



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 695 32 725 T2** 2005.01.27

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 871 715 B1**

(21) Deutsches Aktenzeichen: **695 32 725.9**

(86) PCT-Aktenzeichen: **PCT/US95/06381**

(96) Europäisches Aktenzeichen: **95 921 315.8**

(87) PCT-Veröffentlichungs-Nr.: **WO 95/32281**

(86) PCT-Anmeldetag: **15.05.1995**

(87) Veröffentlichungstag
der PCT-Anmeldung: **30.11.1995**

(97) Erstveröffentlichung durch das EPA: **21.10.1998**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **17.03.2004**

(47) Veröffentlichungstag im Patentblatt: **27.01.2005**

(51) Int Cl.⁷: **C12N 9/22**
C12N 15/55, C12N 15/70

(30) Unionspriorität:

247990	24.05.1994	US
356898	15.12.1994	US

(73) Patentinhaber:

New England Biolabs, Inc., Beverly, Mass., US

(74) Vertreter:

**Kuhnen & Wacker Patent- und
Rechtsanwaltsbüro, 85354 Freising**

(84) Benannte Vertragsstaaten:
DE, FR, GB

(72) Erfinder:

**FOMENKOV, Aleksei, Baltimore, US; DILA,
Deborah Kay, Boston, US; RALEIGH, Elisabeth A.,
Somerville, US; XU, Shuang-yong, Lexington, US**

(54) Bezeichnung: **METHODE ZUR DIREKTER KLONIERUNG VON NUKLEASEGENEN IN E.COLI**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung**HINTERGRUND DER ERFINDUNG**

[0001] Nucleasen sind eine Klasse von Enzymen, die einzel- oder doppelsträngige DNA zerlegen oder zerschneiden. Restriktionsendonucleasen sind eine wichtige Klasse von Nucleasen, die spezielle Sequenzen von Nucleotiden (die „Erkennungssequenz“) entlang dem DNA-Molekül erkennen und sich an sie binden. Nach dem Anbinden spalten sie beide Stränge des Moleküls in oder an einer Seite der Erkennungssequenz. Unterschiedliche Restriktionsendonucleasen erkennen unterschiedliche Erkennungssequenzen. Mehr als einhundertundachtzig Restriktionsendonucleasen mit einmaligen Spezifitäten wurden unter den vielen hundert Bakterienspezies identifiziert, die bisher untersucht wurden.

[0002] Man geht davon aus, dass Restriktionsendonucleasen in der Natur eine schützende Rolle zum Wohle der Bakterienzelle spielen. Sie machen Bakterien widerstandsfähig gegen Infektionen durch Fremd-DNA-Moleküle wie Viren und Plasmide, die sie ansonsten zerstören oder befallen würden. Sie übertragen Widerstandsfähigkeit, indem sie eindringende Fremd-DNA-Moleküle spalten, wenn die passende Erkennungssequenz vorliegt. Durch die stattfindende Spaltung werden viele der infizierenden Gene deaktiviert und die DNA wird für einen weiteren Abbau durch unspezifische Endonucleasen empfänglich gemacht.

[0003] Eine zweite Komponente dieser bakteriellen Schutzsysteme sind die Modifikationsmethylasen. Diese Enzyme sind komplementär zu den Restriktionsendonucleasen und liefern das Mittel, das Bakterien dazu befähigt, ihre eigene DNA vor Spaltung zu schützen und sie von fremder, infizierender DNA zu unterscheiden. Modifikationsmethylasen erkennen und binden sich an die gleiche Nucleotiderkennungssequenz wie die entsprechende Restriktionsendonuclease; anstatt aber die DNA zu spalten, modifizieren sie chemisch das eine oder andere Nucleotid innerhalb der Sequenz durch Hinzufügen einer Methylgruppe. Im Anschluss an die Methylierung wird die Erkennungssequenz nicht mehr durch die Restriktionsendonuclease gebunden oder gespalten. Die DNA einer Bakterienzelle wird aufgrund der Aktivität der Modifikationsmethylase stets völlig modifiziert. Sie ist daher gegenüber der Anwesenheit der endogenen Restriktionsendonuclease völlig unempfindlich. Lediglich unmodifizierte und daher identifizierbare Fremd-DNA ist gegenüber einer Erkennung und Spaltung durch die Restriktionsendonuclease empfindlich.

[0004] Seit dem Beginn der Gentechnologie ist es nun möglich, Gene zu klonieren und die Proteine und Enzyme, die sie kodieren, in größeren Mengen zu produzieren, als es mit konventionellen Reinigungstechniken möglich ist. Der Schlüssel zur Isolierung von Klonen von Restriktionsendonucleasegenen ist die Entwicklung eines einfachen und zuverlässigen Verfahrens zur Identifizierung solcher Klone innerhalb komplexer "Bibliotheken", d.h. Populationen von Klonen, die durch "Shotgun"-Verfahren abgeleitet werden, wenn sie in Frequenzen von nur 10^{-3} bis 10^{-4} auftreten.

[0005] Immer öfter werden Restriktionsmodifikationssysteme des Typs II kloniert. Die ersten klonierten Systeme benutzten eine Bakteriophage-Infektion als Mittel zum Identifizieren oder Selektieren von Restriktionsendonucleaseklonen (EcoRII: Kosykh et al., Molec. Gen. Genet 178: 717–719, (1980); HhaI: Mann et al., Gene 3: 97–112, (1978); PstI: Walder et al., Proc. Nat. Acad. Sci. 78:1503–1507, (1981)). Da die Anwesenheit von Restriktionsmodifikationssystemen in Bakterien, diese dazu befähigt, Infektionen durch Bakteriophagen standhalten zu können, können Zellen, die klonierte Restriktionsmodifikationsgene tragen, prinzipiell als Überlebende von Bibliotheken selektiv isoliert werden, die dem Phage ausgesetzt wurden. Man hat jedoch festgestellt, dass dieses Verfahren nur von begrenztem Nutzen ist. Im Speziellen wurde gefunden, dass klonierte Restriktionsmodifikationsgene nicht immer ausreichend Phagenresistenz aufweisen, um selektives Überleben zu gewährleisten.

