Schritt (5).

7. Wiederholung von Schritt (5) und (6).

8. Endrestriktion von Produkten von Schritt (7).

9. Addition eines Didesoxynucleotidtriphosphats an das 3'-Ende der Produkte von Schritt (8).

10. Amplifikation der Produkte von Schritt (9). Nachweis und Analyse der amplifizierten Produkte.

Solche Schritte können zum Beispiel wie folgt ausgeführt werden:

Schritt 1

Wie in Schritt 1 von Methode I beschrieben.

Schritt 2

Dieser Schritt entspricht der Beschreibung in Schritt 3 von Methode I.

Zum Beispiel:

Die Oligonucleotide 1 bis 7 werden an 14 hybridisiert.

Die Oligonucleotide 11 bis 13 und 18 werden an 17 hybridisiert.
 Die Oligonucleotide 20 bis 25 werden an 19 hybridisiert.
 Die Oligonucleotide 26 bis 32 werden an 40 hybridisiert.
 Die Oligonucleotide 24 bis 39 werden an 41 hybridisiert.
 Die Oligonucleotide 43 bis 48 werden an 42 hybridisiert.

Schritt 3

Dieser Schritt ist der gleiche wie Schritt 4(i), 4(ii) und 4(iii) von Methode I, aber restringierte DNA wird nicht dephosphoryliert.

Schritt 4

Ligation der hybridisierten Oligonucleotide an die gespaltenen DNA-Puffer und andere Standardlösungen:

4.1. 10x Ligationspuffer
 200 mM Tris-HCl, pH 7,4
 100 mM MgCl₂
 100 mM DTT
 10 mM ATP

(i) Die Berechnung von pmol klebriger Enden in 1 µg gespaltenener DNA entspricht den Ausführungen in Schritt 5 von Methode I.

(ii) Ligationen

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen gemischt:

Gespaltene genomische DNA,	7,4 pmol (10 µg)
Hybridisierte Oligonucleotide von Schritt 2	11 pmol
10x Ligationspuffer (4.1)	10 µl
T4-DNA-Ligase (Boehringer Mannheim)	2 Einheiten
Sterilisiertes deionisiertes Wasser, um das Volumen zu bringen auf	100 µl

Das Ligationsgemisch wurde über Nacht bei 4 °C stehen gelassen. Die Effizienz der Ligation wurde durch Gelelektrophorese von 5% der Ligationsprodukte eingeschätzt - siehe Schritt 4(ii) von Methode I bezüglich Verfahrensdetails.

Die Ligationsprodukte wurden durch Phenolextraktion und Ethanolpräzipitation gemäß der Beschreibung in Abschnitt 4(iii) oder Methode I gereinigt.

Schritt 5

Dieser Schritt ist der gleiche wie Schritt 3 oben.

Schritt 6

Dieser Schritt ist der gleiche wie Schritt 4 oben, mit dem Unterschied, daß äquimolare Konzentrationen von DNA und hybridisierten Oligonucleotiden verwendet werden.

Zum Beispiel:

7,4 pmol DNA

7,4 pmol hybridisierte Oligonucleotide

Schritt 7

Wiederholung von Schritt 5 und 6 oben.

Schritt 8

Wie Schritt 5 oben.

Schritt 9Puffer und andere Standardlösungen9.1 T7-DNA-Polymerase-Puffer (5x)

200 mM Tris-HCl, pH 7,5

100 mM MgCl₂

250 mM NaCl

9.2 dNTP-dATP

1 mM dTTP

1 mM dGTP

1 mM dCTP

aufgefüllt mit sterilisiertem
deionisiertem Wasser.

9.3 ddATP

1 mM ddATP

aufgefüllt mit sterilisiertem
deionisiertem Wasser.

(i) Addition eines ddA-Rests an 3'-Enden von Produkten von Schritt 8:

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen gemischt:

- 4 µg DNA von Schritt oben
- 20 µl von 5x T7-DNA-Polymerase-Puffer (9.1)
- 20 µl dNTP-dATP (9.2)
- 20 µl ddATP (9.3)
- 5 Einheiten T7-DNA-Polymerase (Sequenase)

Das Volumen wurde mit destilliertem deionisiertem Wasser auf 100 µl aufgefüllt.

Das Röhrchen wurde 30 Minuten bei 37 °C inkubiert. Das Enzym wurde dann durch 10 Minuten lange Inkubation bei 70 °C denaturiert. Die DNA wurde dann durch Phenolextraktion und Ethanolpräzipitation gemäß vorstehender Beschreibung gereinigt.

Schritt 10

Dieser Schritt ist der gleiche wie Schritt 6 von Methode I, oder die Amplifikation kann auf einer Temperatur-Kreislauf-Maschine (temperature cycling machine), z. B. Techne Programmable Dri-Block PHC-1, durchgeführt werden.

Es versteht sich, daß bei dieser Methode VI Didesoxyadenosin-5'-triphosphat mit den 3'-Enden der ligierten Oligonucleotide in Anwesenheit von T7-DNA-Polymerase (Sequenase) umgesetzt wird, um jeglichen Verlust von 3'-terminalen ddA-Resten als Folge einer unerwarteten Exonucleaseaktivität der für Didesoxynucleotide spezifischen T4-DNA-Ligase zu ersetzen.

Methode VII

Methode VII ist die gleiche wie Methode VI, mit dem Unterschied, daß Schritt 1 von Methode VI weggelassen wird. Außerdem erfolgt bei dieser Methode keine Gelreinigung, aber die Effizienz der Ligation kann durch Agarosegelelektrophorese und Anfärben mit Ethidiumbromid kontrolliert werden.

Methode VIII

Methode VIII ist die gleiche wie Methode VII, mit dem Unterschied, daß Schritt 9 ausgelassen wird. Die bei dieser Methode verwendeten Oligonucleotide sind z. B. wie folgt:

26 bis 32 hybridisiert an 40

36 bis 39 hybridisiert an 41

43 bis 48 hybridisiert an 42

Methode IX

Methode IX umfaßt die folgenden Schritte:

1. Phosphorylierung
2. Hybridisierung
3. Restriktion von DNA
4. Ligation (zum Beispiel bei 25 °C) von hybridisierten Oligonucleotiden an gespaltene DNA
5. Restriktion/Ligation
6. Amplifikation der Produkte von Schritt 5 und Nachweis und Analyse der Amplifikationsprodukte.

Die obigen Schritte können beispielsweise so ausgeführt werden, wie es im Zusammenhang mit den entsprechenden Schritten in Methode VI beschrieben wurde.

Methode X

Die Schritte 1, 2 und 3 Methode X sind die gleichen wie die Schritte 1, 2 und 3 von Methode VIII, Schritt 4(i) von Methode X ist ebenfalls der gleiche wie Schritt 4(i) von Methode VIII, aber Schritt 4(ii) wird wie folgt ausgeführt:

4(ii) Ligation

Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

Restringierte 7387-DNA	0,74 pmol (1µg)
Hybridisierte (annealed) Oligonucleotide	1,65 pmol
10x Ligationspuffer (4.1)	1 µl

109

10x ATP (4.2)	1 μ l
T4-DNA-Ligase (Boehringer Mannheim)	1 Einheit
H ₂ O auf	10 μ l

Das Ligationsgemisch wurde 120 Minuten bei 25 °C (im Techne Dri-Block) belassen.

Das Ligationsgemisch wurde dann in Wasser auf eine DNA-Konzentration von 10 ng/ μ l verdünnt und der Amplifikationsreaktion gemäß der Beschreibung in Schritt 10 von vorstehender Methode VIII unterzogen.

Methode XI

Methode XI entspricht mit folgenden Ausnahmen der Methode X: Nach der 120minütigen Ligationsreaktion bei 25 °C werden die Ligationsprodukte mit der Restriktionsendonuclease digeriert, und zwar durch Zusatz von:

2 μ l Puffer für die entsprechende Restriktionsendonuclease (siehe 3.1)

2 Einheiten Restriktionsnuclease

H₂O auf 20 μ l

und Inkubation bei 37 °C über einen Zeitraum von 60 Minuten.

Die Probe wurde dann in H₂O auf eine DNA-Konzentration von 10 ng/ μ l verdünnt.

Methode XII

Methode XII ist die gleiche wie Methode X, mit dem Unterschied, daß in Schritt 4 (ii) das Ligationsgemisch über Nacht bei 4 °C stehen gelassen wird anstatt 2 Stunden bei 25 °C.

Schließlich wird das Ligationsgemisch in Wasser auf eine DNA-Konzentration von 10 ng/ μ l verdünnt.

Methode XIII

Methode XIII ist die gleiche wie Methode XII, mit dem Unterschied, daß nach der Ligationsreaktion über Nacht bei 4 °C die Ligationsprodukte mit der entsprechenden Restriktionsendonuclease digeriert werden. Das geschieht durch:

Zusatz von
2 μ l Puffer für die entsprechende Endonuclease (siehe 3.1)
2 Einheiten Restriktionsendonuclease (Boehringer Mannheim)
zu dem Ligationsgemisch von 10 μ l
und H_2O auf 20 μ l
sowie Inkubation bei 37 °C über einen Zeitraum von 60 Minuten.

Die Probe wird dann in H_2O auf eine DNA-Konzentration von
10 ng/ μ l verdünnt.

Beispiel 1

Es wurden die Oligonucleotide 1 und 14 nach der Beschreibung
in den Schritten 1 und 2 der obigen Methode I hergestellt. Die-
se Oligonucleotide zum Einsatz bei Methode I wurden nach der
Beschreibung im obigen Schritt 3 hybridisiert (annealed).
Menschliche genomische DNA II wurde hergestellt und entspre-
chend der Beschreibung in Schritt 4 von Methode I gespalten.
Mit Restriktionsendonuclease Xba 1 gespaltene und dephosphory-
lierte (wie im obigen Schritt 4 im einzelnen ausgeführt) menschl-
iche genomische DNA II wurde mit den hybridisierten Oligonucleo-
tiden 1 und 14 ligiert, wie es oben in Schritt (5) beschrie-
ben wurde.

Dann erfolgte die Amplifikation an den Ligationsprodukten.
Figur 11(i), gemäß vorstehender Beschreibung, zeigt den mit Hae
III gespaltenen Marker λ X174 in den Spuren 1 und 11, den Marker
 λ Hind III in Spur 12, eine Amplifikationskontrolle in Spur 2
und das mit den Oligonucleotiden 58 und 61 gewonnene amplifi-
zierte Fragment, das erwartungsgemäß eine Länge von 800 bp auf-
weist.

Figur 11(ii) zeigt auch den Marker λ Hind III in den Spuren 1
und 13, den mit Hae III gespaltenen Marker λ X174 in den Spuren
2 und 12, eine Amplifikationskontrolle in Spur 3 und das unter
Verwendung der Oligonucleotide 58 und 61 gewonnene amplifizierte
Fragment in Spur 10.

Beispiel 2

Verwendung der Oligonucleotide 1 und 14 mit DNA, die mit der Restriktionsendonuclease Xba I gespalten wurde.

a) Oligonucleotid 1 (100 pmol) wurde phosphoryliert, und die kinasierte Species wurde gereinigt, wie es in Schritt 1 von Methode I beschrieben ist.

b) Das kinasierte Oligonucleotid 1 wurde mit terminaler Transferase behandelt (terminal transferased), um einen Dideoxyadenosinrest an das 3'-Ende des Oligonucleotids zu fügen. Das entstehende 15mer Oligonucleotid 1 (+ ddA am 3'-Ende) wurde gereinigt, wie es in Schritt 2 von Methode I im einzelnen ausgeführt ist.

c) Oligonucleotid 1 wurde an Oligonucleotid 14 hybridisiert.

d) Menschliche genomische DNA wurde mit der Restriktionsendonuclease Xba I gespalten und dephosphoryliert, wie es in Methode I im einzelnen ausgeführt wurde.

e) Die hybridisierten (annealed) Oligonucleotide 1 und 14 wurden an mit Xba I restringierte menschliche genomische DNA II ligiert - siehe Methode I bezüglich Details.

f) Die Amplifikation wurde an den Ligationsprodukten unter Einsatz der Oligonucleotide 58 und 61 ausgeführt. Das erwartete amplifizierte Produkt weist in diesem Fall ungefähr 800 bp auf.

Beispiel 3

Verwendung der Oligonucleotide 2 und 14 mit DNA, die mit der Restriktionsendonuclease EcoRI gespalten wurde.

a) Das Oligonucleotid 2 wurde phosphoryliert und ein ddA-Rest, gemäß der obigen Beschreibung in Beispiel (2), an das 3'-Ende gefügt. Dieses modifizierte Oligonucleotid wurde dann an Oligo-

nucleotid 14 hybridisiert.

b) Genomische DNA I wurde mit der Restriktionsendonuclease EcoRI gespalten und, wie in Methode I beschrieben, dephosphoryliert.

c) Die hybridisierten Oligonucleotide 2 + 14 wurden dann an mit EcoRI gespaltene genomische DNA I ligiert.

d) Amplifikation erfolgte an den Ligationsprodukten unter Verwendung von:

(i) Oligonucleotid 58 und 62.

Das erwartete Produkt weist etwa 220 bp auf.

(ii) Oligonucleotid 58 und 63.

Das erwartete Produkt weist etwa 560 bp auf..

Beispiel 4

Verwendung der Oligonucleotide 3 und 14 mit DNA, die mit der Restriktionsendonuclease BamHI gespalten wurde.

a) Das Oligonucleotid 3 wurde phosphoryliert und ein ddA-Nucleotid an das 3'-Ende gefügt, wie es oben in Beispiel 1 beschrieben wurde.

b) Dieses präparierte Oligonucleotid wurde dann an Oligonucleotid 14 hybridisiert (annealed).

c) Genomische DNA I wurde mit der Restriktionsendonuclease BamHI gespalten und gemäß Beschreibung in Methode I dephosphoryliert.

d) Die hybridisierten Oligonucleotide 3 und 14 wurden dann an mit BamHI gespaltene genomische DNA I ligiert (siehe Methode I bezüglich Details).

e) Alle freien 3'-Enden in den Ligationsprodukten wurden dann nach dem in Schritt 6 von Methode 3 beschriebenen Verfahren mit einem ddG-Rest blockiert.

f) Schließlich wurde Amplifikation an diesen Produkten unter Einsatz der Oligonucleotide 58 und 63 ausgeführt.

Das erwartete Produkt weist etwa 200 bp auf.

Beispiel 5

Verwendung der Oligonucleotide 3 und 14 mit DNA, die mit der Restriktionsendonuclease BamHI gespalten wurde.

a) Oligonucleotid 3 wurde hergestellt und gemäß obiger Beschreibung in Beispiel (4) an Oligonucleotid 14 hybridisiert (annealed).

b) pUC8 mit Gehalt an einer 440bp-Insert-Plasmid-DNA (siehe B3, Figur 10) wurde bis zur Vollständigkeit mit der Restriktionsendonuclease BamHI digeriert und, wie im einzelnen in Methode I ausgeführt, dephosphoryliert.

c) Die Oligonucleotide 3 und 14 wurden an mit BamHI gespaltene B3-Plasmid-DNA hybridisiert.

d) Ein ddG-Rest wurde, wie in Beispiel 4 angegeben, an freie 3'-Enden geheftet. Es erfolgte Amplifikation an den Ligationsprodukten unter Einsatz der Oligonucleotide 58 und 65. Vorstehend beschriebene Figur 11(iii) zeigt die Marker Hind III in den Spuren 1 und 11, den mit Hae III gespaltenen Marker ϕ X174 in Spur 6 und das nach Beispiel 5 gewonnene amplifizierte Fragment in den Spuren 2, 3, 4 und 5. Wie erwartet ist das Produkt etwa 440 bp lang.

Beispiel 6

Verwendung der Oligonucleotide 26 und 40 mit DNA, die mit der Restriktionsendonuclease XbaI gespalten wurde.

a) Oligonucleotid 26 wurde phosphoryliert und die kinasierte Species gereinigt, wie es in Methode V beschrieben ist.

- b) Das kinasierte Oligonucleotid 26 wurde unter den in Schritt 3 von Methode I beschriebenen Bedingungen an Oligonucleotid 40 hybridisiert (annealed).
- c) Menschliche genomische DNA II wurde mit Xba I restringiert und, wie in Schritt 4 von Methode I ausgeführt, desphosphoryliert.
- d) Die hybridisierten Oligonucleotide 26 und 40 wurden dann an mit der Restriktionsendonuclease XbaI gespaltene genomische DNA II ligiert (siehe Schritt 5 von Methode I bezüglich Details).
- e) Alle freien 3'-Enden in den Ligationsprodukten wurden unter Anwendung der in Schritt 6 von Methode III beschriebenen Verfahrensweise mit einem ddA-Rest blockiert.
- f) Die Amplifikation erfolgte an den Ligationsprodukten mit:
- (i) den Oligonucleotiden 58 und 61.
Das erwartete Produkt ist etwa 800 bp lang.
 - (ii) den Oligonucleotiden 58 und 64.
Das erwartete Produkt ist etwa 760 bp lang.

Beispiel 7

Verwendung der Oligonucleotide 27 und 40 mit DNA, die mit der Restriktionsendonuclease EcoRI gespalten wurde.

- a) Oligonucleotid 27 wurde phosphoryliert, gereinigt und an Oligonucleotid 40 hybridisiert, wie es oben in Beispiel 5 beschrieben wurde.
- b) Genomische DNA I wurde mit der Restriktionsendonuclease EcoRI gespalten und desphosphoryliert, wie es in Schritt 4 von Methode I im einzelnen ausgeführt ist.
- c) Die hybridisierten Oligonucleotide 27 und 40 wurden dann an mit EcoRI gespaltenen genomischen DNA I ligiert (siehe Schritt 5 von Methode 1 bezüglich Details).

d) Alle freien 3'-Enden in den Ligationsprodukten wurden, wie in Schritt 6 von Methode III beschrieben, mit ddA-Rest blockiert,

e) Die Amplifikation erfolgte an den Ligationsprodukten unter Verwendung von:

(i) Oligonucleotid 58 und 62.

Das erwartete Produkt ist etwa 220 bp lang.

(ii) Oligonucleotid 58 und 63.

Das erwartete Produkt ist 560 bp lang.

Vorstehend beschriebene Figur 11(iii) zeigt den Marker Hind III in den Spuren 1 und 11, den mit Hae III gespaltenen Marker ϕ X174 in Spur 6 und die nach Beispiel 7 gewonnenen amplifizierten Fragmente in den Spuren 7, 8, 9 und 10.

Beispiel 8

Eine EcoRI-Vektorettbibliothek wurde gemäß Methode VI konstruiert, und Vektoretteneinheit 6 (Oligonucleotid 27, 40) wurde zusammen mit DNA der Zelllinie 7387 verwendet.

Die Amplifikationsreaktion wurde an 10 ng der konstruierten Bibliothek ausgeführt, um die in Figur 12 dargestellte Sequenz zu amplifizieren.

Amplifikationsbedingungen

Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

10 ng konstruierte EcoRI-Vektorettbibliothek

100 pmol 58

100 pmol 63

100 μ m (Endkonzentration) dNTPs (siehe M1.7 gemäß nachstehender Definition)

Das Volumen wurde mit sterilem doppelt destilliertem H₂O auf 100 μ m eingestellt. 50 μ l leichtes Mineralöl (Sigma) wurden vorsichtig über das Reaktionsgemisch gegeben, um Verdampfung zu

verhindern. Das Röhrchen wurde dann in einen Techne Dri-Block gegeben, der auf 96 °C eingestellt war. Man ließ die DNA 10 Minuten bei 96 °C denaturieren. Danach ließ man den Block auf 90 °C abkühlen. Zwei Einheiten Taq-Polymerase (Cetus) wurden dem Reaktionsgemisch zugesetzt. Die folgenden Temperaturen und Zeiten wurden zur Amplifizierung der Sequenz zwischen den Oligonucleotiden 58 und 63 angewendet:

91 °	2 Minuten
62 °	2 Minuten
72 °	2 Minuten

Es wurden 40 Zyklen durchgeführt. Beim abschließenden Zyklus wurde die Temperatur des Blocks 9,9 Minuten bei 72 °C gehalten, dann ließ man sie im Verlaufe einer Stunde auf Raumtemperatur abkühlen.