[0006] Ein weiterer Klonierungsansatz schließt das Übertragen von anfänglich als plasmidbärtig gekennzeichneten Systemen in E. coli Klonierungsplasmide ein (EcoRV: Bougueler et al., Nucl. Acid. Res. 12: 3659–3676, (1984); PaeR7: Gingeras und Brooks, Proc. Natl. Acad. Sci. USA 80: 402–406, (1983); Theriault und Roy, Gene 19: 355–359 (1982); PvuII: Blumenthal et al., J. Bacteriol. 164: 501–509, (1985)).

[0007] Ein dritter Ansatz, der zum Klonieren einer zunehmenden Zahl von Systemen angewendet wird, werden jetzt durch Selektion nach einem aktiven Methylasegen kloniert [sic]. Siehe z.B. US-Patent Nr. 5,200,333 und BsuRI: Kiss et al., Nucl. Acid. Res. 13: 6403–6421, (1985)). Da Restriktions- und Modifikationsgene oft eng miteinander gekoppelt sind, können beide Gene oft simultan kloniert werden. Diese Selektion bringt jedoch nicht immer ein komplettes Restriktionssystem hervor, sondern stattdessen nur das Methylasegen (BspRI: Szomolanyi et al., Gene 10: 219–225, (1980); BcnI: Janulaitis et al., Gene 20: 197–204 (1982); BsuRI: Kiss

und Baldauf, Gene 21: 111–119, (1983); und MspI: Walder et al., J. Biol. Chem. 258: 1235–1241, (1983)).

[0008] Ein viertes Klonierungsverfahren (das „Methylase-Indikator“-Verfahren) ist auf methylierungsabhängige Restriktionssysteme McrA, McrBC und Mrr (Raleigh und Wilson, Proc. Natl. Acad. Sci., USA 83:9070–9074, (1986), Heitman und Model, Gene 103:1–9; Kelleher und Raleigh, J. Bacteriol. 173:5220–5223, (1991)) und die *dinD1::lacZ* Operon-Fusion angewiesen, um nach Klonen zu screenen, die Methylasegene enthalten. Der *dinD1* Ort ist ein DNA-Schaden induzierbares Gen, das in *E. coli* exprimiert wird, wenn die SOS-Reaktion ausgelöst wird, wie z.B. durch UV-Behandlung, Mitomycin-Behandlung oder die Wirkung von McrA-, McrBC- oder Mrr-Restriktionsendonucleasen auf methylierte DNA (Kenyon und Walker, Proc. Natl. Acad. Sci. USA, 77:2819–2823, (1980), Heitman und Model, Gene, 103:1–9, (1991); Heitman und Model, J. Bacteriol. 169:3234–3250, (1989); und Piekarowicz et al. J. Bacteriol. 173:150–155 (1991).

[0009] Stämme mit temperaturempfindlichen Mutationen in *mcrA*, *mcrBC* und *mrr*, die die *dinD1::lacZ* Fusion tragen, wurden konstruiert und zur direkten Klonierung von Methylasegenen in *E. coli* von anderen Bakterienquellen (Piekarowicz et al., Nucleic Acids Res. 19:1831–1835 (1991) verwendet.

[0010] Nach der Transformation von ligierter Genom-/Vektor-DNA in einen solchen Stamm bilden Transformanten, die ein Gen enthalten, das eine Methylase exprimiert, die Empfindlichkeit gegen eines der methylierungsabhängigen Restriktionssysteme verleiht, weiße Kolonien bei 42°C und blaue Kolonien bei 30° auf X-Gal-Indikatorplatten infolge einer methylierungsabhängigen Restriktion, die in einer SOS-DNA-Reparaturinduktion und β -Galactosidaseexpression resultiert. Aufgrund einer engen Kopplung zwischen den meisten Restriktionsenzymgenen und den zugehörigen Methylasegenen kann die Klonierung eines Methylasegens in einem DNA-Fragment angemessener Größe zu einer gleichzeitigen Klonierung des zugehörigen Endonucleasegens führen.

ZUSAMMENFASSUNG DER ERFINDUNG

[0011] Die vorliegende Erfindung nutzt ein neuartiges Verfahren zur direkten Klonierung von Nucleasegenen wie Restriktionsendonucleasegenen in *E. coli*. Darüber hinaus nutzt sie einen neuartigen Stamm, der die Anwendung des obigen Verfahrens erleichtert.

[0012] Die Erfindung stellt isolierte DNA bereit, die BsoBI Endonuclease kodiert, sowie Vektoren und transformierte Wirtszellen, um diese Endonuclease zu exprimieren.

[0013] Die vorliegende Erfindung nutzt (i) einen neuen *E. coli* Stamm, der die *dinD1::lacZ* Fusion enthält, wobei es dem Stamm an allen Restriktionssystemen (*EcoKR*⁻, *McrA*⁻, *McrBC*⁻, *Mrr*⁻) mangelt, (ii) ein Verfahren zur direkten Klonierung von Nucleasegenen in *E. coli* unter Verwendung eines Stammes, der eine Fusion eines DNA-Schaden induzierenden Promotors und eines Indikator-/Reportergens, wie die *dinD1::lacZ* Fusion, enthält.