15 µl des Amplifikationsreaktionsgemisches wurden durch Agarosegelelektrophorese analysiert.

Das Resultat ist in Figur 13 dargestellt. Sie zeigt das gemäß diesem Beispiel gewonnene Amplifikationsprodukt in Spur 1 und den mit Hae III gespaltenen Marker ØX174 in spur 2. Das Amplifikationsprodukt besitzt die erwartete Größe (ca. 600 bp).

Beispiel 9

Dieses Beispiel wurde nach Methode IX wie folgt ausgeführt:

Schritt 1

Phosphorylierung des "Top"-Oligonucleotids

Oligonucleotid 49, 51, 52, 53, 55, 27, 57 und 2.

Puffer und andere Standardlösungen

<u>1.1</u>	<u>10x Kinasepuffer</u>
	0,5 M Tris-HCl (pH 7,6)
	0,1 M MgCl ₂
	50 mM Dithiothreitol
	1 mM Spermidin

1 mM EDTA
Aufbewahrung bei -20 °C.

1.2 Adenosin-5'-triphosphat
Von Pharmacia gekauftes Natriumsalz als 10-mM-Lösung.
Verdünt in Wasser auf 20 pmol/µl.
Aufbewahrung bei -20 °C.

Verfahrensweise

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen gemischt:

Oligonucleotid	10 pmol
10x Kinasepuffer (I.1)	1 µl
ATP (I.2)	1 µl (20 pmol)
T4-Polynucleotidkinase (Amersham)	5 Einheiten
H ₂ O	auf 10 µl

Das Molverhältnis der Oligonucleotide zu ATP wurde bei 1:2 gehalten. Die Phosphorylierung der Oligonucleotide erfolgte durch 60minütige Inkubation bei 37 °C.

Am Ende dieser Periode wurde die T4-Polynucleotidkinase durch 10minütige Inkubation bei 90 °C inaktiviert. Die Röhrchen wurden dann 10 Sekunden in einer Mikrozentrifuge geschleudert, um das gesamte Kondensat zu sammeln.

Die phosphorylierten Species werden bei dieser Methode nicht gelgereinigt.

Schritt 2

Hybridisierung (annealing) des phosphorylierten 'Top'-Oligonucleotids an das entsprechende 'Bottom'-Oligonucleotid

Erforderliche Lösung:

2.1 10x Hybridisierungspuffer
100 mM Tris-HCl (pH 8,0)
100 mM MgCl₂

Folgendes wurde in einem 1,5-ml-Sarstedt-Röhrchen mit Schraubkappe gemischt:

- 10 pmol (in 10 µl) phosphoryliertes 'Top'-Oligonucleptid von Schritt 1
- 10 pmol des entsprechenden 'Bottom'- lignonucleotids
- 2 µl 10x Hybridisierungspuffer (2.1)
- H₂O auf 20 µl

Das Röhrchen wurde in ein siedendes Wasserbad gelegt. Nach 5 Minuten wurde die Erhitzung unterbrochen, und man ließ das Wasserbad über einen Zeitraum von etwa 2 Stunden langsam auf Raumtemperatur abkühlen. Schließlich wurde der Inhalt des Röhrchens am Boden durch 10minütiges Zentrifugieren in einer Mikrozentrifuge gesammelt. Die hybridisierten Oligonucleotide mit einer Konzentration von 0,5 pmol/µl sind dann zum Einsatz bei Ligationen bereit.

Die Vektoretteneinheiten 1 - 8 entsprachen der Definition in vorstehender Tabelle 1.

Schritt 3

Herstellung von DNA

Es wurde DNA der Zelllinie 7387 verwendet.

Erforderliche Puffer:

- 3.1 10x EcoRI-Puffer
100 mM Tris-HCl, pH 7,5
100 mM MgCl₂
500 mM NaCl

- 3.2 TE-Puffer
10 mM Tris-HCl, pH 7,5
1 mM EDTA

- 3.3 Agaroseprobenpuffer
15 % (Masse/Vol.) Ficoll 400

tor) sichtbar gemacht.

(iii) (a) Die digerierte DNA wurde durch Phenolextraktion gereinigt:

Ein gleiches Volumen gepuffertes redestilliertes Phenol (BRL, DNA-Qualität mit TE-Puffer, 3.2) wurde zu der restringierten DNA gegeben, und das Gemisch wurde 30 Sekunden wirbelbehandelt. Die gebildete weiße Emulsion durch 5 Minuten lange Mikrozentrifugierung in zwei Phasen getrennt. Die wäßrige Phase wurde in ein frisches Röhrchen transferiert. Alles restliche Phenol wurde entfernt, indem die wäßrige Schicht zuerst mit einem gleichen Volumen Phenol:Chloroform:Isoamylalcohol (25:24:1) - ein gleiches Volumen - und dann mit einem gleichen Volumen Chloroform:Isoamylalcohol (24:1) extrahiert wurde.

(b) Die DNA wurde schließlich durch Ethanolpräzipitation gereinigt:

Zu der DNA-Lösung wurden 3 M Natriumacetat gegeben, damit eine Endkonzentration von 300 mM Natriumacetat entstand. Die DNA wurde durch den Zusatz von 2,5 Volumen 95 %iges Ethanol und 30minütiges Abkühlen bei -80°C präzipitiert, woran sich Zentrifugieren anschloß (Mikrozentrifuge, 5 Minuten, 12 000 U/min). Die pelletierte DNA wurde mit 80 % (Vol./Vol.) kaltem Ethanol gewaschen und unter Vakuum getrocknet. Die DNA wurde in dem TE-Puffer (3.2) gelöst, um eine Konzentration von 0,5 $\mu\text{g}/\text{l}$ zu erhalten.

Schritt 4

Ligation der hybridisierten Oligonucleotide an die gespaltene DNA Puffer und andere Standardlösungen

4.1 10x Ligationspuffer
200 mM Tris-HCl, pH 7,4
100 mM MgCl_2
100 mM DTT

4.2 10xATP
10 mM ATP

4.3 0,5 M NaCl

Die Berechnung der pmol klebriger Enden in 1 µg gespaltener DNA wurde in 4.1(i) von Methode I dargestellt.

Ligationen

Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

Gespaltene 7387-DNA	1,48 pmol (2 µg)
Hybridisierte Oligonucleotide (von Schritt 2)	3 pmol
10x Ligationspuffer (4.1)	2 µl
10x ATP (4.2)	2 µl
T4-DNA-Ligase (Boehringer Mannheim)	2 Einheiten
H ₂ O Endvolumen auf	20 µl

Das Ligationsgemisch wurde 120 Minuten bei 25 °C stehen gelassen. Dann wurde NaCl (4.3) bis zu einer Endkonzentration von 50 mM zu dem Ligationsgemisch hinzugegeben. Alle Fragmente genomischer DNA, die an andere Fragmente genomischer DNA ligiert waren, wurden durch den Zusatz von 2 Einheiten EcoRI (Boehringer Mannheim) und 30minütige Inkubation der Probe bei 37 °C erneut geschnitten. Vektoretteneinheiten wurden dann durch Zusatz von:

ATP auf 1 mM Konzentration (4.2)
1,5 pmol hybridisierte Oligonucleotide
1 Einheit T4-DNA-Ligase
(Boehringer Mannheim)

und 1stündige Inkubation bei 25 °C an diese erneut geschnittenen Fragmente genomischer DNA ligiert. Nach dem Neuschneiden der DNA mit EcoRI bei 37 °C im Verlaufe von 30 Minuten wird die Restriktionsendonuclease nicht inaktiviert.

Deshalb werden während der zweiten und letzten 60minütigen Ligation bei 25 °C alle Fragmente genomischer DNA, die an andere Fragmente genomischer DNA ligieren, wahrscheinlich in der glei-

chen Reaktion neu gespalten, so daß am Ende dieser Restriktions/Ligationsreaktion im Prinzip 100 % der Fragmente genomischer DNA die Vektoretteeinheiten an beiden Enden haben. Ligation der Vektoretteeinheit, um genomische DNA zu schneiden, inaktiviert jene spezielle Restriktionsstelle.

Die DNA (Reaktionsgemisch) wurde dann auf 10 ng/ μ l verdünnt.

EcoRI-Vektoretteeinheiten wurden unter Anwendung dieser Methode und der Vektoretteeinheiten 1 bis 8 zusammen mit DNA der Zelllinie 7387 konstruiert.

Es wurde auch eine Kontroll-EcoRI-Bibliothek konstruiert, wo die Vektoretteeinheiten aus den oben beschriebenen Ligationen ausgelassen wurden.

Schritt 5

Auf die Vektorettebibliotheken angewendete Amplifikation

10 ng von jeder der konstruierten Bibliotheken (für die Vektoretteeinheiten 1 - 8 und die Kontrollbibliothek von DNA ohne Vektoretten wurden der Amplifikation unterzogen. Die in Figur 12 gezeigte Sequenz wurde unter Verwendung der Oligonucleotide 58 und 63 als Amplimer amplifiziert, und die Amplifikationsbedingungen waren wie in Beispiel 8 beschrieben. 15 μ l des Amplifikationsproduktes wurden durch Agarosegelelektrophorese analysiert.

Das Ergebnis ist wie in Figur 14 dargestellt. Das Produkt der erwarteten Größe (ca. 600 bp) wurde bei allen Reaktionen erzeugt, ausgenommen dort, wo die Substrat-DNA von der Bibliothek stammt, die ohne Vektoretten im Ligationsgemisch konstruiert wurde. Das entspricht der Erwartung, weil die Sequenz des Oligonucleotids (einer der obigen Amplimer) sich innerhalb der Vektoretteeinheiten befindet.

Diese Ergebnisse zeigen, daß die Konstruktion der Vektoretteeinheiten keinen signifikanten Einfluß auf die Amplifikation der Nucleotidsequenz zwischen den Vektoretteeinheiten und einer ge-

wünschten genomischen Sequenz hat (d. h. Oligonucleotid 63 oben).

Die Ergebnisse sind in Figur 14 dargestellt. Sie ist eine Photographie von einem mit Ethidiumbromid angefärbten Gel, welches die gewonnenen Produkte zeigt, die nachstehend aufgeführt sind:

1. mit 7387-DNA + Vektoretteeinheit 1 konstruierte Bibliothek
2. mit 7387-DNA + Vektoretteeinheit 2 konstruierte Bibliothek
3. mit 7387-DNA + Vektoretteeinheit 3 konstruierte Bibliothek
4. mit 7387-DNA + Vektoretteeinheit 4 konstruierte Bibliothek
5. mit 7387-DNA + Vektoretteeinheit 5 konstruierte Bibliothek
6. mit 7387-DNA + Vektoretteeinheit 6 konstruierte Bibliothek
7. mit 7387-DNA + Vektoretteeinheit 7 konstruierte Bibliothek
8. mit 7387-DNA ohne Vektoretten konstruierte Bibliothek
9. ϕ X174 HaeIII-Marker
10. ϕ X174 HaeIII-Marker
11. mit 7387-DNA + Vektoretteeinheit 8 konstruierte Bibliothek

In jedem Fall wird das Produkt mit der erwarteten Größe gewonnen.

Beispiel 10

Dieses Beispiel wurde nach Methode X entsprechend vorstehender Beschreibung durchgeführt.

Es wurden EcoRI-Vektorettebibliotheken unter Anwendung dieser Methode und mit den Vektoretteeinheiten 1 bis 7 gemäß hierin gegebener Definition und mit DNA der Zelllinie 7387 konstruiert. Es wurde auch eine Kontroll-EcoRI-Bibliothek konstruiert, wo die Vektoretteeinheiten aus den oben beschriebenen Ligationen ausgeschlossen wurden.

Auf die Vektorettebibliotheken angewendete Amplifikation

10 ng von jeder der konstruierten Bibliotheken (für die Vektoretteeinheiten 1 bis 7 und die Kontrollbibliothek von DNA ohne Vektoretten) wurden der Amplifikation unterzogen. Die in Beispiel 8 gezeigte Sequenz wurde unter Verwendung der Oligonucleotide 58 und 63 als Amplimer amplifiziert. Die in Beispiel 8 be-

schriebenen Amplifikationsbedingungen wurden angewendet.

15 µl des Produktes wurden durch Agarosegelelektrophorese analysiert.

Das Resultat ist wie in Figur 15 dargestellt. Das Produkt der erwarteten Größe (ca. 600 bp) wurde bei allen Reaktionen erzeugt, ausgenommen dort, wo die Substrat-DNA von der Bibliothek stammt, die ohne Vektoretten im Ligationsgemisch konstruiert wurde. Das entspricht der Erwartung, weil sich die Sequenz von Oligonucleotid 58 (einer der obigen Amplimer) innerhalb der Vektoretteeinheiten befindet.

Die Spuren der in Figur 15 gezeigten Photographie sind wie folgt:

1. ØX174 HaeIII-Marker
2. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 1
3. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 2
4. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 3
5. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 4
6. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 5
7. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 6
8. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 7
9. Amplifikationsprodukt von der Bibliothek von 7387-DNA ohne Vektoretten
10. ØX174 HaeIII-Marker.

Für jede Bibliothek wird das erwartete Resultat erzielt.

Diese Ergebnisse zeigen, daß die Konstruktion der Vektoretteeinheiten keinen signifikanten Einfluß auf die Amplifikation der

Nucleotidsequenz zwischen den Vektoretteneinheiten und einer gewünschten genomischen Sequenz hat (d. h. Oligonucleotid 63 oben).

Beispiel 11

Dieses Beispiel wurde nach Methode XI entsprechend vorstehender Beschreibung durchgeführt.

Es wurden EcoRI-Vektorettbibliotheken unter Anwendung dieser Methode und mit den Vektoretteneinheiten 1 bis 7 und DNA der Zelllinie 7387 konstruiert.

Es wurde auch eine Kontroll-EcoRI-Bibliothek konstruiert, wo die Vektoretteneinheiten aus den oben beschriebenen Ligationen ausgelassen wurden.

Auf die Vektorettbibliotheken angewendete Amplifikation

10 ng von jeder der konstruierten Bibliotheken (für die Vektoretteneinheiten 1 bis 7 und die Kontrollbibliothek von DNA ohne Vektoretten) wurden der Polymerasekettenreaktion unterzogen. Die in Beispiel 1 gezeigte Sequenz wurde unter Verwendung der Oligonucleotide 58 und 63 als Amplimer amplifiziert. Die in Beispiel 8 beschriebenen Amplifikationsbedingungen wurden angewendet.

15 µl des Produktes wurden durch Agarosegelelektrophorese analysiert.

Das Ergebnis ist wie in Figur 16 dargestellt. Das Produkt mit der erwarteten Größe (ca. 600 bp) wurde bei allen Reaktion erzeugt, ausgenommen dort, wo die Substrat-DNA von der Bibliothek stammt, die ohne Vektoretten im Ligationsgemisch konstruiert wurde. Das entspricht der Erwartung, weil die Sequenz von Oligonucleotid 58 (einer der obigen Amplimer) sich innerhalb der Vektoretteneinheiten befindet.

Die Spuren der in Figur 16 gezeigten Photographie sind wie folgt:

1. ØX174 HaeIII-Marker
2. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 1

3. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 2
4. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 3
5. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 4
6. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 5
7. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 6
8. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 7
9. Amplifikationsprodukt von der Bibliothek von 7387-DNA ohne Vektoretten
10. ϕ X174 HaeIII-Marker.

Die Ergebnisse zeigen, daß die Konstruktion der Vektoretteneinheiten keinen signifikanten nachteiligen Einfluß auf die Amplifikation der Nucleotidsequenz zwischen den Vektoretteneinheiten und einer gewünschten genomischen Sequenz hat (d. h. Oligonucleotid 63 oben).

Beispiel 12

Dieses Beispiel wurde nach Methode XII entsprechend vorstehender Beschreibung durchgeführt.

Es wurden EcoRI-Vektorettbibliotheken unter Anwendung dieser Methode und mit den Vektoretteneinheiten 1 bis 7 und DNA der Zelllinie 7387 konstruiert. Es wurde auch eine Kontroll-EcoRI-Bibliothek konstruiert, wo die Vektoretteneinheiten aus den oben beschriebenen Ligationen ausgelassen wurden.

Auf die Vektorettbibliotheken angewendete Amplifikation

10 ng von jeder der konstruierten Bibliotheken (für die Vektoretteneinheiten 1 bis 7 und die Kontrollbibliothek von DNA ohne Vektoretten) wurden der Amplifikation unterzogen. Die in Beispiel 1 gezeigte Sequenz wurde unter Verwendung der Oligonucleo-

tide 58 und 63 als Amplimer amplifiziert. Die in Beispiel 8 beschriebenen Amplifikationsbedingungen wurden angewendet.

15 µl des Amplifikationsproduktes wurden durch Agarosegelelektrophorese analysiert.

Das Ergebnis ist wie in Figur 17 dargestellt. Das Produkt mit der erwarteten Größe (ca. 600 bp) wurde bei allen Reaktionen erzeugt, ausgenommen dort, wo die Substrat-DNA von der Kontrollbibliothek stammt, die ohne Vektoretten im Ligationsgemisch konstruiert wurde. Das entspricht den Erwartungen, weil die Sequenz von Oligonucleotid 58 (einer der obigen Amplimer) sich innerhalb der Vektoretteeinheiten befindet.

Die Spuren der in Figur 17 gezeigten Photographie sind wie folgt:

1. ØX174 HaeIII-Marker
2. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 1
3. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 2
4. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 3
5. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 4
6. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 5
7. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 6
8. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteeinheit 7
9. Amplifikationsprodukt von der Bibliothek von 7387-DNA ohne Vektoretten
10. ØX174 HaeIII-Marker

Diese Resultate zeigen, daß die Konstruktion der Vektoretteeinheiten keinen signifikanten Einfluß auf die Amplifikation der Nucleotidsequenz zwischen den Vektoretteeinheiten und einer gewünschten genomischen Sequenz hat (d. h. Oligonucleotid 63 oben).

Beispiel 13

Dieses Beispiel wurde nach Methode XIII entsprechend vorstehender Beschreibung durchgeführt.

Es wurden EcoRI-Vektorettbibliotheken unter Anwendung dieser Methode und mit den Vektoretteneinheiten 1 bis 7 und DNA der Zelllinie 7387 konstruiert.

Es wurde auch eine Kontroll-EcoRI-Bibliothek konstruiert, wo die Vektoretteneinheiten aus der Ligationsreaktion ausgelassen wurden.

Auf die Vektorettbibliotheken angewendete Amplifikation

10 ng von jeder der konstruierten Bibliotheken (für die Vektoretteneinheiten 1 bis 7 und die Kontrollbibliothek von DNA ohne Vektoretten) wurden der Amplifikation unterzogen. Die in Beispiel 1 gezeigte Sequenz wurde unter Verwendung der Oligonucleotide 58 und 63 als Amplimer amplifiziert. Die in Beispiel 8 beschriebenen PCR-Bedingungen wurden angewendet.

15 µl des Amplifikationsproduktes wurden durch Agarosegelelektrophorese analysiert.

Das Resultat ist wie in Figur 18 dargestellt. Das Produkt der erwarteten Größe (600 bp) wurde bei allen Reaktionen erzeugt, ausgenommen dort, wo die Substrat-DNA von der Bibliothek stammt, die ohne Vektoretten im Ligationsgemisch konstruiert wurde. Das entspricht der Erwartung, weil die Sequenz von Oligonucleotid 58 (einer der obigen Amplimer) sich innerhalb der Vektoretteneinheiten befindet.

Die Spuren der in Figur 18 gezeigten Photographie sind wie folgt:

1. ØX174 HaeIII-Marker
2. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 1
3. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 2
4. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 3
5. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 4

6. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 5
7. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 6
8. Amplifikationsprodukt von der Bibliothek von 7387-DNA + Vektoretteneinheit 7
9. Amplifikationsprodukt von der Bibliothek von 7387-DNA ohne Vektoretten
10. ϕ X174 HaeIII-Marker

Diese Ergebnisse zeigen, daß die Konstruktion der Vektoretteneinheiten keinen signifikanten Einfluß auf die Amplifikation der Nucleotidsequenz zwischen den Vektoretteneinheiten und einer gewünschten genomischen Sequenz hat (siehe Oligonucleotid 63 oben).

Beispiel 14

Es wurden fünf verschiedene EcoRI-Vektorettbibliotheken auf folgende Weise unter Verwendung von DNA der Zelllinie 7387 und Vektoretteneinheit 8 konstruiert:

(1)

- (a) Eine Bibliothek wurde nach Methode IX konstruiert.
- (b) Eine Bibliothek wurde nach Methode X konstruiert.
- (c) Eine Bibliothek wurde nach Methode XI konstruiert.
- (d) Eine Bibliothek wurde nach Methode XII konstruiert.
- (e) Eine Bibliothek wurde nach Methode XIII konstruiert.

(2)

Vor der Amplifikation wurden die in (1) oben konstruierten Bibliotheken einer 3'-Endblockierreaktion unterzogen:

Puffer und andere Standardlösungen

<u>7.2.1.</u>	<u>T7-DNA-Polymerase-Puffer</u>	<u>5x</u>
	200 mM Tris-HCl	pH 7,5
	100 mM MgCl ₂	
	250 mM NaCl	

7.2.2. dNTPs-dATP

1 mM dTTP

1 mM dGTP

1 mM dCTP

aufgefüllt mit sterilisiertem deionisiertem Wasser.

7.2.3. ddATP

1 mM ddATP, aufgefüllt mit sterilisiertem deionisiertem Wasser.

Addition eines ddA-Rests an 3'-Enden von Produkten von (1) oben. Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

0,5 µl DNA von Schritt (1) oben

4 µl 5x T7-DNA-Polymerase-Puffer (7.2.1.)

2 µl dNTPs-dATP-Lösung (7.2.2.)

2 µl ddATP-Lösung (7.2.3.)

1 Einheit T7-DNA-Polymerase (Sequense)

H₂O auf 20 µl

Das Röhrchen wurde 30 Minuten lang bei 37 °C gehalten. Das Enzym wurde dann durch 10minütige Inkubation bei 70 °C denaturiert. Die Konzentration von DNA in allen Proben wurde dann in Wasser auf 10 ng/µl verdünnt.

Auf die Vektorettbibliotheken angewendete Amplifikation

10 ng von jeder der fünf konstruierten Bibliotheken wurden der Polymerasekettenreaktion unterzogen. Die in Beispiel 8 gezeigte Sequenz wurde unter Verwendung der Oligonucleotide 58 und 63 als Amplimer amplifiziert. Die Amplifikationsbedingungen von Beispiel 8 wurden angewendet.

15 µl des Amplifikationsproduktes wurden durch Agarosegelelektrophorese analysiert.

Das Resultat ist wie in Figur 19 dargestellt.

Die Spuren des in Figur 19 gezeigten Autoradiogramms sind wie folgt:

1. ØX174 HaeIII-Marker
2. Amplifikationsprodukt von 7387-DNA + Vektoretteneinheit 8, Methode IX + 3'-Endblockierung
3. Amplifikationsprodukt von 7387-DNA + Vektoretteneinheit 8, Methode X + 3'-Endblockierung
4. Amplifikationsprodukt von 7387-DNA + Vektoretteneinheit 8, Methode XI + 3'-Endblockierung
5. Amplifikationsprodukt von 7387-DNA + Vektoretteneinheit 8, Methode XII + 3'-Endblockierung
6. Amplifikationsprodukt von 7387-DNA + Vektoretteneinheit 8, Methode XIII + 3'-Endblockierung

Alle Produkte haben die erwartete Größe.

Beispiel 15

Es wurden 5 verschiedene EcoRI-Vektorettbibliotheken unter Verwendung von DNA der Zelllinie 7387 und Vektoretteneinheit 8 konstruiert.

- (a) Eine Bibliothek wurde nach Methode IX konstruiert.
- (b) Eine Bibliothek wurde nach Methode X konstruiert.
- (c) Eine Bibliothek wurde nach Methode XI konstruiert.
- (d) Eine Bibliothek wurde nach Methode XII konstruiert.
- (e) Eine Bibliothek wurde nach Methode XIII konstruiert.

Der Unterschied bestand darin, daß Oligonucleotid 2 nichtphosphoryliert war.

Die Bibliotheken (b) bis (e) wurden hergestellt, wie es für die Methoden IX bis XII beschrieben worden ist.

Die Bibliotheken (b), (c), (d), (e) und eine Aliquote von Bibliothek (a) wurden mit ddATP und T7-DNA-Polymerase (Sequenzase) behandelt, um freie 3'-Enden zu blockieren. Die Bedingungen waren, wie in Beispiel 14 beschrieben.

Die Amplifikation wurde an all diesen Bibliotheken ausgeführt. Die in Figur 12 gezeigte Sequenz wurde unter Verwendung der Oligonucleotide 63 und 58 amplifiziert. Die in Beispiel 8 beschriebenen Bedingungen wurden angewendet.

15 µl des Amplifikationsproduktes wurden durch Agarosegelelektrophorese (1,4 % (Masse/Vol.)Gel) analysiert.

Das erzielte Resultat war wie in Figur 20 dargestellt.

Die Spuren der in Figur 20 gezeigten Photographie sind wie folgt:

1. Amplifikationsprodukt von 7387-DNA + Vektoretteeinheit 8, Methode IX + 3'-Endblockierung (Oligonucleotid 2 nichtphosphoryliert)
2. Amplifikationsprodukt von 7387-DNA + Vektoretteeinheit 8, Methode IX + 3'-Endblockierung. (Oligonucleotid 2 nichtphosphoryliert)
3. Amplifikationsprodukt von 7387-DNA + Vektoretteeinheit 8, Methode X + 3'-Endblockierung. (Oligonucleotid 2 nichtphosphoryliert)
4. Amplifikationsprodukt von 7387-DNA + Vektoretteeinheit 8, Methode XII + 3'-Endblockierung (Oligonucleotid 2 nichtphosphoryliert)
5. Amplifikationsprodukt von 7387-DNA + Vektoretteeinheit 8, Methode XII + 3'-Endblockierung (Oligonucleotid 2 nichtphosphoryliert)
6. Amplifikationsprodukt von 7387-DNA + Vektoretteeinheit 8, Methode XIII + 3'-Endblockierung (Oligonucleotid nichtphosphoryliert)
7. ØX174 HaeIII-Marker

Alle Produkte haben die erwartete Größe.

Beispiel 16

Dieses Beispiel wurde ausgeführt, um zu zeigen, daß das in Figuren 13 - 20 gezeigte Produkt ein Produkt der Oligonucleotide 58 und 63 (siehe Beispiel 8) und nicht ein Produkt von einem der Amplimer ist.

Für jede Amplifikation wurden 10 ng von jeder der folgenden Vektorettebibliotheken verwendet:

Eine Bibliothek, die unter Verwendung von 7387-DNA und den Vektoretteneinheiten 1 bis 8 in Methode VII hergestellt wurde.

Eine Bibliothek, die unter Verwendung von 7387-DNA und Vektoretteneinheit 8 nach Methode VII + 3'-Endblockierung (bezüglich Details siehe Beispiel 14) hergestellt wurde.

Eine Bibliothek, die unter Verwendung von 7387-DNA und Vektoretteneinheit 8 (1b Oligonucleotid nichtphosphoryliert) nach Methode VII + 3'-Endblockierung hergestellt wurde (bezüglich Details siehe Beispiel 15).

Eine Bibliothek, die unter Verwendung von 7387-DNA ohne Vektorette nach Methode VII hergestellt wurde.

Um zu zeigen, daß die in Beispiel 8 aus den mit den Vektoretteneinheiten 1 bis 8 konstruierten Bibliotheken hergestellten Produkte alle identisch waren und die richtige Sequenz hatten, wurden die in Figur 16 (Produkte 2 bis 8) gezeigten Amplifikationsprodukte und das in Spur 11 von Figur 14 gezeigte Produkt unter Einsatz der Oligonucleotide 59 und 66 sequenziert.

Figur 21 zeigt eine Photokopie des von der Sequenzierung der Produkte 2 bis 8 von Figur 16 mit Oligonucleotid 66 erzeugten Autoradiogramms.

Das Autoradiogramm zeigt auch die Sequenzierung von Teilen von Intron 8 und Exon 9 des Phenylalaninhydroxylasegens (Phenylketonurie - PKU) als Kontrolle. Die Amplifikationsprodukte 2 bis 8 erscheinen von links nach rechts (wie in Figur 16 gezeigt) und sind alle identisch. Die Spurfolge ist CGTA.

Figur 22 zeigt die Nucleotidsequenzdaten, die von den in Beispiel 16 beschriebenen Produkten stammen.

Die Basen 1 bis 16 sind am 3'-Ende der veröffentlichten PKU-Intron-8-Sequenz. Die Basen 17 bis 452 bilden neue PKU-Intron-8-Sequenzdaten, die nach der erfindungsgemäßen Methode bestimmt wurden.

Beispiel 17

Das Produkt mit ca. 220 bp von Beispiel 7(i), Bahn (c), Spuren 7, 8, 9, 10, 11 wurde unter Verwendung der Oligonucleotide 62 und 67 sequenziert.

So wurde das 220-bp-Produkt von Beispiel 7(i)

1) in 3-Schritt-Zyklen amplifiziert, und zwar: a) 60 °C, 2 Minuten, b) 72 °C, 3 Minuten, und c) 91 °C, 2 Minuten, ausgeführt wurden 40 Zyklen (Verwendung von 0,5 µg DNA)/100 µl Beladung/Bahn (siehe Figur 24a)

2) und das Produkt von (1) oben wurde dann gelgereinigt und in 2-Schritt-Zyklen reamplifiziert, und zwar: a) 60 °C, 4 Minuten, und b) 91 °C, 2 Minuten. Es wurden 40 Zyklen durchgeführt (siehe Figur 24b).

Die Positionen der Oligonucleotide 58, 62 und 67 sind in Figur 23 dargestellt. In Figur 25 ist die in (a) angegebene Sequenz die mit Oligonucleotid 62 gelesene Sequenz, die in (b) angegebene Sequenz ist die mit Oligonucleotid 67 gelesene Sequenz. In (a) und (b) ist die unterstrichene Sequenz die mit Oligonucleotid 62 und 67 gelesene komplementäre/reverse Sequenz. Die gesamte 5' 3'-Sequenz ist in Figur 25(c) gezeigt, worin:

die Positionen 2 - 31 die 5'-3'-Sequenz von Oligonucleotid 62 zeigen,

die Positionen 75 - 104 die reverse komplementäre Sequenz von Oligonucleotid 67 zeigen,

die Positionen 95 - 104 (durch eine durchgehende Linie über der Sequenz gekennzeichnet) die ersten 10 Basen der veröffentlichten Intron-Sequenz sind (siehe Biochemistry 1985 24 S. 556-561 Kwok et al.),

die Positionen 2 bis 104 mit der veröffentlichten Sequenz für PKU Exon I und Teilen der Intronsequenzen übereinstimmen,

die Positionen 105 aufwärts die neuen Sequenzdaten für PKU Intron 1 darstellen.

Beispiel 18

Exon V des alpha-Antitrypsingens wurde der erfindungsgemäßen Methode (siehe Figur 26) unterzogen, wobei folgende Oligonucleotide eingesetzt wurden:

58 (Universalvektoretteprimer), 60 (Universalvektorettenucleotide, ineinandergeschachtelt), 68 (Exon V 5'), 69 (Exon V 5', ineinandergeschachtelt) und 70 (Exon V 3').

Die Vektorettebibliothek wurde hergestellt mit mit EcoRI geschnittener genomischer DNA mit hybridisierten (annealed) Vektoretteoligonucleotiden (87/40 entsprechend der Definition in dieser Beschreibung) und ligiert.

Die Amplifikation erfolgte folgendermaßen:

100 ng Vektorettebibliothek-DNA

Oligonucleotide 58 + 68 (100 pmol)

10 µl Puffer : 500 mM KCl

100 µM dNTP's

100 mM MgCl₂

2 Einheiten Taq-Polymerase

14 mM MgCl₂

Auffüllung des Volumens auf 100 µl

0,1 % Gelatine

mit doppelt destilliertem Wasser

Anfänglicher Denaturierungsschritt

93 °C, 10 Minuten lang

39 Zyklen von 93 °C, 2 Minuten

65 °C, 2 Minuten

72 °C, 8 Minuten

Abschließender Zyklus: 72 °C, 10 Minuten lang

Das Ergebnis wird in Figur 27(a) in Spur 4 gezeigt. Spur 3 zeigt eine Kontrollamplifikationsreaktion mit den Oligonucleotiden 68 und 70 zur Erzeugung eines Produktes mit ca. 200 bp.

Die primäre Amplifikationsrunde ergab eine Substanz ('smear') von etwa der erwarteten Größe. Das primäre Reaktionsgemisch wurde dann wie folgt verdünnt.

1 µl des Produktes von Spur 4 in Figur 27(a) wurde mit ddH₂O

(d. h. 10³ Verdünnung)

auf 1 ml aufgefüllt.

1 μ l davon wurde bei anschließenden PC-Reaktionen verwendet (gesamt = 100 μ l) (d. h. Gesamtverdünnung = 10^5) und unter Einsatz einer Reihe von Primerpaaren wie folgt reamplifiziert:

Zweite PCR (mit ineinandergeschachtelten Oligonucleotiden)

1 μ l Produkt von Spur 4 in Figur 27(a) (10^3)DNA

100 pmol Oligonucleotide 60 + 69

10 μ l 10X Puffer

10X: 500 mM KCl

100 μ m dNTP's

100 mM Tris pH 8,3

2 Einheiten Taq-Polymerase

10 mM MgCl₂

0,1 % Gelatine

Auffüllung des Volumens auf 100 μ l mit doppelt destilliertem Wasser. Reaktionsbedingungen:

Anfänglicher Denaturierungsschritt \rightarrow 95 °C, 10 Minuten lang

Das sekundäre Amplifikationsprodukt wird als 3,0-kb-Produkt in den Spuren 1, 2 und 3 von Figur 27(b) gezeigt.

Figur 27(b) ist eine Photographie eines Agarosegels, welches das Produkt der sekundären Amplifikation zwischen den Oligonucleotiden 60 und 69 zeigt:

Spur 1 das 3,0-kb-Produkt der sekundären Amplifikation zwischen den Oligonucleotiden 60 und 69,

Spur 2 das 3,0-kb-Produkt der sekundären Amplifikation zwischen den Oligonucleotiden 60 und 69,

Spur 3 das 3,0-kb-Produkt der sekundären Amplifikation zwischen den Oligonucleotiden 60 und 69,

Spur 4 den mit Hae III gespaltenen Marker λ 174 und den mit Hind III gespaltenen Marker λ .

Das Gel wurde auf eine Nylonmembran geblottet und mit einer α -1 Antitrypsinsonde untersucht. Es wurde eine einzelne Bande der richtigen Größe (3,0 kb) beobachtet und in Figur 27(c) gezeigt.

Figur 27(c) ist ein Autoradiogramm eines Nylonfilters (geblottet von dem in Figur 27(c) gezeigten Gel), der mit einer α -1 Anti-

trypsinsonde untersucht worden war. Die Position von DNA-Markern wird gezeigt.

Spur 1 ist das 3,0-kb-Produkt sekundärer Amplifikation zwischen den Oligonucleotiden 60 und 69.

Spur 2 ist das 3,0-kb-Produkt sekundärer Amplifikation zwischen den Oligonucleotiden 60 und 69.

Spur 3 ist das 3,0-kb-Produkt sekundärer Amplifikation zwischen den Oligonucleotiden 60 und 69.

Beispiel 19

D5 ist eine anonyme genomische Sequenz, die als ein 474-bp-XbaI-EcoRI-Fragment kloniert worden ist. Die Sequenz von D5 wurde durch Standardmethoden bestimmt (Newton et al., Nucleic Acids Research 16, 8233-8243, 1988) und wird in Figur 29 gezeigt. Es wurden zwei Oligonucleotidprimer, nämlich 71, 72, synthetisiert.

71 5' AAGTTTGAGCATAGGAAAAGTTCTGTGCCC

72 5' AGTTCTGTGCCCAAATTGCATCCAAG

Die Position der Oligonucleotide 71, 72 ist in Figur 29 angegeben.

Genomische DNA wurde von der Lymphoblastoidzelllinie, GM 7387, (NIGMS Human Genetic Mutant Cell Repository) isoliert, und Aliquoten (20 µg) wurden mit den folgenden Restriktionsenzymen, nämlich Dra I, Bcl I, Hind III und Hae III, digeriert. Es wurden dann Vektorettbibliotheken nach Methode X hergestellt.

Die primären Amplifikationen wurden an Aliquoten jeder Bibliothek gemäß nachfolgender Beschreibung durchgeführt.

Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

10 ng konstruierte Vektorettbibliothek

100 pmol Oligonucleotid 58

100 pmol Oligonucleotid 71

100 µm dNTPs

10 µl 10X-Puffer

10X Puffer ist:

500 mM KCl
 100 mM Tris-HCl,
 pH 8,3
 10 mM MgCl
 0,1 % Gelatine

Das Volumen wurde mit sterilem doppelt destilliertem Wasser auf 100 µl eingestellt. 50 µl leichtes Mineralöl (Sigma) wurde vorsichtig über das Reaktionsgemisch pipettiert, um Verdampfung zu verhindern.

Das Röhrchen wurde dann auf 96 °C erhitzt (in einem Techne Programmable Dri-Block PHC-1, der auf 96 °C äquilibriert worden war). Man ließ die DNA 10 Minuten bei 96 °C denaturieren. Den Block ließ man auf 90 °C abkühlen. Zwei Einheiten Taq-Polymerase (Cetus) wurden zu dem Reaktionsgemisch hinzugegeben. Die folgenden Temperaturen und Zeiten wurden angewendet, um die Sequenz zwischen den Oligonucleotiden 58 und 71 zu amplifizieren.

92 °C	2 Minuten
65 °C	2 Minuten
72 °C	5 Minuten

Es wurden 40 Zyklen durchgeführt. Während des abschließenden Zyklus wurde das Reaktionsgemisch 10 Minuten bei 72 °C gehalten. Aliquoten (15 µl) wurden von den Reaktionsgemischen abgenommen. Farbstoffgemisch (5 µl) wurde zugesetzt, und die Proben wurden auf 1,4 %igen Agarosegelen analysiert.

Farbstoffgemisch	155 (Masse/Vol.)	Ficoll 400
(Dye Loading Mix)	0,05 % (Masse/Vol.)	Bromphenolblau
	0,05 % (Masse/Vol.)	Xylencyanol
	Gelöst in 1X TBE	

Eine sekundäre Amplifikation wurde an Aliquoten der primären Reaktionsgemische durchgeführt. Eine Aliquote des primären Reaktionsgemisches (2 µl) wurde mit Wasser auf 400 µl verdünnt.