[0014] Da gezeigt wurde, dass DNA-Brüche oder -Schnitte, die durch die T7.3 Endonuclease, *EcoRI* oder *BamHI* Restriktionsenzyme eingeführt werden, die SOS-Reaktion in *E. coli* induzieren (Panayotatos und Fontaine, J. Biol. Chem. 260:3173–3177 (1985), Heitman und Model, Gene, 103:1-9, (1991), Xu und Schildkraut, J. Biol. Chem. 266:4425–4429, (1991)), gingen die Autoren der vorliegenden Erfindung davon aus, dass, wenn ligierte genomische DNA-Fragmente und Vektor in einen Indikatorstamm wie *dinD1::lacZ* eingeführt werden, dem es an allen bisher beschriebenen Restriktionssystemen fehlt, und Transformanten auf X-Gal-Platten plattiert werden, nucleasehaltige Klone direkt durch Abimpfen blauer Kolonien gefunden werden können. Bei der Verwendung zum Klonieren von Genen, die eine Restriktionsendonuclease kodieren, ist es im Gegensatz zum Methylase-selektionsansatz nicht nötig, dass das Methylasegen das Wirtschromosom völlig schützt. In der Tat kann das Methylasegen ganz abwesend sein. Dies trifft besonders auf thermostabile Enzyme zu, bei denen die Transformanten gemäß der vorliegenden Erfindung bei niedrigeren Temperaturen wachsen gelassen werden, d.h. zwischen etwa 30°C und 37°C. Bei dieser niedrigeren Temperatur sind thermostabile Restriktionsendonucleasen weniger aktiv und transformierte Wirtszellen können mit teilweiser oder sogar ohne schützende Methylierung überleben.

[0015] Mit anderen Worten, gemäß der vorliegenden Erfindung trifft Folgendes zu: wenn Wirtszellen wie die bevorzugten *E. coli* Zellen (*dinD1::lacZ*, *EcoKR*⁻, *McrA*⁻, *McrBC*⁻, *Mrr*⁻) mit ligierter Genom-/Vektor-DNA transformiert und Transformanten auf X-Gal-Platten plattiert werden, dann bilden Zellen, die das Restriktionsendonucleasegen tragen, blaue Kolonien, da das Restriktionsenzym DNA in vivo schädigt und die SOS-DNA-Reparaturreaktion induziert. Dieses Verfahren (das „Endonuclease-Indikator-Verfahren“) unterscheidet sich von

dem „Methylase-Indikator-Verfahren“, Piekarowicz et al., supra, dadurch, dass das frühere Verfahren eine Expression der Methylase und nicht der Endonuclease nachwies. Dieses Verfahren war auf die DNA-schädigende Wirkung endogener Restriktionsenzyme (McrA, McrBC oder Mrr) angewiesen, die auf spezifische methylierte Sequenzen wirken. Die Expression einer geeigneten Fremd-Methylase kann eine Sequenz hervorbringen, die für eine oder mehrere dieser methylationsspezifischen Endonucleasen empfänglich ist, was zur SOS-Induktion und blauen Farbe führt. Das Gen für die zugehörige Endonuclease kann das Methylasegen begleiten oder auch nicht. Das in der Erfindung angewendete Endonuclease-Indikator-Verfahren weist nur die Endonuclease und nicht die Methylase nach, da die relevanten methylierungsabhängigen Restriktionssysteme im Wirt nicht vorliegen.

KURZE BESCHREIBUNG DER FIGUREN

[0016] (Fig. 1 und 2: GESTRICHEN)

[0017] Fig. 3 zeigt einen BsoBI-Aktivitätstest. Ein µg Lambda-DNA-Substrat wurde mit µl 10 Zellextrakt bei 65°C eine Stunde lang in einem Puffer aus 50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT inkubiert. Die verdauten DNA-Produkte wurden in einem 0,8%igen Agarosegel aufgetrennt und durch Ethidiumbromidfärbung nachgewiesen. Man stellte fest, dass der Zellextrakt vom Isolat Nr. 16 BsoBI Endonucleaseaktivität enthielt. Erste Bahn: BsoBI positive Kontrolle. Die Zahl über den jeweiligen Bahnen ist die Isolatnummer.

[0018] Fig. 4A ist die BsoBI Methylase (SEQ ID Nr. 3). Fig. 4B ist das BsoBI-Restriktionsendonucleasegen (SEQ ID Nr. 4). Fig. 4C ist die entsprechende Aminosäuresequenz für das BsoBI Methylasegen (SEQ ID Nr. 5). Fig. 4D ist die entsprechende Aminosäuresequenz für das BsoBI-Restriktionsendonucleasegen (SEQ ID Nr. 6).

[0019] Hinweis: In Fig. 4B (SEQ ID Nr. 4) und Fig. 4D (SEQ ID Nr. 6) treten Sequenzierungsfehler auf. Informationen über den fehlerhaften Inhalt sind in der öffentlichen Akte der Anmeldung als ergänzende technische Informationen enthalten. Fig. 5 ist eine Restriktionskarte des DNA-Inserts, das das BsoBI Restriktionsendonucleasegen enthält. Die Insert-Größe liegt bei etwa 8,0 kb. Das Insert wurde von mit Sau3AI teilweise verdauter *Bacillus stearothermophilus* Genom-DNA abgeleitet und in den BamHI Ort von pUC19 kloniert.

AUSFÜHRLICHE BESCHREIBUNG DER ERFINDUNG

[0020] Die Erfindung nutzt ein neuartiges Verfahren zur direkten Klonierung von Nucleasegenen. Im Allgemeinen umfasst das Verfahren die folgenden Schritte, allerdings wird es der fachkundigen Person verständlich sein, dass Abänderungen dieser Schritte möglich sind, ohne dass das Ergebnis negativ beeinflusst wird.