Eine Aliquote des verdünnten Materials (2 μ l) wurde in ein 0,5-ml-Sarsted-Röhrchen gebracht.

Folgendes wurde ebenfalls in das Röhrchen gegeben:

- 100 pmol Oligonucleotid 60
- 100 pmol Oligonucleotid 72
- 100 μ M dNTPs
- 10 μ l 10X-Puffer

10X-Puffer ist:

- 500 mM KCl
- 100 mM Tris-HCl, pH 8,3
- 10 mM MgCl
- 0,1 % Gelatine

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 μ l eingestellt. 50 μ l leichtes Mineralöl (Sigma) wurde vorsichtig über das Reaktionsgemisch pipettiert, um Verdampfung zu verhindern. Dann wurde das Röhrchen auf 90 $^{\circ}$ C erwärmt (im Techne Programmable Dri-Block PHC-1, der auf 96 $^{\circ}$ C äquilibriert worden war). Man ließ die DNA 10 Minuten bei 96 $^{\circ}$ C denaturieren. Den Block ließ man auf 90 $^{\circ}$ C abkühlen. Zwei Einheiten Taq-Polymerase (Cetus) wurden dem Reaktionsgemisch zugesetzt.

Die folgenden Temperaturen und Zeiten wurden angewendet, um die Sequenz zwischen den Oligonucleotiden 60.und 72 zu amplifizieren:

92 $^{\circ}$ C	2 Minuten
65 $^{\circ}$ C	2 Minuten
72 $^{\circ}$ C	5 Minuten

Es wurden 37 Zyklen durchgeführt. Während des abschließenden Zyklus wurde das Reaktionsgemisch 10 Minuten bei 72 $^{\circ}$ C gehalten. Aliquoten (15 μ l) wurden von den Reaktionsgemischen abgenommen. Farbstoffgemisch (Dye Loading Mix) (5 μ l) wurde zugesetzt, und die Proben wurden auf einem 1,4 %igen Agarosegel analysiert.

Farbstoffgemisch: 15 % (Masse/Vol.) Ficoll 400
(Dye Loading Mix) 0,05 % Bromphenolblau
0,05 % Xylencyanol
Gelöst in 1x TBE

Die Bibliotheken (Hae III, Hind III, Bcl I, Dra I) ergaben bestimmte Produkte im Größenbereich von 0,4 bis 1,5 kb (Figur 30). Aus der beobachteten Größe der Produkte konnte eine provisorische Restriktionskarte erstellt werden (Figur 31). Die Produkte wurden weiter durch Restriktionsdigests analysiert. Authentische Produkte sollten gespalten werden, um die Fragmente der vorhergesagten Größe zu gewinnen. So könnte das 0,82-kb-Bcl I-Produkt durch Hind III gespalten werden, um ein Produkt von 0,78 kb zu gewinnen, und könnte auch durch Hind III gespalten werden, um ein Fragment von 0,75 kb zu gewinnen. Die Analyse des 0,82-kb-BclI-Produktes durch Hind III und Hae III wird in Figur 30 gezeigt. Die vorausgesagten Produkte wurden festgestellt, und das BclI-Produkt wurde dann für die Sequenzanalyse eluiert.

So ist in Figur 30 das aus der Amplifikation einer HaeIII-Vektorettbibliothek resultierende Produkt in Spur 2 gezeigt, das aus der Amplifikation einer HindIII-Vektorettbibliothek resultierende Produkt ist in Spur 3 gezeigt, das aus der Amplifikation einer BclI-Vektorettbibliothek resultierende Produkt ist in Spur 4 gezeigt, und das aus der Amplifikation einer DraI-Vektorettbibliothek resultierende Produkt ist in Spur 5 zu sehen. Spur 7 zeigt die nach der Digestion des BclI-Produktes mit dem Restriktionsenzym Hae III gewonnenen Fragmente. Spur 8 zeigt die nach Digestion des BclI-Produktes mit dem Restriktionsenzym Hind III gewonnenen Fragmente. Spur 9 zeigt das BclI-Produkt, die Spuren 1 und 6 enthalten 1-kb-Marker (BRL).

Restriktionsdigestanalyse von Vektorettprodukten:

16 µl Amplifikationsreaktionsgemisch
2 µl 10 X Restriktionspuffer
2 µl Restriktionsenzym

Die Probe wurde 1 stunde bei 37 °C inkubiert. 5 µl Farbstoffgemisch (Dye Loading Mix) wurden der Probe zugesetzt, und die Pro-

dukte wurden auf 1,4 %igem Agarosegel analysiert.

Sequenzierung

Das BclI-Produkt wurde aus dem Agarosegel eluiert und unter Verwendung von Oligonucleotid 59 und Standardmethodologie sequenziert (Newton et al., Nucleic Acids Research 16, 8233-8243, 1988). Die Sequenz wird in Figur 31 veranschaulicht. Ein weiterer Primer (73) wurde aus dieser Sequenz konstruiert und in nachfolgenden Schritten verwendet.

73 5' GGCCTTTGANNAAGAGAAGAGTCAAGGATG

Die Untersuchung der Sequenz bestätigt das Vorhandensein der HaeIII- und HindIII-Stellen, die durch Schritte entsprechend der erfindungsgemäßen Methode vorausgesagt wurden.

Beispiel 20

Untersuchung des Chlamydia-Genoms mit der erfindungsgemäßen Methode

Chlamydia-trachomatis-Serotyp L2 wurde in mit Cycloheximid behandelten McCoy-Zellen gezüchtet. Elementarkörper wurden auf Glycerol-Tartrat-Gradienten gereinigt, wobei sich 10^9 Elementarkörper ergaben. DNA wurde durch Behandlung mit Proteinase K/SDS, Phenol/Chloroform behandelt und durch Ethanolpräzipitation rückgewonnen. Es wurden 90 mg DNA rückgewonnen, wovon 900 ng als Chlamydien-DNA eingeschätzt wurden und der Rest von den McCoy-Zellen abgeleitete Mäuse-DNA war.

DNA (5 µg, die äquivalent waren zu 50 ng Chlamydien-DNA) wurde entweder mit EcoRI oder BamHI digeriert. Vektorettbibliotheken wurden unter Verwendung der geeigneten Oligonucleotide (27, 40, EcoRI, 31, 40, BamHI) nach Methode X hergestellt. Die Bibliotheken wurden auf ein Endvolumen von 100 µl aufgefüllt.

Zwei Oligonucleotidprimer wurden auf der Basis der Consensus-Sequenz von MOMP (Major Outer Membrane Protein) von Chlamydia L2

synthetisiert. (Stephens et al., J Bacteriology 169, 3879-3885, 1977).

74 5' CTGCTCACGTAAATGCACAATTCCG

75 5' AACCGAGAGCAAAAGCCATATTGGC

Amplifikation

Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

10 µl	EcoRI-Vektorettbibliothek
20 µl	5X Puffer
	5X Puffer = 335 Tris-HCl, pH 8,5
	83 mM NH ₄ SO ₄
	10 mM MgCl
	50 mM 2-Mercaptoethanol
	0,8 mg/ml Rinderserumalbumin
100 pmol	Oligonucleotid 58
100 pmol	Oligonucleotid 74
200 µM	dNTPs

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 µl eingestellt. 50 µl leichtes Mineralöl (Sigma) wurden vorsichtig über das Reaktionsgemisch pipettiert, um Verdampfung zu verhindern. Das Röhrchen wurde dann auf 96 °C erhitzt (im Techne Programmable Dri-Block PHC-1, der auf 96 °C äquilibriert worden war). Man ließ die DNA 10 Minuten bei 96 °C denaturieren. Dann ließ man den Block auf 90 °C abkühlen. Dem Reaktionsgemisch wurden zwei Einheiten Taq-Polymerase (Cetus) zugesetzt. Die folgenden Temperaturen und Zeiten wurden angewendet, um die Sequenz zwischen den Oligonucleotiden 60 und 72 zu amplifizieren:

92 °C	0,7 Minuten
65 °C	0,7 Minuten
72 °C	1 Minute

Es wurden 40 Zyklen durchgeführt. 10 % des Reaktionsgemisches wurden für die Analyse auf Agarosegel entnommen. Die Resultate sind in Fig. 32 gezeigt, worin Spur 1 den mit Hae III gespaltenen Marker ϕ X174 zeigt.

Spur 2 zeigt das Produkt der Amplifikation zwischen den Oligonucleotiden 58 und 74 unter Verwendung von 10 µl einer EcoRI-Bibliothek.

Spur 3 zeigt das Produkt der Amplifikation zwischen den Oligonucleotiden 58 und 74 unter Verwendung von 1 µl einer EcoRI-Bibliothek.

Das Produkt aus der Amplifikation einer EcoRI-Bibliothek mit den Oligonucleotiden 58 und 74 wurde gereinigt. Asymmetrische Amplifikation (Proc. Natl. Acad. Sci. USA Bd. 85, S. 7652-7656, Oktober 1988, U.B. Gyllenstein und H. Ehrlich) des Produktes erfolgte mit 50 pmol Oligonucleotid 74 und Oligonucleotid 58 in 35 Zyklen gemäß obiger Beschreibung. Das amplifizierte Produkt wurde auf einer G50-Schleudersäule (G50 spun column) (wie beschrieben in: Molecular Cloning - A laboratory manual (Molekulares Klonieren - Ein Laborhandbuch); Maniatis, F., Fritsch, E.F., Sambrook, J., Cold Spring Harbor Laboratory, 1982) gereinigt und mit Oligonucleotid 59, wie in Beispiel 17 beschrieben, sequenziert. Die Sequenz wird in Figur 3A gezeigt. Die Digestion einer Aliquote des PCR-Produktes entweder mit Hinf I oder Cla I bestätigt die Anwesenheit dieser Restriktionsstellen (Figur 34).

In Figur 34 sind die Spuren der Photographie die folgenden:

- Spur 1 ØX174 HaeIII-DNA-Marker
- Spur 2 mit Cla I digeriertes EcoRI-Produkt
- Spur 3 EcoRI-Produkt
- Spur 4 mit Hinf I digeriertes EcoRI-Produkt
- Spur 5 ØX174 HaeIII-DNA-Marker

Beispiel 21

Auf das Chlamydia-Genom angewendete lineare Amplifikation

Folgendes wurde in ein 0,5-ml-Sarstedt-Röhrchen gebracht:

- 10 µl BamHI-Vektorettbibliothek
- 20 µl 5X Puffer (in Beispiel 20 definiert)
- 100 pmol Oligonucleotid 75
- 200 µm dNTPs

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 µl eingestellt. 50 µl leichtes Mineralöl (Sigma) wurde vorsichtig über das Reaktionsgemisch pipettiert, um Verdampfung zu verhindern.

Dann wurde das Röhrchen auf 96 °C erhitzt (im Techne Programmable Dri-Block PHC-1, der auf 96 °C äquilibriert worden war). Man ließ die DNA 10 Minuten bei 96 °C denaturieren. Man ließ den Block auf 90 °C abkühlen.

Zwei Einheiten Taq-Polymerase (Cetus) wurden dem Reaktionsgemisch zugesetzt. Die folgenden Temperaturen und Zeiten wurden angewendet, um die Sequenz zwischen den Oligonucleotiden 75 und der BamHI-Stelle zu amplifizieren:

93 °C	0,7 Minuten
65 °C	0,7 Minuten
72 °C	1,5 Minuten.

Es wurden 40 Zyklen ausgeführt. 100 pmol Oligonucleotid 58 und zwei Einheiten Taq-Polymerase wurden dann dem Reaktionsgemisch zugesetzt, und die Amplifikation wurde für weitere 20 Zyklen gemäß obiger Beschreibung fortgesetzt. Eine Aliquote Reaktionsgemisch wurde durch Agarosegelelektrophorese analysiert (Figur 35).

Es wurde eine zusätzliche Reaktion ausgeführt, bei der eine BamHI-Bibliothek unter Einsatz der Oligonucleotide 75 und 58 unter den oben beschriebenen Reaktionsbedingungen amplifiziert wurde. Eine Aliquote des Reaktionsgemisches wurde durch Agarosegelelektrophorese (Figur 35) analysiert.

In Figur 35 zeigt:

- Spur 1 den mit Hae III gespaltenen Marker ØX174,
- Spur 2 das Produkt der linearen Amplifikation einer BamHI-Vektorbibliothek unter Einsatz des Oligonucleotids 75,
- Spur 3 das aus 20 Zyklen sekundärer Amplifikation mit den Oligonucleotiden 58 und 75 resultierende Produkt,
- Spur 4 das aus 35 Zyklen sekundärer Amplifikation mit den Oligonucleotiden 58 und 75 resultierende Produkt,
- Spur 5 das aus 40 Zyklen primärer Amplifikation mit den Oligonucleotiden 58 und 75 resultierende Produkt.

Beispiel 22

Eine Humangenombibliothek wurde unter Verwendung des Cosmidvektors p Cos EMBL2 nach den veröffentlichten Verfahrensweisen konstruiert (PFR Little (1987) in: DNA cloning - A practical approach (DNA-Klonierung - eine praktische Methode), Band III, S. 19-42, herausgegeben von DM Glover). Die Bibliothek wurde durch Hybridisierung an Sonde KM19 gescreent (X. Estivill et al. (1987) Nature 326, 840 - 845). Mehrere positive Klone wurden identifiziert, und es wurde ein Klon 4.17 selektioniert, der die KM19-Sequenz in einem 40-kb-Insert enthielt.

Cosmid 4.17 wurde mit Hind III digeriert, und eine Vektorettbibliothek wurde nach Methode X hergestellt.

Es wurden zwei Oligonucleotidprimer, 76, 77, synthetisiert:

76 5' GGAAGGCCTTCAAATAACACTGTAGCC

77 5' GCTGCATCATATAAGTTGCC

Die Positionen dieser Primer sind in Figur 36(a) gezeigt. PCR-Amplifikation wurde an Aliquoten der Bibliothek gemäß nachstehender Beschreibung ausgeführt.

Folgendes wurde in einem 0,5-ml-Sarstedt-Röhrchen gemischt:

0,1 µg Cosmidvektorettbibliothek

100 pmol Oligonucleotid 58

100 pmol Oligonucleotid 76

100 µM dNTPs

10 µl 10x Puffer

10x Puffer ist: 500 mM KCl

100 mM Tris-HCl,
pH 8,3

10 mM MgCl₂

0,1 % Gelatine

Es wurde eine zweite Reaktion mit folgendem Inhalt eines 0,5-ml-Sarstedt-Röhrchen durchgeführt:

0,1 µg Cosmidvektorettbibliothek
 100 pmol Oligonucleotid 58
 100 pmol Oligonucleotid 77
 100 µM dNTPs
 10 µl 10x Puffer

Das Volumen der Reaktionsgemische wurde mit sterilem doppelt destilliertem Wasser auf 100 µl eingestellt. 50 µl leichtes Mineralöl (Sigma) wurden vorsichtig über das Reaktionsgemisch pipettiert, um Verdampfung zu verhindern. Dann wurden die Röhren in einen Thermal Cycler (Techne Programmable Dri-Block PHC-1) gelegt, der auf 96 °C äquilibriert worden war. Man ließ den Block auf 96 °C abkühlen. Dem Reaktionsgemisch wurden zwei Einheiten Taq-Polymerase (Cetus) zugesetzt. Die folgenden Temperaturen und Zeiten wurden angewendet, um die Sequenzen zwischen den Oligonucleotiden 58 und 76 und zwischen 58 und 77 zu amplifizieren:

92 °C	2 Minuten
65 °C	2 Minuten
72 °C	5 Minuten

Es wurden 40 Zyklen durchgeführt. Während des abschließenden Zyklus wurde das Reaktionsgemisch 10 Minuten lang bei 72 °C gehalten.

Aliquoten (15 µl) wurden von den Reaktionsgemischen abgenommen. Farbstoffgemisch (Dye Loading Mix) (5 µl) wurde zugesetzt, und die Proben wurden auf 1,4 %igen Agarosegelen analysiert.

Farbstoffgemisch: 15 % (Masse/Vol.) Ficoll 400
 (Dye Loading Mix) 0,05 % (Masse/Vol.) Bromphenolblau
 0,05 % (Masse/Vol.) Xylencyanol

Gelöst in 1 x TBE

Die Ergebnisse sind in Figur 36(b) gezeigt.

Spur 1 ØX174 HaeIII-DNA-Marker
 Spur 2 230-bp-Produkt (Oligonucleotid 58, Oligonucleotid 78)
 Spur 3 1200-bp-Produkt (Oligonucleotid 58, Oligonucleotid 77)
 Spur 4 ØX174 HaeIII-DNA-Marker

Beispiel 23

Die folgenden Oligonucleotide wurden synthetisiert.

78 5' CTCTCCCTTCTCCGGTGATGCCGGCCACGATGCGTCCGGCGGTCTCTCCTTC

Ein modifiziertes Oligonucleotid vom Typ E, das eine Sequenz enthält, die die Nucleotidreste 385 - 414 des Tetracyclinresistenzgens von Plasmid pBR322 darstellt (T. Maniatis, E.F. Fritsch und J. Sambrock (1982) Molecular Cloning - A laboratory manual (Molekulares Klonieren - Ein Laborhandbuch) S. 479 - 503). Zur Verwendung in Verbindung mit Oligonucleotiden vom Typ D.

79 5' CCGGTGATGCCGGCCACGATGCGTCCGGCG

Ein modifizierter Universalvektoretteprimer, der die Nucleotidreste 385-414 des Tetracyclinresistenzgens von Plasmid pBR322 (T. Maniatis et al. Molecular Cloning - A laboratory manual (Molekulares Klonieren - Ein Laborhandbuch, S. 479 - 503) darstellt. Zur Verwendung mit Oligonucleotid 78 und Oligonucleotiden 80/81, 82/81 und 83/81 gemäß nachstehender Beschreibung.

80 5' AATTGGCGGCCGCCATCCTAATTCTGTCTGAAGGTAAGGAACGGAGGAGAGAACT

Ein modifiziertes Oligonucleotid vom Typ D mit einem kohäsiven EcoRI-Ende, einer Not I-Erkennungsstelle und einer Fok I-Erkennungsstelle. Zur Verwendung mit Oligonucleotid 81.

81 5' AGTTCTCCCGGTGATGCCGGCCACGATGCGTCCGGCGGATGGCGGCCGCC

Ein modifiziertes Oligonucleotid vom Typ E.

82 5' AGCTGGCGGCCGCCATCCTAATTCTGTCTGAAGGTAAGGAACGGAGGAGAGAACT

Ein modifiziertes Oligonucleotid vom Typ D mit einem kohäsiven HindIII-Ende, einer Not I-Erkennungsstelle und einer Fok-I-Erkennungsstelle. Zur Verwendung mit Oligonucleotid 81.

83 5' AGCTGGCGGCCGCCATCC -NH₂

Ein modifiziertes Oligonucleotid vom Typ A, das ein kohäsives Hind-

III-Ende mit einer den 3'-Terminus blockierenden Aminohexylmodifikation enthält. Zur Verwendung mit Oligonucleotid 81.

84 5' GGCTCAACAGTCAGTTTGAAGTAGC

Ein einer EcoRI-Stelle in der KM19-Region benachbarter Primer.

Restriktionskarten, die für jede der folgenden Amplifikationen geeignet sind, sind in Figur 37 dargestellt. Es zeigen

a) eine Restriktionskarte eines Segments genomischer DNA, die an eine EcoRI-Vektorette ligiert ist. Die Positionen der Oligonucleotide 27, 76, 78 und 79 sind angegeben.

b) eine Restriktionskarte eines Segments genomischer DNA, das an eine EcoRI-Vektorette ligiert ist. Die Positionen der Oligonucleotide 78, 79 und 84 sind angegeben.

c) ein Segment des an eine EcoRI-Vektorette ligierten Phenylalaninhydroxylasegen-Exons 9. Die Positionen der Oligonucleotide 63 und 79 sind angegeben.

d) eine Restriktionskarte des an eine modifizierte EcoRI-Vektorette ligierten Segments genomischer DNA (in 37(a) oben beschrieben). Die Positionen der Oligonucleotide 76, 79, 80 und 81 sind angegeben.

e) eine Restriktionskarte eines an eine modifizierte EcoRI-Vektorette ligierten Segments genomischer DNA (in 37(b) oben beschrieben). Die Positionen der Oligonucleotide 79, 80, 81 und 84 sind angegeben.

f) ein Segment des an eine modifizierte EcoRI-Vektorette ligierten Exons 9 des Phenylalaninhydroxylasegens. Die Positionen der Oligonucleotide 63, 79, 80 und 81 sind angegeben.

g) eine Restriktionskarte eines an eine modifizierte HindIII-Vektorettebibliothek ligierten Segments genomischer DNA. Die Positionen der Oligonucleotide 76, 78 und 79 sind angegeben.

h) eine Restriktionskarte eines an eine modifizierte HindIII-Vektorettebibliothek ligierten Segments genomischer DNA. Die Positionen der Oligonucleotide 78, 79, 81 und 82 sind angegeben.

i) eine Restriktionskarte eines an eine modifizierte HindIII-Vektorettebibliothek ligierten Segments genomischer DNA. Die Po-

sitionen der Oligonucleotide 76, 79, 81 und 83 sind angegeben.

(a) Eine EcoRI-Vektorettbibliothek wurde nach Methode X aus EcoRI-digierter genomischer DNA (Zelllinie 7387) und den Oligonucleotiden 27 und 79 hergestellt. Eine Aliquote der Bibliothek (1 ng) wurde mit den Oligonucleotiden 76 und 79 amplifiziert. Die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 μ l) wurde auf einem 1,4 %igem Agarosegel analysiert. Die Resultate sind in Figur 38 gezeigt:

Spur 1 \emptyset X174 HaeIII-DNA-Marker

Spur 2 870-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 76 und 79 resultiert.

(b) Eine Aliquote (1 ng) der in dem vorhergehenden Beispiel beschriebenen EcoRI-Vektorettbibliothek wurde mit den Oligonucleotiden 79 und 84 amplifiziert. Die Amplifikation erfolgte nach der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 μ l) wurde auf einem 1,4 %igen Agarosegel analysiert. Das Resultat ist in Figur 38 gezeigt.

Spur 3 520-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 79 und 84 resultiert.

(c) Eine Aliquote (1 ng) der in den vorhergehenden Beispielen beschriebenen EcoRI-Vektorettbibliothek wurde mit den Oligonucleotiden 63 und 79 amplifiziert. Die Amplifikation erfolgte gemäß der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 μ l) wurde auf einem 1,4 %igen Agarosegel analysiert. Das Resultat ist in Figur 38 gezeigt.

Spur 4 600-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 63 und 79 resultiert.

(d) Es wurde eine EcoRI-Vektorettbibliothek nach Methode X aus EcoRI-digierter DNA (Zelllinie 7387) und den Oligonucleotiden 80 und 81 hergestellt. Eine Aliquote der Bibliothek (1 ng) wurde mit den Oligonucleotiden 76 und 79 amplifiziert. Die Amplifika-

tion erfolgte entsprechend der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 µl) wurde auf 1,4 %igem Agarosegel analysiert. Die Resultate sind in Figur 38 gezeigt.

Spur 870-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 76 und 79 resultiert.

(e) Eine Aliquote der in (4) oben beschriebenen Bibliothek wurde mit den Oligonucleotiden 79, 84 amplifiziert. Die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches wurde auf einem 1,4 %igen Agarosegel analysiert. Das Resultat ist in Figur 38 gezeigt.

Spur 6 520-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 79 und 84 resultiert.

(f) Eine Aliquote der in (4) oben beschriebenen Bibliothek wurde mit den Oligonucleotiden 63 und 79 amplifiziert. Die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches wurde auf einem 1,4 %igen Agarosegel analysiert. Das Resultat ist in Figur 38 gezeigt.

Spur 7 600-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 63 und 79 resultiert.

(g) Es wurde eine HindIII-Vektorettbibliothek nach Methode X aus HindIII-digierter genomischer DNA (Zelllinie 7387) und den Oligonucleotiden 30 und 81 hergestellt. Eine Aliquote der Bibliothek (1 ng) wurde mit den Oligonucleotiden 76 und 79 amplifiziert. Die Amplifikation erfolgte nach der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 µl) wurde auf einem 1,4 %igen Agarosegel analysiert. Das Resultat ist in Figur 38 gezeigt.

Spur 8 230-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 76 und 79 resultiert.

(h) Es wurde eine HindIII-Vektorettbibliothek nach Methode X aus HindIII-digierter genomischer DNA (Zelllinie 7387) und den

Oligonucleotiden 82 und 81 hergestellt. Eine Aliquote der Bibliothek (1 ng) wurde mit den Oligonucleotiden 76 und 79 amplifiziert. Die Amplifikation erfolgte nach der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 µl) wurde auf einem 1,4 %igen Agarosegel analysiert. Die Resultate sind in Figur 38 gezeigt.

Spur 9 230-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 76 und 79 resultiert.

(i) Es wurde eine HindIII-Vektorettebibliothek nach Methode X aus HindIII-digierter genomischer DNA (Zelllinie 7387) und den Oligonucleotiden 83 und 81 hergestellt. Eine Aliquote der Bibliothek (1 ng) wurde mit den Oligonucleotiden 76 und 79 amplifiziert. Die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Eine Aliquote des Reaktionsgemisches (15 µl) wurde auf einem 1,4 %igen Agarosegel analysiert. Die Resultate sind in Figur 38 gezeigt.

Spur 9 230-bp-Produkt, das aus der Amplifikation zwischen den Oligonucleotiden 76 und 79 resultiert.

Spur 10 ØX174 HaeIII-DNA-Marker.

Beispiel 24

Die folgenden Oligonucleotide wurden synthetisiert:

85 5' ATTCTGTCCAGGAACTTTGTGTTTTGTCA

86 5' AGCCGAAGACAAAGGGCATAATCTC

87 5' TCACTACCAGCTTCCCCACTTCCCTGGC

88 5' CCCTAATTAGTCTCAGCTCCAGGTAAGGT

DNA wurde aus Zelllinie 7382 (von der vorstehend gezeigt wurde, daß sie den ++Haplotyp für den KM19-PstI-Polymorphismus enthält) extrahiert und mit PstI digiert. Nach der Extraktion mit Phenol/Chloroform wurde DNA durch Ethanolpräzipitation rückgewonnen (siehe Figur 39(a)).

284053

151 a

Zirkularisierung

Folgendes wurde in ein 0,5-ml-Sarstedt-Röhrchen gegeben:

10 ng Pst1-digiertes 7382-DNA

10 μ l 10X Ligationspuffer

10 μ l 10 mM ATP

10 0,5 μ l T4-DNA-Ligase (Boehringer Mannheim, 8 Einheiten/ μ l).

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 μ l eingestellt und 2 Stunden bei 25 $^{\circ}$ C inkubiert. Es wurden 5 identische Reaktionen durchgeführt und nach 2 Stunden bei 25 $^{\circ}$ C gepoolt. Das gepoolte Gemisch wurde mit Phenol/Chloroform extrahiert, und DNA wurde durch Ethanolpräzipitation rückgewonnen. Das DNA-Pellet wurde in 100 μ l Wasser gelöst und durch Gelfiltrationen auf vorgepackten Säulen gereinigt (G 50-Sephadex-Nick-Säulen, Pharmacia). Die DNA wurde in 400 μ l Wasser eluiert. Die Lösung wurde lyophilisiert und in Wasser (9 μ l) erneut gelöst, um eine Endkonzentration von 50 ng/ μ l zu erhalten. 1 μ l dieser Lösung wurde mit Wasser auf 50 μ l verdünnt (siehe Figur 39(b)).

Konstruktion der Vektorettbibliothek

HindIII-Digestion

2 μ l	PstI-Kreise (100 ng)
0,5 μ l	10X Puffer
1,5 μ l	H ₂ O
1 μ l	HindIII (2 Einheiten)

Das Reaktionsgemisch wurde 30 Minuten bei 37 $^{\circ}$ C inkubiert.

Ligation

Folgendes wurde dem Restriktionsdigest zugesetzt:

1 μ l	10X Puffer
1 μ l	10 mM ATP
1 μ l	Hybridisierte Oligonucleotide 81 und 82 (0,25 pmol) oder 81 und 83 (0,25 pmol)
1 μ l	Ligase (2 Einheiten)
0,5 μ l	10X HindIII-Puffer

Das Gemisch wurde 30 Minuten bei 37 $^{\circ}$ C inkubiert. Hybridisierte Oligonucleotide (81/82 oder 81/83 0,25 pmol) wurden zugesetzt, und die Reaktion wurde eine weitere Stunde bei 37 $^{\circ}$ C inkubiert. Das Volumen des Reaktionsgemisches wurde dann mit Wasser auf 100 μ l eingestellt (siehe Figur 39(c)).

1. Folgendes wurde in ein 0,5-ml-Sarstedt-Röhrchen gegeben:

PstI-Kreise (5 ng)
Oligonucleotid 85 (100 pmol)
Oligonucleotid 86 (100 pmol)
10 µl 10X PCR-Puffer
100 µM dNTPs
2 Einheiten Taq-Polymerase (Cetus)

Das Volumen wurde mit Wasser auf 100 µl eingestellt, und die Amplifikation wurde entsprechend der Beschreibung in Beispiel 8 durchgeführt. Das vorausgesagte 710-bp-Produkt wurde gewonnen und bestätigte, daß sich Kreise gebildet hatten. Die Resultate sind in Figur 40 gezeigt.

Spur 1 ØX174 HaeIII-DNA-Marker

Spur 2 710-bp-Produkt der Amplifikation zwischen den Oligonucleotiden 85 und 86.

2. Folgendes wurde in ein 0,5-ml-Sarstedt-Röhrchen gegeben:

10 ng PstI-kreis-HindIII-Vektorette-Oligonucleotid 81/82
Oligonucleotid 85 (100 pmol)
Oligonucleotid 79 (100 pmol)
10 µl 10 x PCR-Puffer
100 µM dNTPs
2 Einheiten Taq-Polymerase (Cetus)

Das Volumen wurde mit Wasser auf 100 µl eingestellt, und die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Das vorausgesagte 720-bp-Produkt wurde gewonnen. Die Resultate sind in Figur 40 gezeigt.

Spur 3 720-bp-Produkt der Amplifikation zwischen Oligonucleotid 85 und 79.

3. Das in (2) oben beschriebene Amplifikationsreaktionsgemisch wurde durch einen Faktor von 10^7 verdünnt. 1 µl des verdünnten

Gemisches wurde in ein 0,5-ml-Sarstedt-Röhrchen gebracht.
Folgendes wurde ebenfalls in das Sarstedt-Röhrchen gegeben:

Oligonucleotid 79 (100 pmol)
Oligonucleotid 88 (100 pmol)
10 µl 10X PCR-Puffer
100 µM dNTPs
2 Einheiten Taq-Polymerase (Cetus)

Das Volumen wurde auf 100 µl eingestellt, und die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Das vorausgesagte 500-bp-Produkt wurde gewonnen. Die Resultate sind in Figur 40 gezeigt.

Spur 4 500-bp-Produkt der Amplifikation zwischen den Oligonucleotiden 79, 88.

4. Das Amplifikationsprodukt von (3) oben wurde mit PstI digestiert. Die vorausgesagten Produkte (340 bp, 160 bp) wurden gewonnen. Das Resultat ist in Figur 40 gezeigt.

Spur 5 Von der Digestion des PCR-Produktes (Spur 4) mit PstI resultierende Produkte. Es werden Produkte mit 340 bp und 160 bp gewonnen.

5. Folgendes wurde in ein 0,5-ml-Sarstedt-Röhrchen gegeben:

10 ng PstI-Kreis-HindIII-Vektorette-Oligonucleotid 83/82
Oligonucleotid 85 (100 pmol)
Oligonucleotid 79 (100 pmol)
10 µl 10X PCR-Puffer
100 µM dNTPs
2 Einheiten Taq-Polymerase (Cetus).

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 µl eingestellt, und die Amplifikation erfolgte entsprechend der Beschreibung in Beispiel 8. Das vorausgesagte 720-bp-Produkt wurde gewonnen. Die Resultate sind in Figur 40 gezeigt.

Spur 6 720-bp-Produkt der Amplifikation zwischen Oligonucleotid 79 und 85.

6. Das in (5) oben beschriebene Amplifikationsreaktionsgemisch wurde durch einen Faktor von 10^6 verdünnt. 1 μ l des verdünnten Gemisches wurde in ein 0,5 ml Sarstedt-Röhrchen gebracht. Folgendes wurde ebenfalls in das Sarstedt-Röhrchen gegeben:

Oligonucleotid 79 (100 pmol)
Oligonucleotid 88 (100 pmol)
10 μ l 10X PCR-Puffer
100 μ M dNTPs
2 Einheiten Taq-Polymerase (Cetus)

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 μ l eingestellt, und die Amplifikation wurde entsprechend der Beschreibung in Beispiel 8 durchgeführt. Das vorausgesagte 500-bp-Produkt wurde gewonnen. Die Resultate sind in Figur 40 gezeigt.

Spur 7 500-bp-Produkt der Amplifikation zwischen Oligonucleotid 79, 88.

7. Das Amplifikationsprodukt von (6) oben wurde mit PstI digestiert. Die vorausgesagten Produkte (340 bp, 160 bp) wurden gewonnen. Die Resultate sind in Figur 40 gezeigt.

Spur 8 Produkte, die von der Digestion von PCR-Produkt (Spur 6) mit PstI resultieren. Produkte von 340 bp und 160 bp werden gewonnen.

8. Das in (3) oben beschriebene Amplifikationsreaktionsgemisch wurde durch einen Faktor von 10^7 verdünnt. 1 μ l des verdünnten Gemisches wurde in ein 0,5-ml-Sarstedt-Röhrchen gegeben. Folgendes wurde dann in das Sarstedt-Röhrchen gebracht:

Oligonucleotid 79
Oligonucleotid 85

- 10 µl 10X PCR-Puffer
- 100 µM dNTPs
- 5 2 Einheiten Taq-Polymerase (Cetus)

Das Volumen wurde mit sterilem, doppelt destilliertem Wasser auf 100 µl eingestellt, und die Amplifikation wurde entsprechend der Beschreibung in Beispiel 1 durchgeführt. Das vorausgesagte 720-bp-Produkt wurde gewonnen. Die Resultate sind in Figur 40 gezeigt.

Spur 9 720-bp-Produkt der Amplifikation zwischen Oligonucleotid 79, 85.

15 9. Das Amplifikationsprodukt von (8) oben wurde mit PstI digeriert. Die vorausgesagten Produkte (370 bp, 350 bp) wurden gewonnen.

Spur 10 Produkte, die von der Digestion von PCR-Produkt (Spur 9) mit PstI resultieren.

20 Produkte von 370 bp und 350 bp wurden gewonnen.

Spur 11 ØX174 HaeIII-DNA-Marker.

25

30

35

39

Fig. 1.

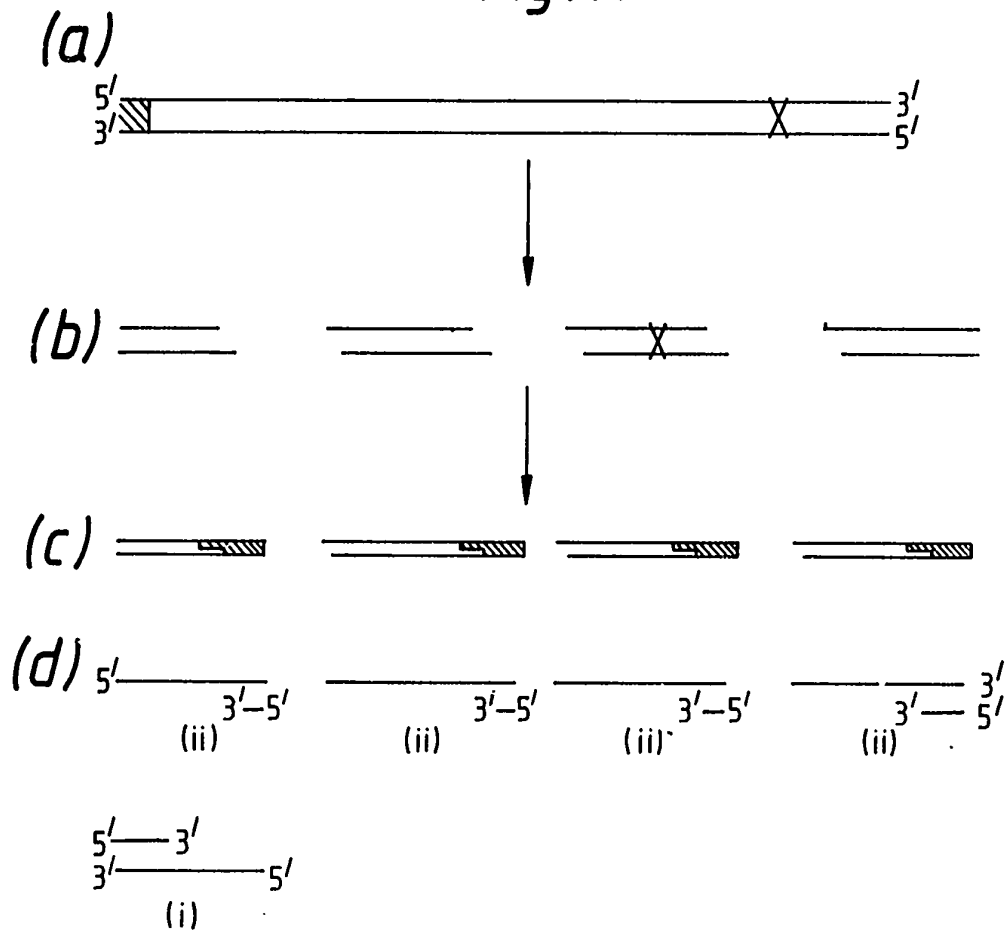


Fig. 2.