1) Genomische DNA wird von dem Nuclease produzierenden Stamm hergestellt und komplett oder teilweise gespalten, um etwa 500 bp – 20.000 bp an klonierbaren DNA-Fragmenten zu erzeugen. Diese Fragmente können zum Beispiel mit Restriktionsenzymen erhalten oder durch Beschallung abgesichert werden. Die so erhaltenen Fragmente werden anschließend mit Klonierungsvektoren mit kompatiblen kohäsiven oder stumpfen Enden in pBR322, pUC19, pACYC187, pSC101 oder deren Derivaten ligiert.

2) Das ligierte DNA-Gemisch wird in einen Bakterienstamm übertragen, der einen DNA-Schaden produzierenden Promotor enthält, der mit einem Indikator/Reporter verschmolzen ist, wie die bevorzugte *dinD::lacZ* Fusion, und dem es an methylierungsabhängigen Restriktionssystemen (*dinD::lacZ*, *hsdR*, *mcrA*, *mcrBC*, *mrr*) mangelt. Ein bevorzugter Stamm ist *E. coli* ER1992 (NEB Nr. 907), von dem eine Probe bei der Amerikanischen Typus-Kultursammlung unter dem Budapest Vertrag am 24. Mai 1994 mit der ATCC Beitrittsnummer 55582 hinterlegt wurde. Zu weiteren DNA-Schaden induzierenden Promotoren, die verwendet werden können, gehören *dinA* (Iwasaki, M., et al., J. Bacteriol., 172:6268–6273 (1990) und *dinG* (Lewis, L.K., J. Bacteriol. 174:5110–5116 (1992).

Zu anderen Indikator-/Reportergen, die mit einem beliebigen der obigen Promotoren verschmolzen werden können, gehören alkalische Phosphatase (*phoA*) (Proc. Natl. Acad. Sci. USA, 82:5107–5111 (1985), Luciferase (*lux*) (Engelrecht, J., Science, 227:1345–1347 (1985), (3-Glucuronidase (Metcalf, W.W., Gene, 129:17–25 (1993), Aminoglycosid-Phosphotransferase (Ward, J.M., et al., Mol. Gen. Genet., 203:468–478 (1986) und Endoglucanase (Bingle, W.W., et al., Can. J. Microbiol., 39:70–80 (1993).

Nach der Transformation in *E. coli* ER1992 werden die Zellen auf Indikatorplatten plattiert, die X-Gal und geeignete Antibiotika enthalten, und über Nacht bei etwa 30°C bis 42°C inkubiert. Manchmal braucht keines oder brauchen nicht alle der methylierungsabhängigen Restriktionssysteme des Wirts inaktiviert zu werden, was von dem speziellen zu klonierenden.

Restriktionsendonucleasegen oder den Nucleasegenen abhängig ist. Ein totaler Mangel an diesen Systeme-

men ist zwar nicht erforderlich, doch wird es bevorzugt, die *dinD::lacZ* Stämme zu verwenden, denen es an allen methylierungsabhängigen Systemen mangelt, um ein Restriktionsmodifikationssystem zu klonieren, wenn die modifizierten Basen (unüblicherweise C⁵ Cytosin, N⁴ Cytosin oder N⁶ Adenin) nicht bekannt sind.

3) Individuelle mittel-/dunkelblaue Kolonien werden aufgenommen und in LB-Medium (10 ml bis 1000 ml) plus geeigneter Antibiotika inokuliert und über Nacht bei etwa 30°C bis 42°C geschüttelt.

4) Zellen werden durch Zentrifugation geerntet, in Beschallungspuffer plus Lysozym resuspendiert, und die Lyse von Zellen wird durch Beschallung abgeschlossen. Zelltrümmer und unlösliche Komponenten werden durch Zentrifugation entfernt.

5) Wenn das zu klonierende Nucleasegen von einem thermostabilen Bakterium stammt, dann wird das Ly-sat über einen Zeitraum (z.B. 30 Minuten), der ausreicht, um native E.coli Proteine zu denaturieren, auf etwa 65°C erwärmt. Dieser Schritt inaktiviert wirksam native E. coli Nucleasen.

6) Der Überstand (Zellextrakt) wird hinsichtlich Nucleaseaktivität auf geeigneten DNA-Substraten wie pUC19, pBR322, M13mp18/19 replikativer Form oder einzelsträngige DNA oder λ DNA bei 37°C bis 68°C in einem geeigneten Puffer getestet.

7) DNA-Verdaumuster oder Fragmente werden durch Agarose-Gelelektrophorese oder PAGE aufgetrennt und durch Ethidiumbromidfärbung nachgewiesen.

[0021] Das oben beschriebene Verfahren ist besonders zum Klonieren der thermostabilen Nuclease BsoBI vom *Bacillus stearothermophilus* (NEB Nr. 882) geeignet. BsoBI ist ein Isoschizomer von Aval, das CPyCG-PuG (SEQ ID Nr. 7) erkennt.

[0022] Die Erfindung kann einen neuartigen Stamm nutzen, der in dem oben beschriebenen Endonuclease-Indikator-Verfahren verwendet werden kann. Dieser Stamm, E. coli ER1992, enthält die *dinD::lacZ* Fusion und es mangelt ihm an allen Restriktionssystemen (EcoKR⁻, McrA⁻, McrBC⁻ und Mrr⁻).