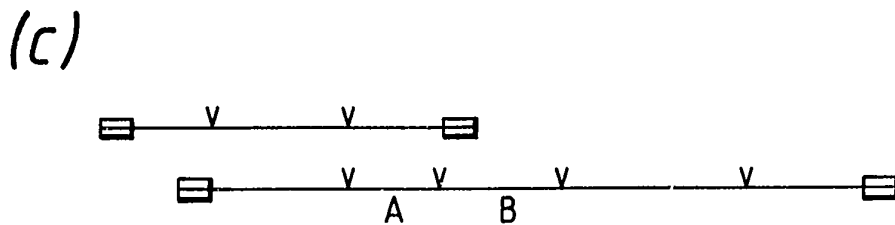
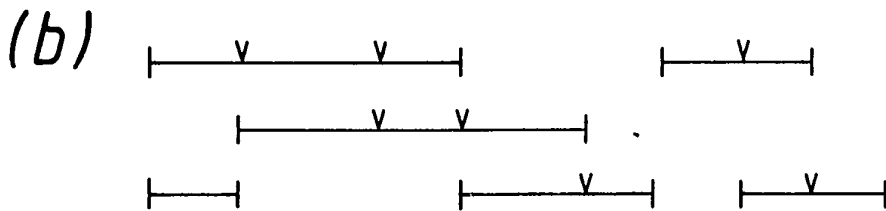
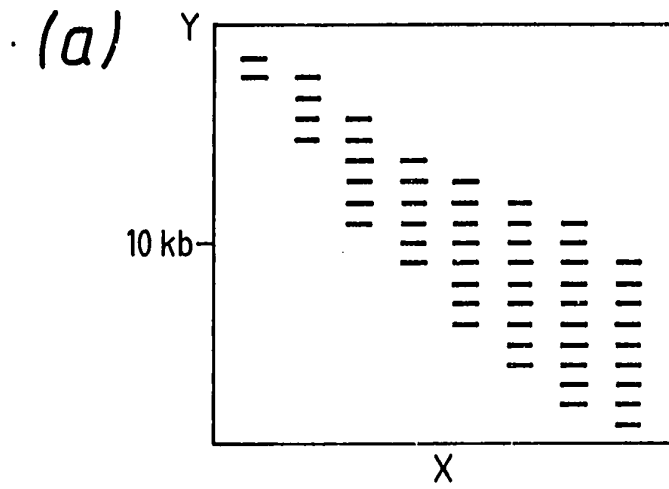
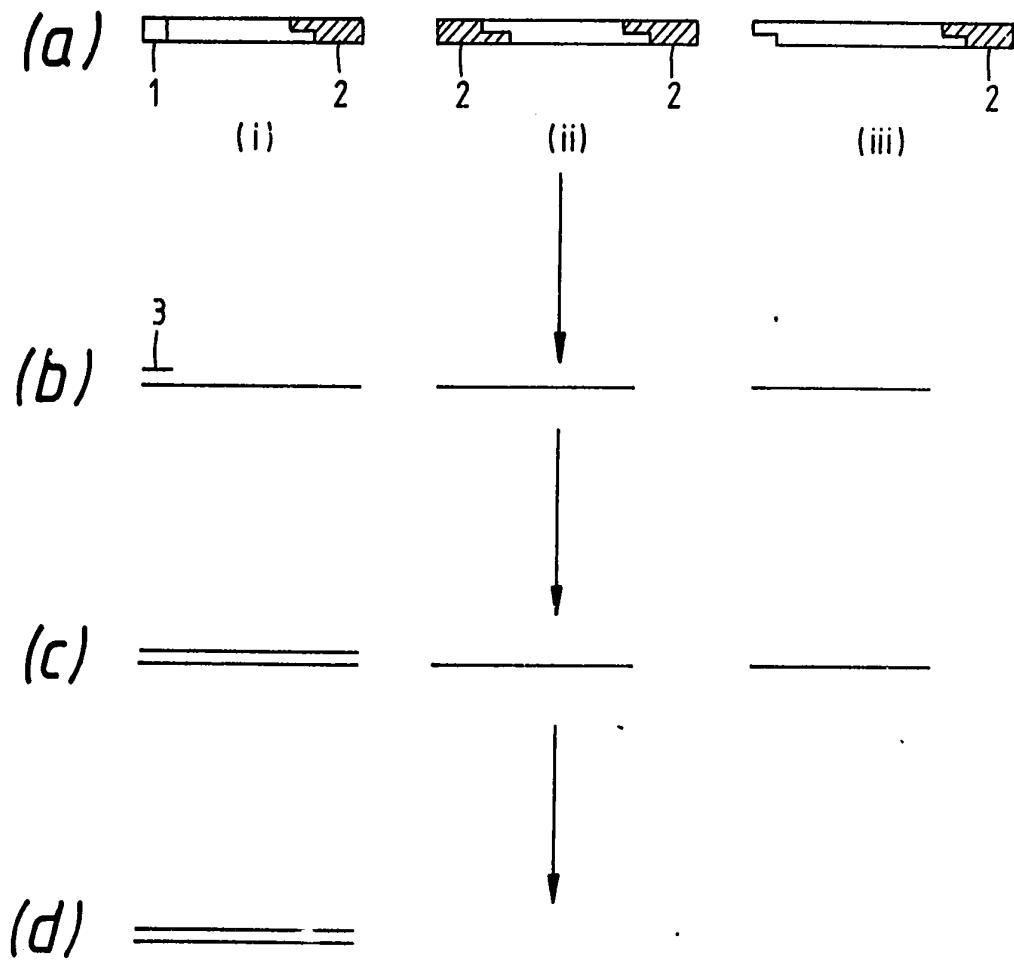
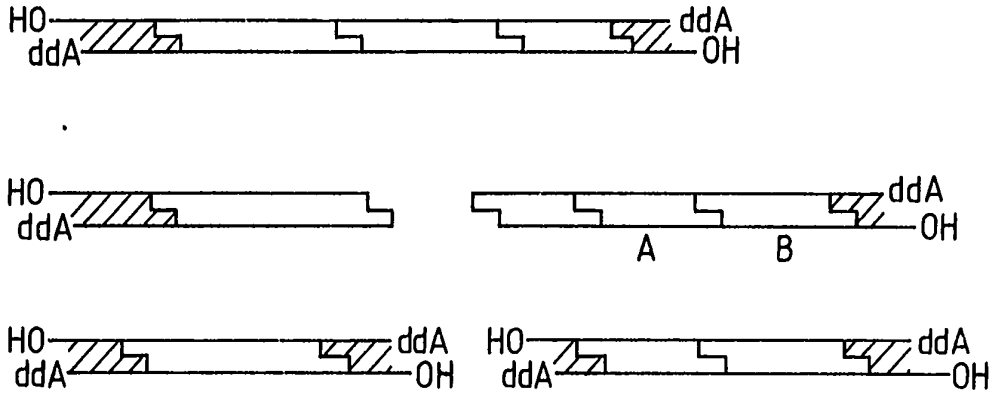


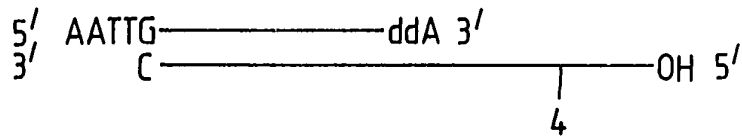
Fig. 3.



(a) Fig. 4.



(b)



(c)

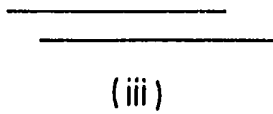
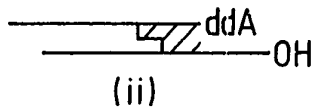
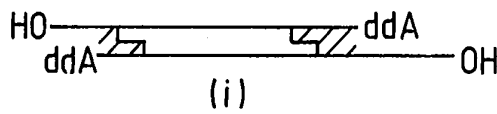


Fig. 5.

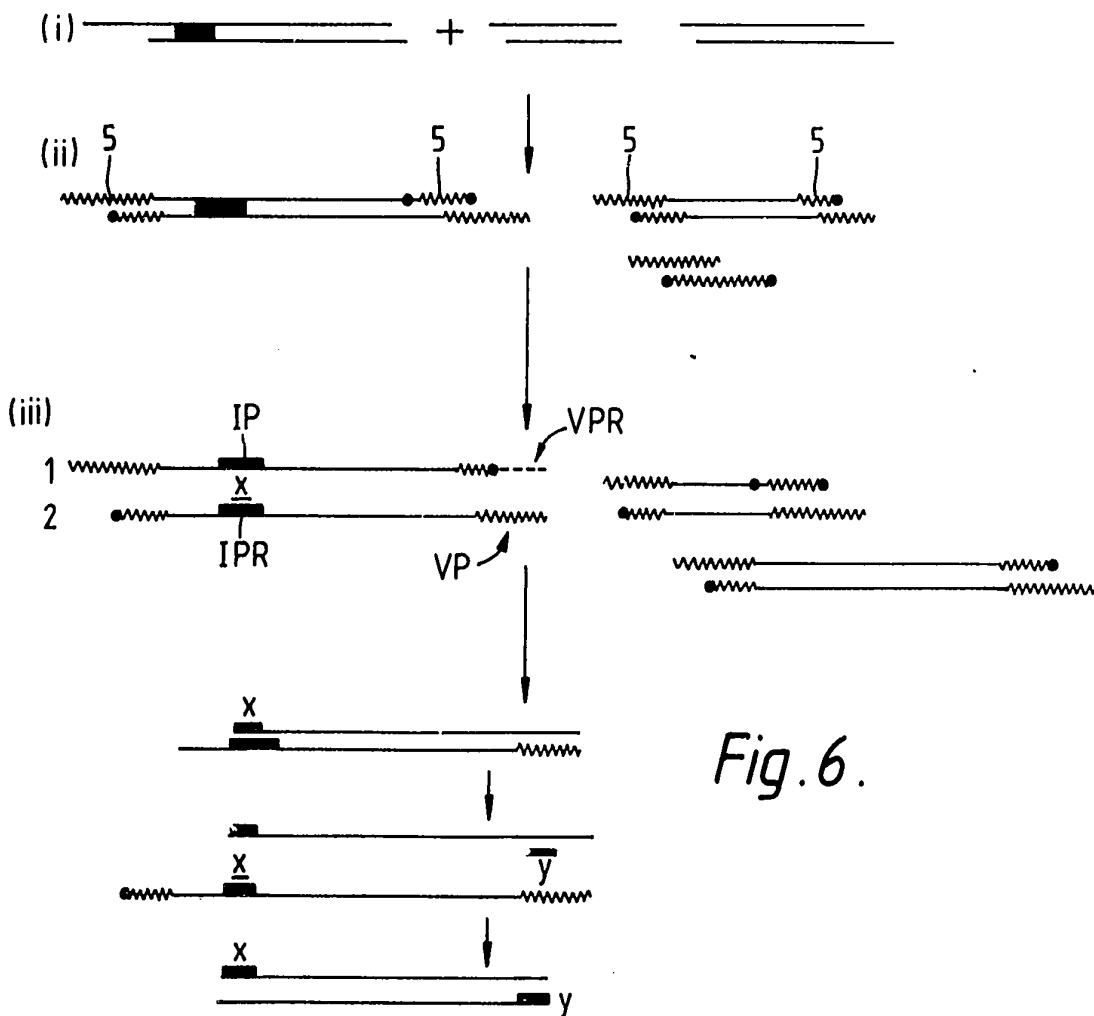
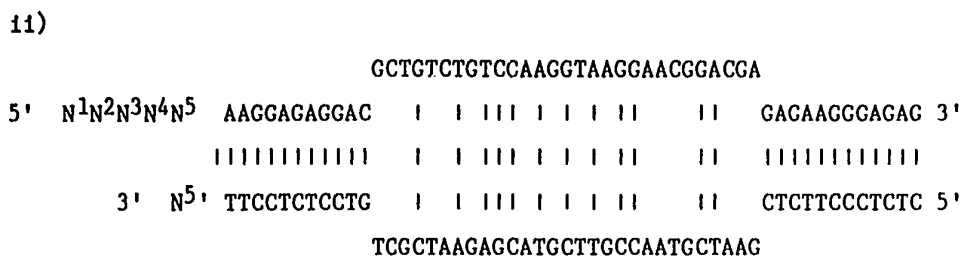
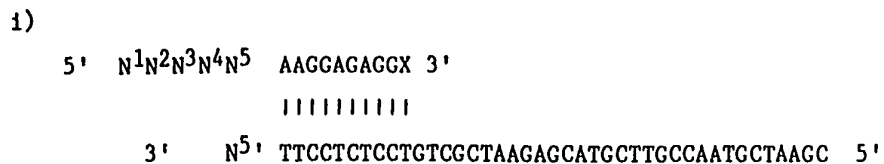
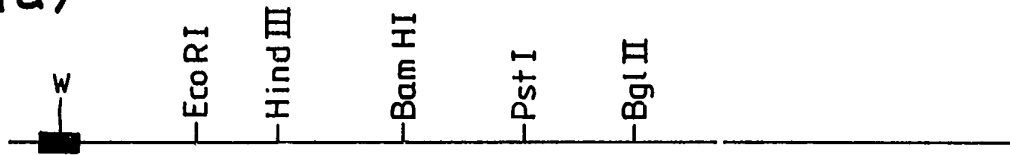


Fig. 6.

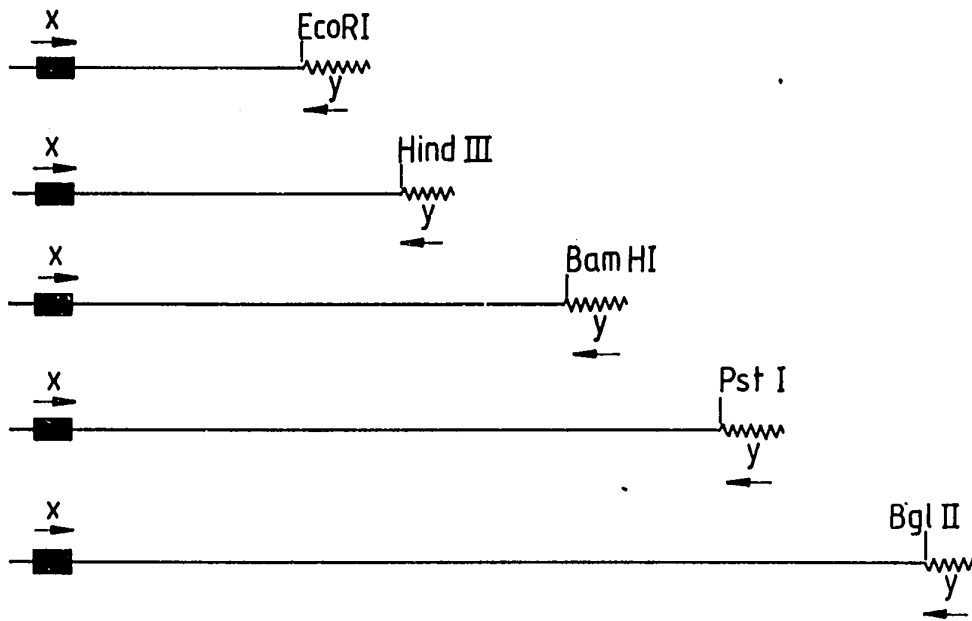
28 FEB 1980 07:00:07

Fig. 7.

(a)



(b)



(c)

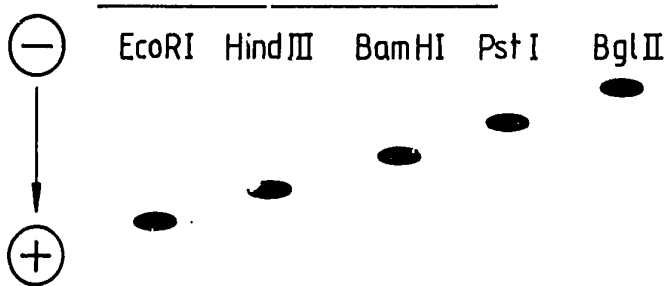


Fig. 8.

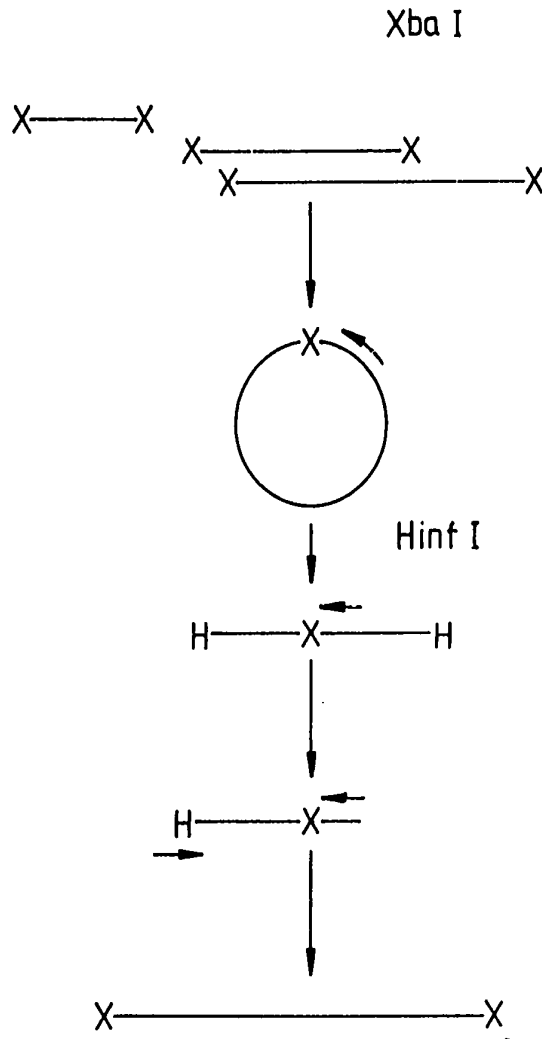
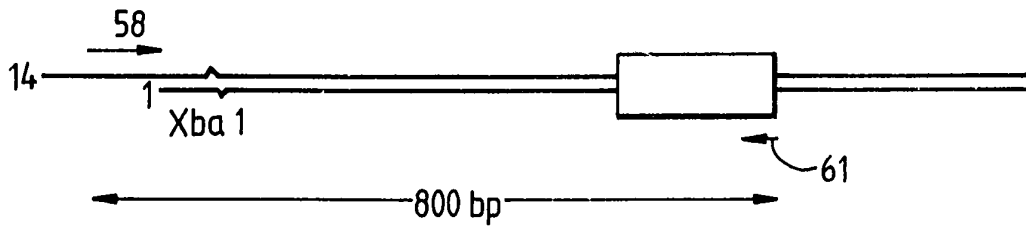


Fig. 9.



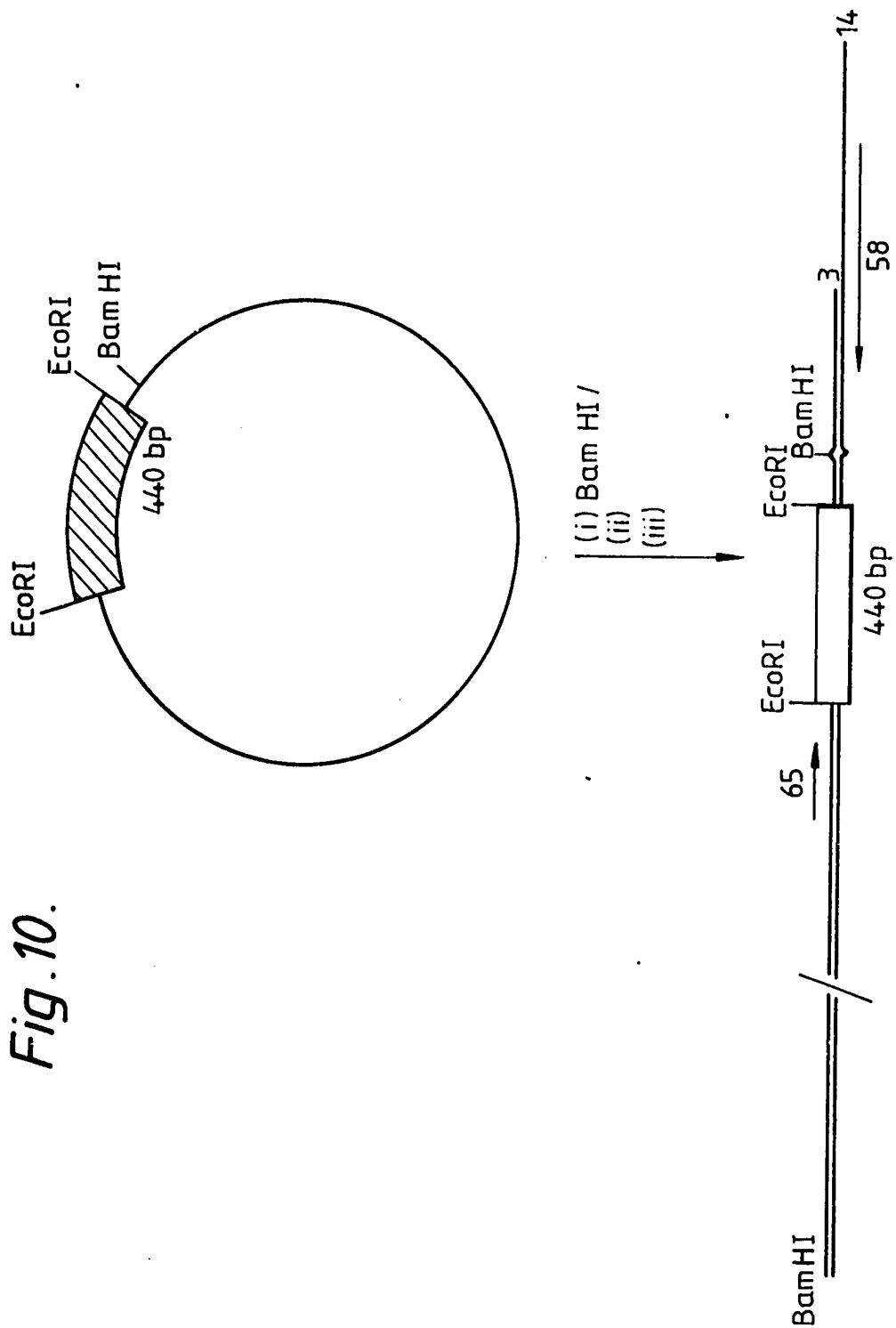
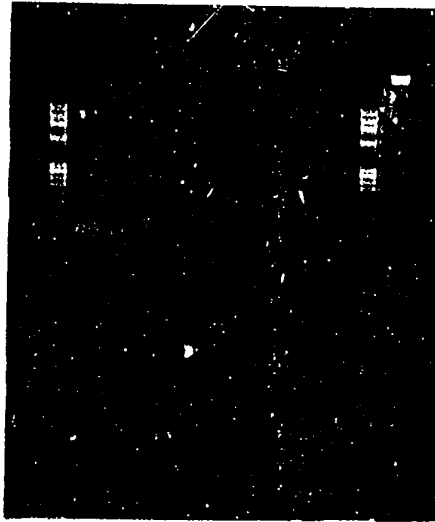


Fig. 10.

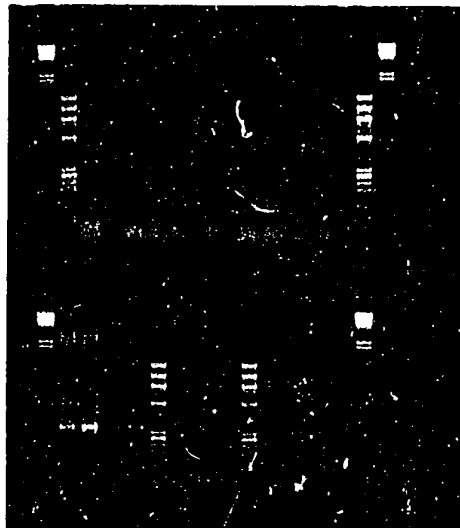
Fig. 11.

(i)



1 2 3 4 5 6 7 8 9 10 11 12

(ii)



1 2 3

10 12 13

(iii)

1 2 3 4 5 6 6 7 8 9 11
10

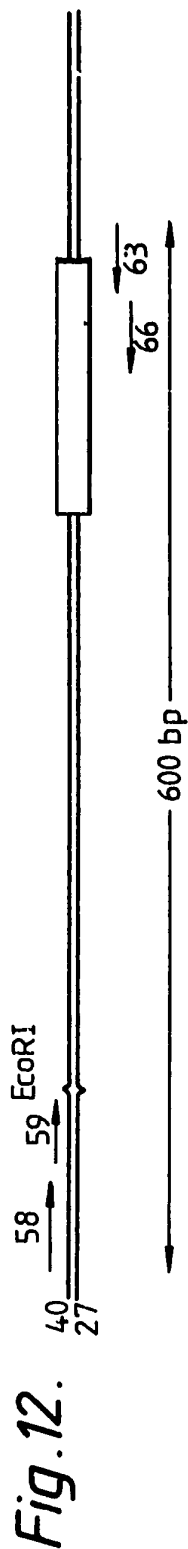


Fig. 22.

10	20	30	40	50
5' TGTAAATGGG	GGAAAATAGA	ACCTGTTCTG	TTCCTGTAAT	TGGAACCACA
60	70	80	90	100
GAACCAACCT	ACTACTTGGC	CATAAAAAGG	TGATGGGTGG	CCAGATGCCT
110	120	130	140	150
TCAGAACAGC	CCACATAGAC	CCTGAGTGTG	TTATCAAGTC	TTTTCTGCCC
160	170	180	190	200
ATATCTTATA	GAATCACAGA	TCTCCAGGGT	AGGGGGATCC	TTCCACCCCTC
210	220	230	240	250
ACTTGGCCCT	TTGTTTTTTCAC	TTTATAAGTG	AGGAAATTGA	GGCACAGAGA
260	270	280	290	300
AGGSAAGTGA	CTTGCTCATG	ATCACAGGGT	CGGTTAGTAT	CTAAAGCTTC
310	320	330	340	350
TTCTCTTGAC	TCCAGCTCAG	TGATTATTCA	TCAGCAGAAC	ATTCTTGGAA
360	370	380	390	400
TGTTAAATGA	GTCCTTAATCA	TGTACCCCTGG	AAACATCACA	GGAGAACTCC
410	420	430	440	450
AGATAAGTGG	TGACATCGGT	CCAGGCATAT	ATATAGCTGG	AGAGTCTAAC

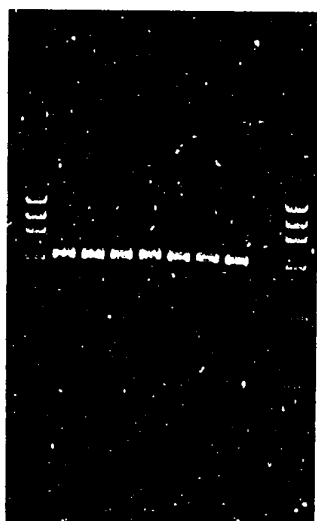
28 FEB 1980 030007

Fig. 13.



1 2

Fig. 14.

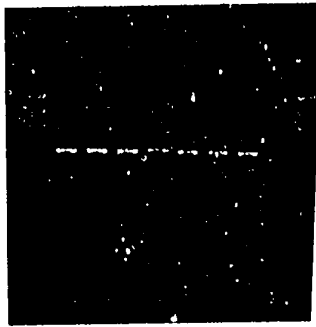


1 2 3 4 5 6 7 8 9



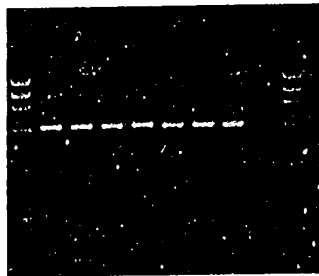
10 11

Fig. 15.



1 2 3 4 5 6 7 8 9 10

Fig. 16.



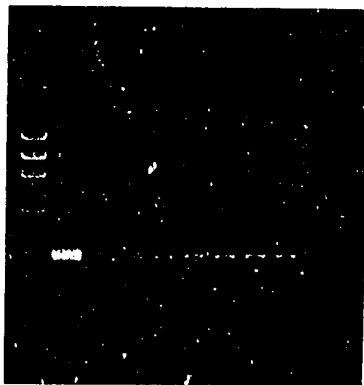
1 2 3 4 5 6 7 8 9 10

Fig. 17.



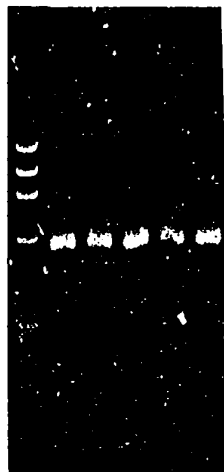
1 2 3 4 5 6 7 8 9 10

Fig. 18.



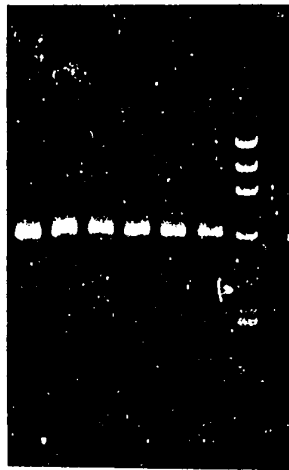
1 2 3 4 5 6 7 8 9 10

Fig. 19.



1 2 3 4 5 6

Fig. 20.



1 2 3 4 5 6 7

Fig. 21.

CGTA	CGTA	CGTA	CGTA	CGTA	CGTA	CGTA
1	2	3	4	5	6	7

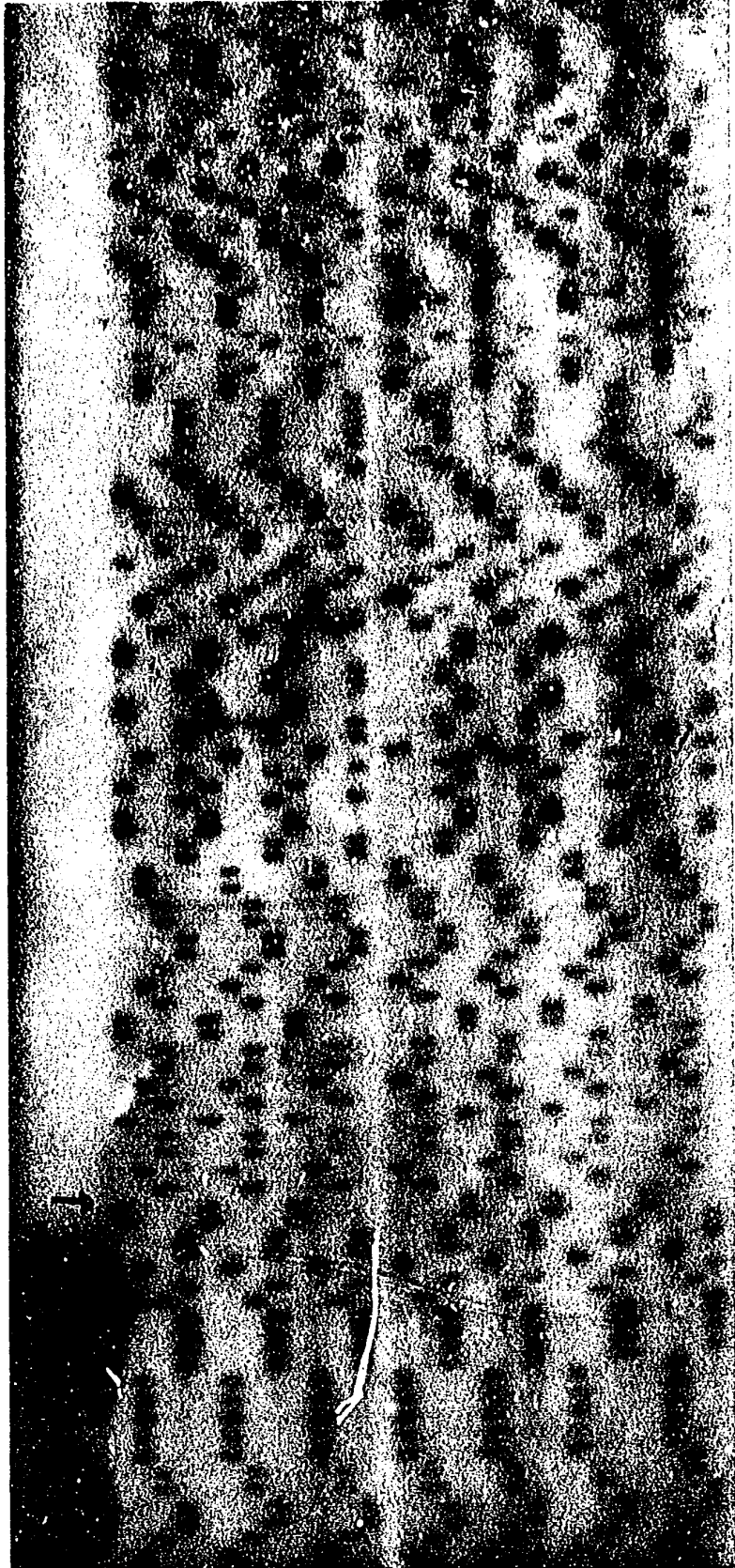


Fig. 23.

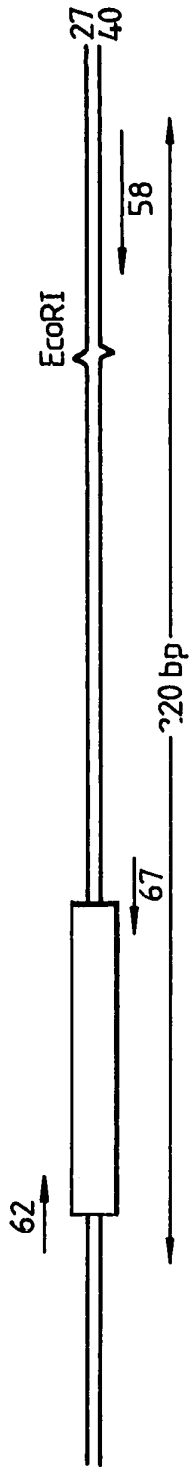


Fig. 25.

- a) CTG GAA AAC CCA GGC TTG GGC AGG AAA CTC TCT GAC TTT GGA CAG GTG AGC
 CAC GGC ACC CTG AGC TGT CAG TTA GGG GAA TTT GGG CCT CCA GAG AAA GAG ATC
 CGA AGA CTG CTG CTG CTT CTG GTT TAA TAA
- b) GGG TTT TCC AGG ACC GCA GTG GAC ATG CTG GCT CCC CGG GAG TGA GGT CTC
 TGG CTT T
- c) AAA GCCAG AGACC TCACT20 CCCGG GCAGC CAGCA TGTC40 ACTGC

 GCTCC TGGAA AACCC60 AGGCT TGGGC AGSAA ACTCT80 CTGAC TTTGG

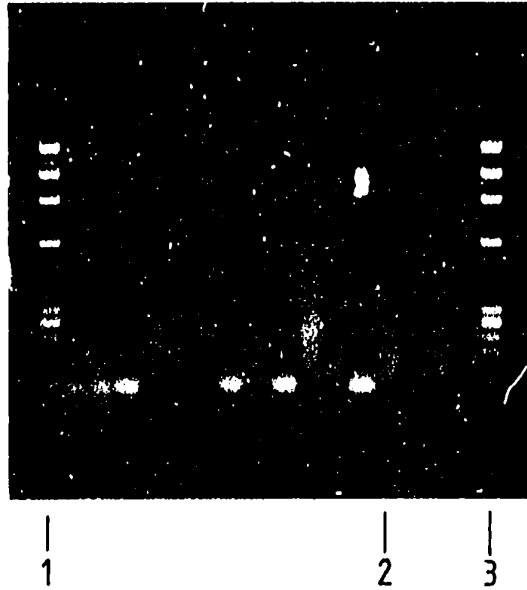
ACAGG TGACC100 CACGG CACCC TGAGC TGTCAL20 GTTAG GGGAA TTTGG

 GCCTC140 CACAG AAAGA GATCC GAAGA160 CTGCT GGTCG TTCTG

 GTTTA180 ATAA ----- etc.

Fig. 24.

(a)



(b)

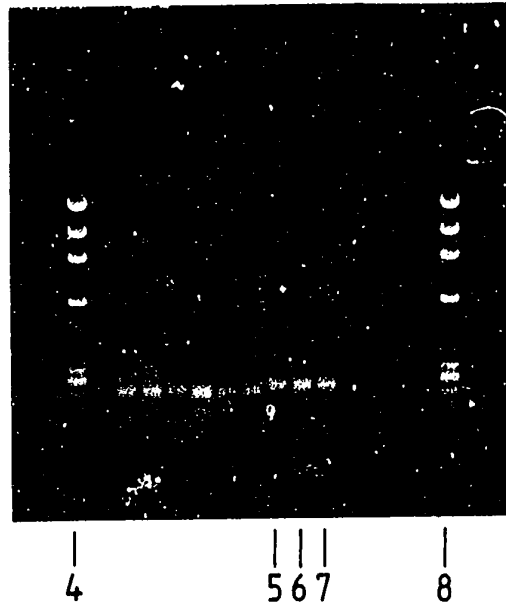


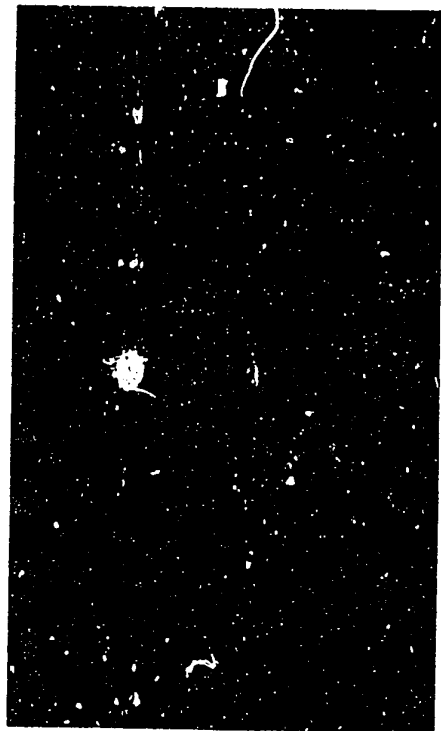
Fig. 27.

(a)



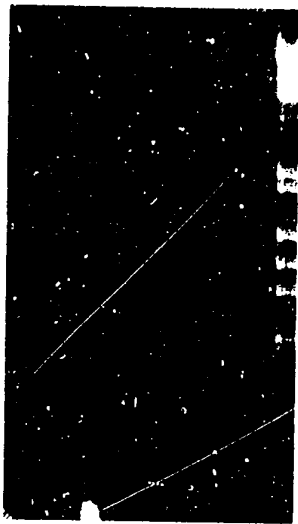
3
4

(c)



1 2 3

(b)



1 2 3 4

Fig. 28.

1 ATGCCTGCAG GTCGACTCTA GAGGATTCCC CTAGAGCATA T₁AAAT¹TATT

51 TTCAAGGGAA GATGTAAAAA TAGGTATGAA GAAGTTCTGG TACTTTTTTC

101 CCCACCCAGC AGATCACTGT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

151 TATCACTTGA GTGTTATGCA CTGCTCTTTA AAGATTCCCTG CTCTCAGGAG

201 CACCTAAGGG TAGAATAGGA ACTGGCACTT CCATGCCCAG GACATGCCTT

251 CTGGAGGTGC TTATGGAGAT GGA¹CTGGAGT CTGGGTCTTG GAAAGTAGGC

301 ATGTTTGATG GGCCAGTTTT ACATTAAGTT CTGTAGTCTC ATTGCATCTG

351 GCACCCCTCAA ATTTCTTATT AAATACTACA GCCTGGTTTT TAGCTTTGTT

401 CCATAGTAGG AAAGTTTGA GCATAGGAAA AGTTCTGTGC CAAAAATTGC

71

451 ATCCAAGAAT TGCCTACCGA GCTC 72

Fig. 33.