[0023] E. coli ER1992 (F⁻ λ^- Δ (argF-lac)U169 supE44e14⁻ *dinD1::Mu* dl1734 (Kan^R, LacZ⁺) *rfbD1?* *relA1?* *endA1* *spoT1?* *thi-1* Δ (*mcrC-mrr*)114::IS10) wurde in drei Schritten konstruiert (i) ein Lac⁻ Derivat von ER1821 wurde durch Transduktion mit *proC::Tn5*, verknüpft mit Δ (argF-lac)U169 von NK6993, erhalten, wobei nach Kan^R selektiert und nach Lac⁻ Pro⁻ gescreent wurde, um ER1984 zu gewinnen; (ii) dieses Derivat wurde zu Pro⁺ Kan^s gemacht durch Transduktion von ER1578, so dass ER1991 gewonnen wurde; (iii) *dinD1::Mu* dl1734 (Kan^R, LacZ⁺) wurde durch Transduktion von JH140 (ein E. coli Stamm, der *dinD1::Mu* dl1734 (Kan^R, LacZ⁺) ist) eingeführt (J. Heitman, et al., *Gene* 103:1-9 (1991)), wobei Kan^R selektiert und nach einer Nalidixinsäure induzierbaren Expression von β -Galactosidase gescreent wurde, vermittelt durch die *dinD*-Fusion. Dies wurde auf X-Gal-Platten mit einem zentralen Well getestet, das dieses DNA-schädigende Agens enthielt. Gereinigte Transduktanten wurden von dem Well aus radial ausgestrichen. Einer, der einen Gradienten von dunkelblauer Farbe gewann, wurde als E. coli ER1992 bezeichnet. Dieser Stamm wies eine hellblaue Farbe auf X-Gal auf, wenn kein DNA-Schaden vorlag.

1. BEISPIEL

KLONIERUNG DES *bsobIR* GENS, DAS DIE *bsobI* RESTRIKTIONSENDONUCLEASE KODIERT, IN E. coli

[0024] Bakterien-DNA-Reinigung: Fünf Gramm *Bacillus stearothermophilus* (NEB Nr. 882) Zellen wurden in 25 ml Puffer resuspendiert, der 25 % Saccharose, 50 mM Tris-HCl, pH 8,0, enthielt. Fünf Milliliter 0,5 M EDTA, pH 8,0, und 6 ml Lysozym (10 mg/ml) wurden zur Zellsuspension gegeben. Nach einer 10-minütigen Inkubation bei Raumtemperatur wurden 36 ml Lysepuffer (1 % Triton X-100, 50 mM Tris-HCl, pH 8,0, 62 mM EDTA) und 5 ml von 10 % SDS zugegeben, um die Zellen vollständig zu lysieren. Proteine wurden durch zweimalige Phenol-CHCl₃-Extraktion und zweimalige CHCl₃-Extraktion entfernt. Genomische DNA wurde durch Zugabe von 1/10 Volumen von 3,5 M Natriumacetat und einem gleichen Volumen Isopropanol präzipitiert und bei 15.000 rpm/min zentrifugiert. Das DNA-Pellet wurde mit 50 ml 70%igem Ethanol gewaschen und unter Vakuum getrocknet. Das DNA-Pellet wurde in 10 ml TE-Puffer resuspendiert und in 2 Liter TE-Puffer (10 mM Tris-HCl, pH 7,8, 1 mM EDTA) über Nacht bei 4°C dialysiert.

[0025] Teilweiser Verdau genomischer DNA: Fünfzig Mikrogramm genomische DNA wurden jeweils mit 2 Einheiten, 1 Einheit, 0,5 Einheiten, 0,25 Einheiten, 0,125 Einheiten von Sau3AI bei 37°C 30 Minuten lang verdaut. Die verdaute DNA wurde durch zweimalige Phenol-CHCl₃-Extraktion und zweimalige CHCl₃-Extraktion und Ethanolpräzipitation gereinigt. Vektor pUC19 DNA wurde durch das Restriktionsenzym BamHI linearisiert und durch alkalische Kälberdarm-Phosphatase (CIP) dephosphoryliert. Nach der CIP-Behandlung wurde die Vektor-DNA durch zweimalige Phenol-CHCl₃-Extraktion und durch zweimalige CHCl₃-Extraktion und Ethanolpräzipitation gereinigt.

[0026] Ligation und Transformation: Zehn Mikrogramm von mit Sau3AI teilweise verdauter Genom-DNA wurden mit 1 µg BamHI-gespaltener und CIP-behandelter pUC19 DNA über Nacht bei 16°C ligiert. Die ligierte DNA wurde unter Zugabe eines gleichen Volumens TE-Puffer verdünnt und zum Transformieren von kompetenten ER1992 (dinD1::lacZ⁺, hsdR, mcrA, mcrBC, mrr) Zellen verwendet. Transformanten wurden auf Ampicillin (Ap) plus X-Gal-Platten (5-Bromo-4-chloro-3-indolyl-D-galactopyranosid, X-Gal, 0,016 % Gewicht/Vol.) plattiert und über Nacht bei 37°C inkubiert. Unter 10.000 Transformanten wurden 24 blaue Kolonien gefunden und jede blaue Kolonie wurde in 10 ml LB plus Ap inokuliert und über Nacht in einem Schüttler bei 37°C inkubiert. Zellen wurden durch Zentrifugation geerntet und in 1 ml Beschallungspuffer (10 mM Tris-HCl, pH 7,8, 10 mM β-Mercaptoethanol, 10 mg/ml Lysozym) resuspendiert. Die Lyse der Zellen wurde durch zweimalige Beschallung über einen Zeitraum von jeweils 30 Sekunden abgeschlossen. Unlösliche Komponenten wurden durch Zentrifugation entfernt und der Überstand wurde aufgenommen und wie folgt für einen Endonucleaseaktivitätstest verwendet: 1 µg Lambda-DNA-Substrat wurde mit 10 µl Zellextrakt bei 65°C eine Stunde lang in einem Puffer aus 50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT inkubiert. DNA-Fragmente wurden in einem 0,8%igen Agarosegel aufgelöst und durch Ethidiumbromidfärbung nachgewiesen. Es wurde gefunden, dass der Zellextrakt von einem Isolat BsoBI Endonucleaseaktivität enthielt (siehe **Fig. 3**, Isolat Nr. 16). Plasmid-DNA wurde von diesem Stamm extrahiert und zur Prüfung der Resistenz mit BsoBI Restriktionsendonuclease verdaut. Die Plasmid-DNA wurde durch BsoBI gespalten, was ein Hinweis darauf war, dass entweder das BsoBI Methylasegen nicht in dem Insert vorlag oder das Methylasegen überhaupt nicht exprimiert war. Das Plasmid, das das BsoBI Endonucleasegen (BsoBIR) trug, wurde als pBsoR1 bezeichnet. Zur Bestimmung der Stabilität des BsoBI Klons wurden ER1992 [pBsoR1] Zellen bei 37°C wachsen gelassen und Plasmid-DNA wurde von den Zellen isoliert und zur Retransformation in kompetente ER1992 Zellen verwendet. Im Retransformations-experiment wurden Transformanten bei 30°C ausplattiert, um weitere Mutationen zu minimieren. Es wurde gefunden, dass 67 % der Transformanten blaue Kolonien bildeten (356 blaue Kolonien von 528 Transformanten), wohingegen die anderen 33 % weiße Kolonien bildeten (172 weiße Kolonien von 528 Transformanten). Als Plasmid-DNA von den weißen Kolonien isoliert und durch Restriktionsverdau analysiert wurde, stellte man fest, dass die DNA extensive Deletionen aufwies. Man gelangte zu dem Schluss, dass ER1992 [pBsoR1] Zellen bei 37°C nicht sehr stabil sind. ER1992 [pBsoRI] Zellen waren stabiler, als die blauen Kolonien, die pBsoRI trugen, über Nacht bei 30°C anstatt bei 37°C inkubiert wurden. Als von der 30°C-Kultur isolierte Plasmid-DNA für die Retransformation verwendet wurde, bildeten 99,6 % der Transformanten blaue Kolonien bei 30°C (480 blaue Kolonien, 2 weiße Kolonien). Eine Probe von pBsoRI (NEB NR. 951) wurde am 13. Dezember 1994 unter dem Budapester Vertrag bei der Amerikanischen Typus-Kultursammlung mit der ATCC-Beitriffs-Nr. 75966 hinterlegt.