1 AGATGCTGTA CACAFAAAAG GACAAGCTGA ACTCCTGCAA CTCACTTAAC

51 GAGAACTATT ^{HinfI}AGAATCGGCA ^{HinfI}CATTTGAACA GACAAGTCGC TGAATCCAAG

101 ATGAAGCCTT TTATTTTTAA GAAAAAATG GCCTTTACAT CACCGACAAA

151 TTAGGTTACG ATGAAAAAGG CTGTTTC

Fig. 29.

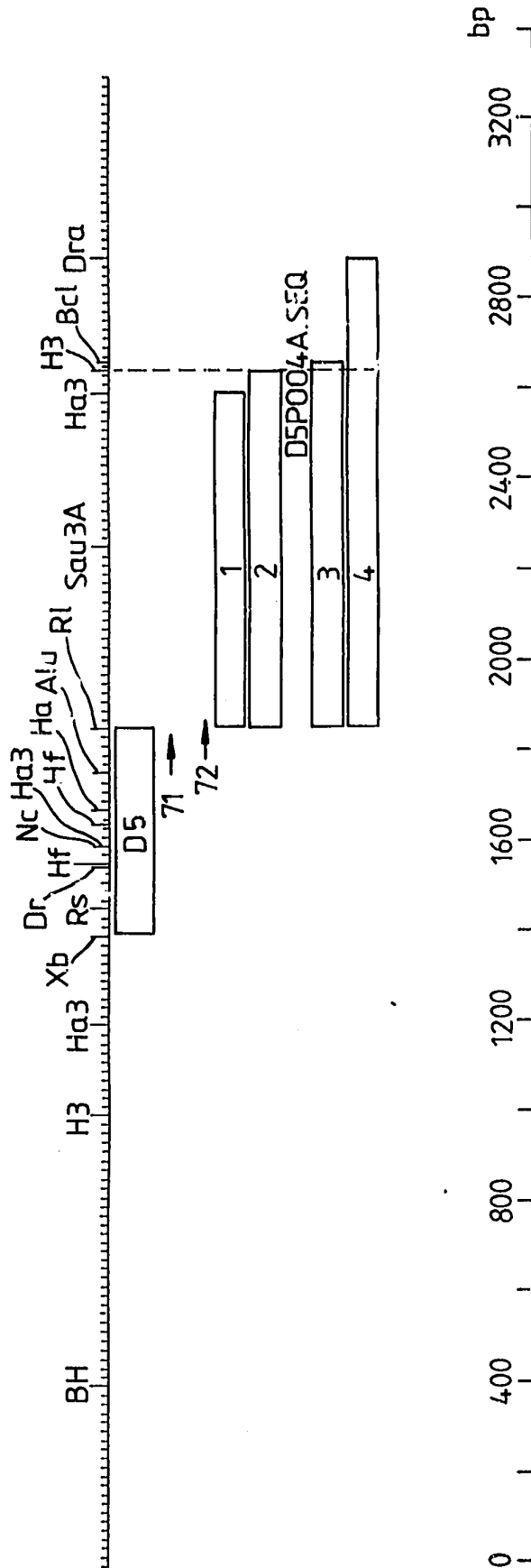


Fig. 30.



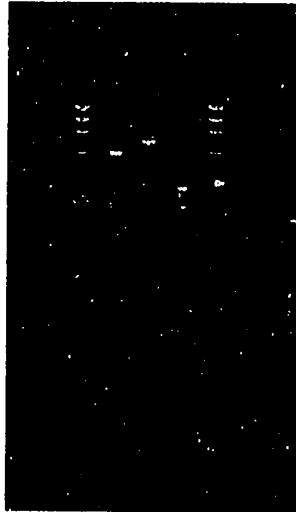
||| | | | | | |
9 8 7 6 5 4 3 2 1

Fig. 32.



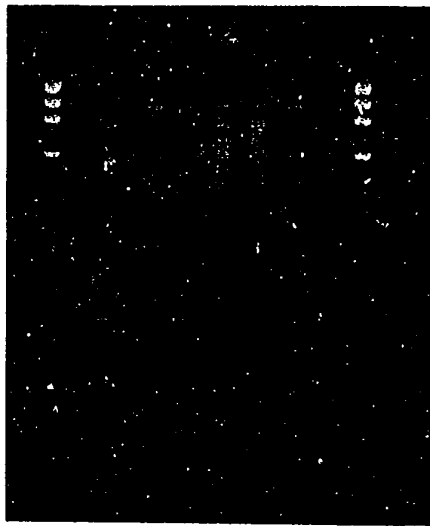
|||
1 2 3

Fig. 34.



| | | | |
1 2 3 4 5

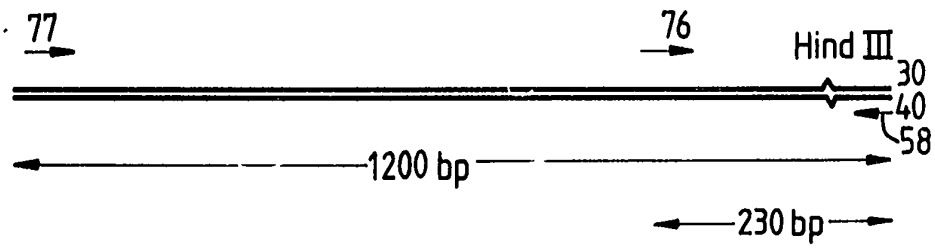
Fig. 35.



| | | | |
1 2 3 4 5

Fig. 36.

(a)



(b)

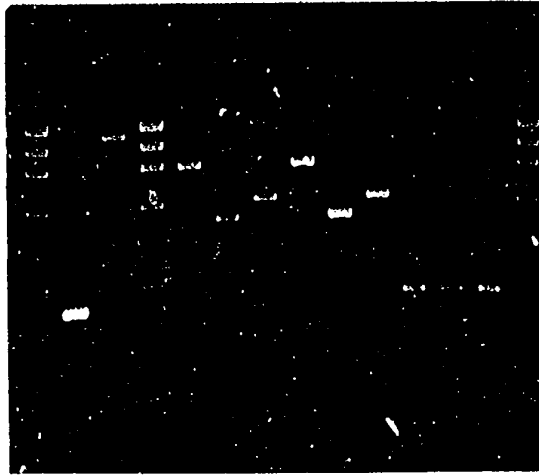


Fig. 37.

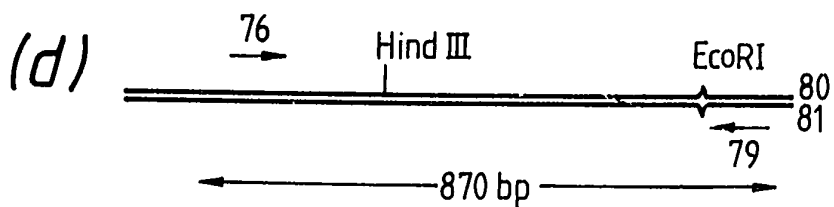
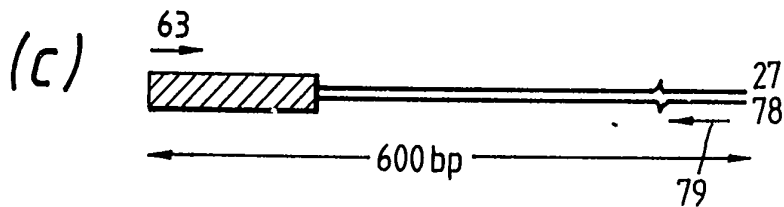
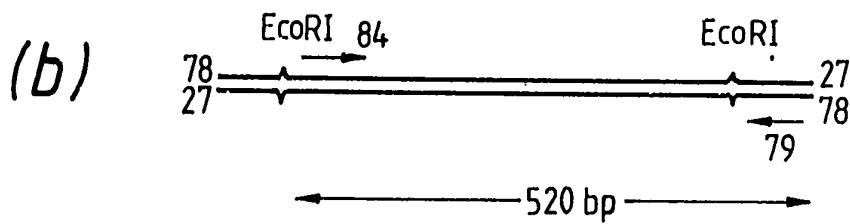
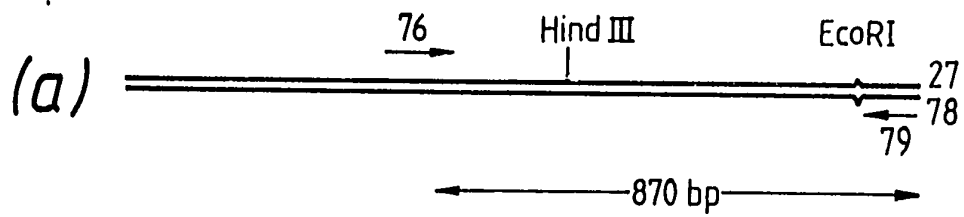


Fig. 37.

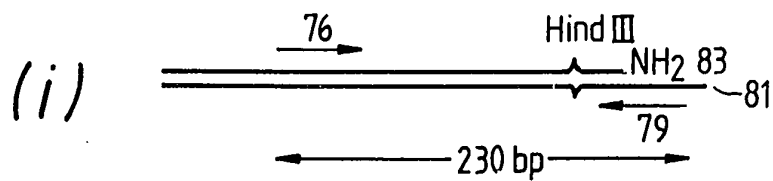
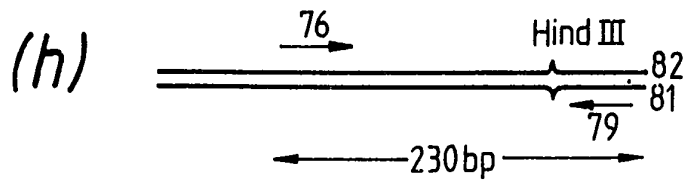
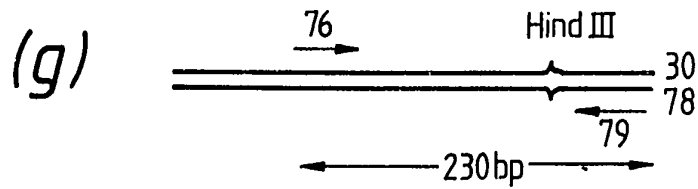
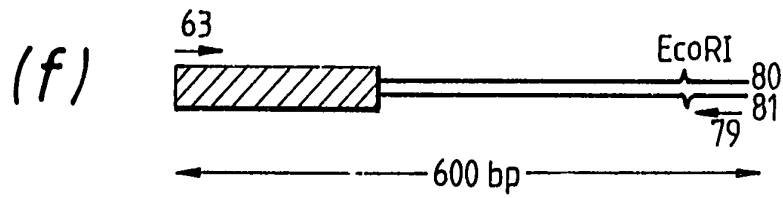
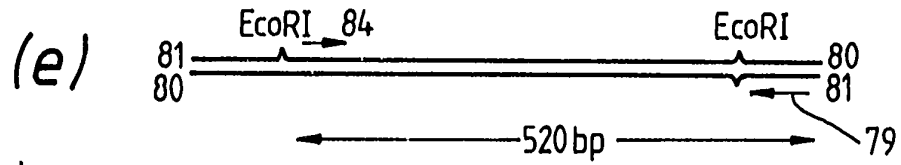
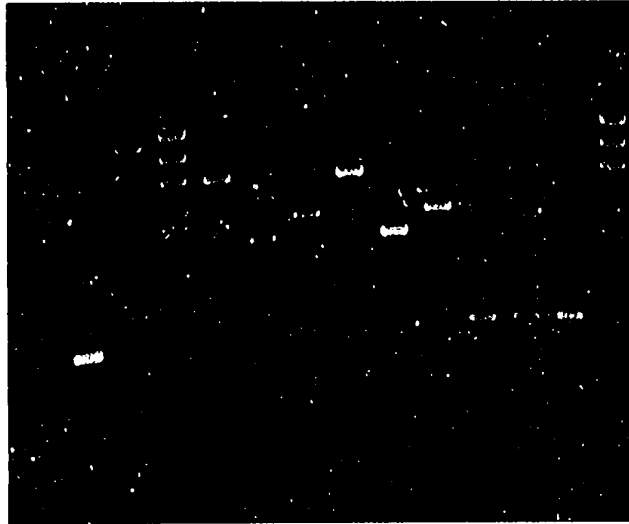
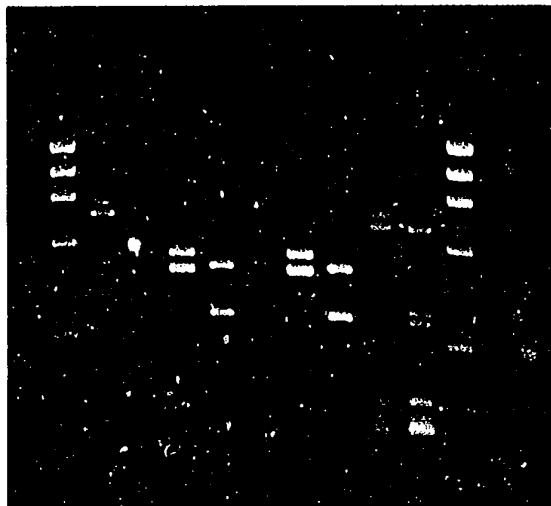


Fig. 38.



1 2 3 4 5 6 7 8 9 10 11

Fig. 40.



1 2 3 4 5 6 7 8 9 10 11

Fig. 39.

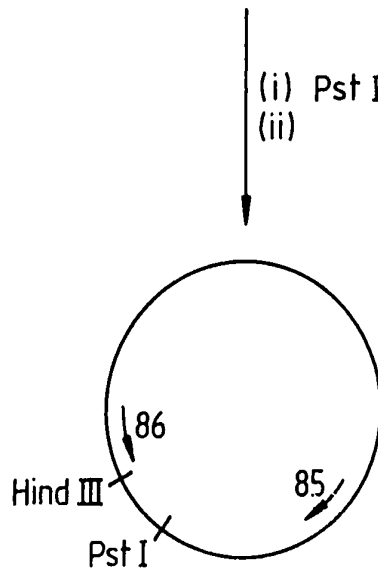
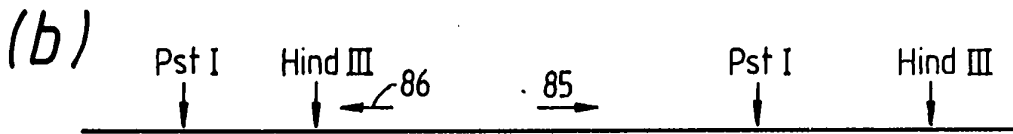
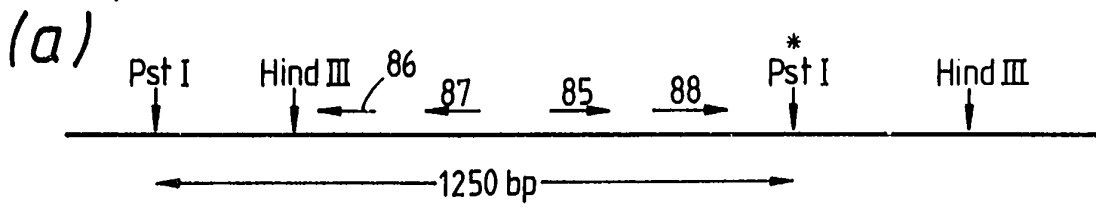


Fig. 39.

(c)

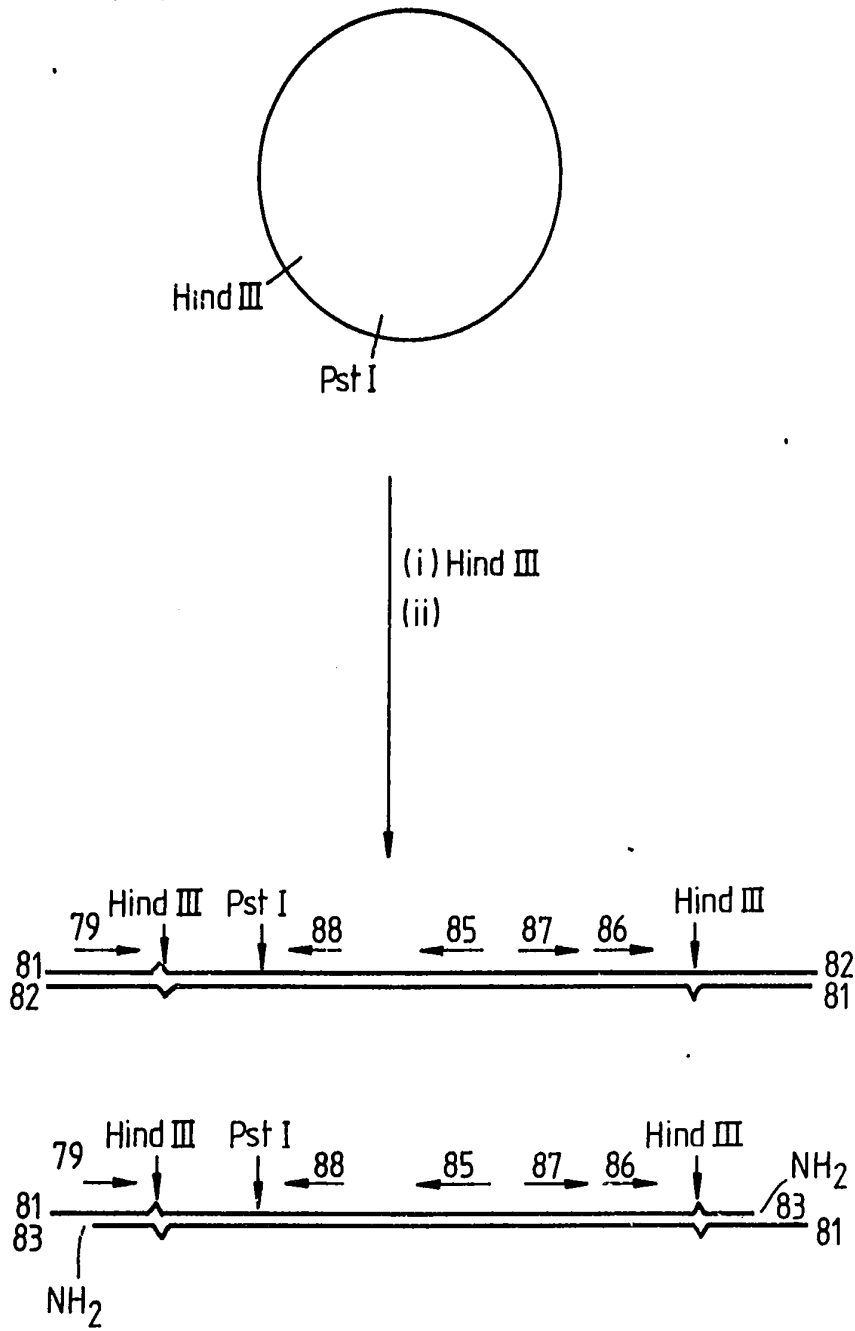
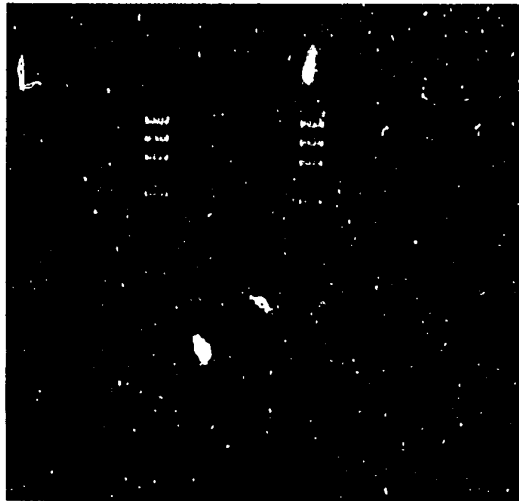


Fig.41.



| | | |
1 2 3 4

Fig.42.



| | | | | | | | |
1 2 3 4 5 6 7 8 9