[0027] Die das Aval Restriktions-Modifikationssystem kodierenden Gene wurden bei New England Biolabs kloniert. Aval und BsoBI haben die gleiche Erkennungssequenz 5'CPyCGPuG3'(SEQ ID Nr. 7). Ein 1,8 kb PstI Fragment, das das Aval Methylasegen (avaIM) trug, wurde von pAvalRM12 verdaut und in den pR976 Vektor (ein pACYC184 Derivat, Tc^R) kloniert. Das Plasmid pR967-AvalM⁺ wurde in ER1992 (dinD::lacZ) Zellen transformiert, um das E. coli Chromosom zu prämodifizieren. Ein zweites Plasmid, pBsoRI, das das BsoBI Endonucleasegen trug, wurde anschließend in die Zellen eingeführt, die pR976-AvalM⁺ enthielten, und auf eine X-Gal-Indikatorplatte plattiert. Es wurde gefunden, dass die Transformanten noch immer blaue Kolonien bildeten, was ein Hinweis darauf war, dass die Aval Methylase das E. coli Chromosom nicht vor BsoBI Endonuclease-schaden schützt.

[0028] Plasmid-Mini-Präparationsverfahren: Eine Übernacht-Kultur von 1,5 ml wurde 3 Minuten lang bei 14.000 rpm/min pelletiert. Der Überstand wurde abgegossen und das Zellpellet wurde in 200 µl STET-Puffer (50 mM Tris-HCl, pH 7,8, 50 mM EDTA, 0,5 % Triton-X100, 8 % Saccharose) resuspendiert. 50 µl Lysozym (10 mg/ml) wurden zur Zellsuspension gegeben. Die lysierten Zellen wurden 1 Minute lang in kochendem Wasser gekocht und das Präzipitat wurde 10 Minuten lang bei 14.000 rpm/min zentrifugiert. 200 µl Überstand wurden aufgenommen und mit 100 µl 7,5 M NH₃Ac und 600 µl 95%igem Ethanol vermischt. DNA wurde durch eine 10-minütige Zentrifugation bei 14.000 rpm/min bei Raumtemperatur präzipitiert. Das DNA-Pellet wurde mit 1 ml 70%igem Ethanol gewaschen und 15 Minuten lang unter Vakuum getrocknet. Nach dem Trocknen wurde das Pellet in 100 µl TE-Puffer plus 5 µl RNaseA (10 mg/ml) resuspendiert.

[0029] Restriktionskartierung des Inserts: Die Plasmid-DNA pBsoR1 wurde mit verschiedenen Restriktions-enzymen verdaut und die verdauten DNA-Produkte wurden durch Agarosegel-Elektrophorese analysiert. Die Restriktionskarte ist in **Fig. 5** dargestellt.

[0030] Das bsoBIR Gen wurde innerhalb eines 1,4 kb EcoNI/XbaI Fragments in dem 8 kb Genom-DNA-Insert durch Deletionskartierung kartiert. Das gesamte bsoBIR Gen wurde von Subklonen und Deletionsklonen (SspI Fragment-Subklon, AvrII/XbaI Subklon, NlaIII Fragment-Subklon, Sau3AI Fragment-Subklon, EcoNI/AvrII Sub-

klon, AvrII/XhoI) und „Primer-Walking“ sequenziert. Das Startcodon des bsoBIR Gens wurde einem TTG-Codon zugeordnet, da die N-terminale Proteinsequenz der gereinigten BsoBI Endonuclease mit der vorhergesagten Proteinsequenz hinter diesem Startcodon übereinstimmt ((M) NTQKPFENHLKSVDL (entsprechend den Aminosäuren 1–17 der SEQ ID Nr. 6)). Die DNA auf beiden Seiten des bsoBIR Gens wurde auch sequenziert, um ggf. die Existenz eines Teils des Cytosin-Methylasegens zu bestimmen, das im 8,0 kb DNA-Fragment kloniert wurde. Das bsoBIM Gen oder ein Teil des bsoBIM Gens könnte auf der Basis der konservierten Methylasemotive identifiziert werden. Als die DNA-Region hinter dem bsoBIR Gen sequenziert und auf allen sechs Rastern translatiert wurde, wurde ein konserviertes N4 Cytosin-Methylasemotiv gefunden. Diese Sequenz lautet DPFLGSGTT (entsprechend den Aminosäuren 106–114 der SEQ ID Nr. 5). Nur 727 bp des bsoBIM Gens lagen auf dem ursprünglichen 8,0 kb Insert vor. Daher wurde eine inverse PCR angewendet, um den restlichen Teil des Methylasegens zu klonieren. Klonierung und Sequenzierung der inversen PCR-Produkte von HaeII und NlaIII gespaltenen/selbstligierten Genom-DNA brachten weitere 902 bp des Methylasegens hervor, in denen ein weiteres konservatives Methylasemotiv TSPPY (entsprechend Aminosäuren 319–323 der SEQ ID Nr. 5) gefunden wurde. Das gesamte Methylasegen (1629 bp) wurde durch lac^R von Genom-DNA amplifiziert und in pSX33lacI^q kloniert und in kompetente E. coli RR1 Zellen transformiert, um den Wirt zu prämodifizieren. Ein das bsoBIR Gen enthaltendes BbsI/XbaI Fragment wurde in einen Expressionsvektor pRRS kloniert, um das überproduzierende Konstrukt E. coli RR1 (pSX33lacI^q-BsoBIM⁺, pRRS-BsoBIR⁺) zu gewinnen. Die Sequenz der Gene bsoBIM und bsoBIR ist jeweils in den **Fig. 4–1 bis 4–4** dargestellt.

SEQUENZAUFLISTUNG

[0031] (1) ALLGEMEINE INFORMATIONEN

(i) ANMELDERIN:

(A) NAME: NEW ENGLAND BIOLABS, INC.

(B) STRASSE: 32 TOZER ROAD

(C) STADT: BEVERLY

(D) STAAT: MA

(E) LAND: US

(F) POSTLEITZAHL: 01915

(ii) TITEL DER ERFINDUNG: VERFAHREN ZUR DIREKTEN KLONIERUNG VON NUCLEASEGENEN IN E. COLI

(iii) ANZAHL SEQUENZEN: 7

(v) COMPUTERLESBARES FORMULAR:

(A) MEDIENTYP: Diskette

(B) COMPUTER: IBM-kompatibel

(C) BETRIEBSSYSTEM: Windows 2000

[0032] (2) INFORMATIONEN ZUR SEQ ID NR: 1

(i) SEQUENZCHARAKTERISTIKEN:

(A) LÄNGE: 4

(B) TYP: NUCLEINSÄURE

(C) STRÄNGIGKEIT: EINZEL

(D) TOPOLOGIE: LINEAR

(XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 1: 1 TCGA

[0033] (2) INFORMATIONEN ZUR SEQ ID NR: 2

(i) SEQUENZCHARAKTERISTIKEN:

(A) LÄNGE: 9

(B) TYP: NUCLEINSÄURE

(C) STRÄNGIGKEIT: EINZEL

(D) TOPOLOGIE: LINEAR

(XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 2: 1 GACNNNGTC

[0034] (2) INFORMATIONEN ZUR SEQ ID NR: 3

(i) SEQUENZCHARAKTERISTIKEN:

(A) LÄNGE: 1629

(B) TYP: NUCLEINSÄURE

(C) STRÄNGIGKEIT: UNBEKANNT

(C) TOPOLOGIE: UNBEKANNT

(XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 3:

```

1  GTGGAAACCG TACAAATGAG CTGCGATTTA CAACCAGAAG ACACAAAAGC
51  TAATATTGAG TGTTCCTCTC CTGATTACGA AATAAAGTGT TTCTCCATA
101 ATCAATTGTC TCCTCAAATT GAGCGTTTAA AAATTGAAAA AAAATACATT
151 CATATTGTAG AAGAGACAGA AACATTTAAC AGAAAATTAG TGAGTTACCA
201 AGCCAACAAG AACCAATCAA TCCACAATTG GATTAGATAT AAAGAGGGGT
251 TCTCCTCCGA GTTGGTTCCA AATCTAATTG AAGAGTTTGG ATTATCAAAA
301 GGTGACACTA TATTAGACCC TTTTCTAGGC TCAGGAACTA CAAGTCTGAC
351 TCCGAGAATA CTCGGAATAG ATAGTATTGG AATAGACATT TTGCCAATTT
401 CACACATAGC ATTTGAACCC AAATCCTTTA TCTTTGAATA TAATCTCGAA
451 GAACTAGATA GGGCCTATAG AGAAATATGC GAAATATCGC CAACGAAGAT
501 AGACCAAAAAG TTTAATCATT TAAGTATAGC TGAGGGCGCT TTTCCAGAAG
551 AAACAGAAAA TGACCTTTTA TTTTTTACCC ATTGGGATAA TAACTCACAG
601 TATAGTTATC AAACCAAAAC GTTAATAAAA TTAATCTTGG TAAGCATACT
651 AGAAGAGATT AGCTATACTA GAAAAGATGG GCAATATTTA AGATGGGATT
701 ATAGATCACA GAAGGTTATC GAAACTAATA AAAAAAGACT TGAACAAGGA
751 AAAGAGCCAA TTAAAACCAT ATTAGATAAG GGCGAACTTC CTACTGTGAA
801 AGAGTCCCTT TTAAACACTC TTTTAACAAT AAAAGAAGAT ATAAAAGAAA
851 TACAACAAAA ATGTTTACCT AATGAGTCCG TTCATGAACT TATAAAAGAC
901 AGTGCACTTA ATGCACTACC TAAAATTAAT GATAATACTT TTGACGCTGT
951 TATTACTTCC CCCCCATATT GCAATAGATA CGACTACACT AGAACATATG
1001 CACTTGAACT TGCTTACTTA GGGGTAGATG AAAAGAAGAT AAGAGAATTA
1051 AGACAAGCAC AACTTTCGTG TACAGTAGAA AATAAATCTA AACTTAAACA

```

1101 ACTAAAGGAC TACTACCATT CACTTTTCCT AGAAAGTAGA TACGCAGAAA
 1151 TTGAACGACT TGTAACAGGA AATGAGGTAT TAAATGAAAT TAATTACGCT
 1201 TTAAGGAAGC GCTGGGAAAA CGGTGAGGTT AATAACAAGG GTATACTTTC
 1251 AATGGTTGAC GGATATTTTA CAGAGTTAAC CTTTATATTT TATGAACTAT
 1301 TTAGAACGTG TAAGCCAGGA GCTAAAGTTG CCTTTGTAA TGATAATGTT
 1351 CGGTATGCAG GTGAAATTAT CCTGTGTGAT TTCTTATCAA CGAAAATTGC
 1401 AGAAGATATC GGTTTTAAAC CAATTAAAAT CTATACACTA AAACAAAGAA
 1451 AAGGCAATAG CAGTCAACAA ATGGGTAAGT ATGGCAAGAG TAGCGTTAAG
 1501 AAAAAGCATC ACAATATGGG AAAAACCATA AGAGTGAGCA CAGTCTCACT
 1551 CTCTTTTAT TTTTATCGCA ACTTATCCCC CACATATAAC GATTACTATG
 1601 GCGGAAGGGG GAAGAAAGAT GATAATTGA

[0035] (2) INFORMATIONEN ZUR SEQ ID NR: 4

(i) SEQUENZCHARAKTERISTIKEN:

(A) LÄNGE: 972

(B) TYP: NUCLEINSÄURE

(C) STRÄNGIGKEIT: UNBEKANNT

(D) TOPOLOGIE: UNBEKANNT

(XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 4:

1 TTGAACAETC AAAAACCTTT TGAAATCAT CAAAAAGTG TAGACGATTT
 51 AAAAACGACA TATGAAGAAT ACAGAGCAGG ATTTATTGCT TTTGCTTTAG
 101 AGAAAAATAA ACGCTCTACC CCATATATTG AGAGAGCACG AGCATTAAAA
 151 GTTGCTGCAT CTGTTGCAA AACACCAAAA GATTTACTTT ACTTGAAGA
 201 TATTCAAGAC GCTCTCCTTT ATGCGTCGGG TATTTCTGAT AAAGCAAAGA
 251 CATTCTTAAC GGAAGATGAT AAAAAGGAGT CTATAAATAA TTTAATTGAG
 301 AACTTTTTTG GACCTGCCGG AGAGGAGTTT ATAGACGAGT TAATATTCCG
 351 CTATTTATTA TTCCAAGGCC TAGGAGGAAC AATGAGAAAC ATAGCAGGTG
 401 CACTCGCTCA ACAAAAATTA ACAAGAGCAA TAACTCTCTGC CTCGGATATT
 451 GCTAATATCC CTTATAAATG GTTAGACTCA CGAGATAAAA AGTACACTAA
 501 TTGGATGGAC AAGCCTGAGG ATGATTACGA ATTAGAAACA TTTGCTAAAG
 551 GTATTAGCTG GACTATTAAC GGAAAACACA GAACTTTAAT GTATAATATT
 601 ACTGTTCCAC TTGTAAAAAA GAATGTTGAT ATATGTTTAT TTAAGTGTGA
 651 ACCTGAAATA TATACGCCAC AAAAAGTACA CCAACAGCCT GAAAAGTATC

701 TACTGCTTGG AGAGTTAAAG GGC GGAATAG ACCCTGCTGG AGCGGATGAA
 751 CATTGGAAAA CACCAAATAC TGCTCTTACT AGGATTAGAA ATAATTCATC
 801 CGAAAAGGGA TTATCACCTA AAACAATTTT TATAGGCGCT GCTATTGAGC
 851 ATAGTATGGC AGAAGAAATA TGGGACCAAC TTCAATCCGG TTCATTAACA
 901 AACTCAGCAA ATCTTACTAA AACTGAACAA GTAGGCTCTC TTTGTAGATG
 951 GATAATTAAT ATATAA

[0036] (2) INFORMATIONEN ZUR SEQ ID NR: 5

- (i) SEQUENZCHARAKTERISTIKEN:
- (A) LÄNGE: 542
- (B) TYP: AMINOSÄURE
- (C) STRÄNGIGKEIT: UNBEKANNT
- (D) TOPOLOGIE: UNBEKANNT
- (XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 5:

1 METVQMSLDL QPEDTKANIE CSSPDYEIKC FSHNQFAPQI ERLKIEKKYI
 51 HIVEETETFN RKLVSQANK NQSIHNWIRY KEGFSSELVP NLIEEFGLSK
 101 GDTILDPPFLG SGTSLTPRI LGIDSIGIDI LPISHIAFEP KSFIFEYNLE
 151 ELDRAYREIC EISPTKIDQK FNHLSIAEGA FPEETENDLL FFTHWDNNSQ
 201 YSYQTKTLIK LILVSILEEI SYTRKDGQYL RWDYRSQXVI ETNKKRLEQG
 251 KEPIKTILDK GELPTVKESL UNTLLTIKED IKEIQKCLP NESVHELKID
 301 SALNALPKIN DNTFDVITS PPYCNRYDYT RTYALELAYL GVDEKKIREL
 351 RQAQLSCTVE NKS KLQKLD YYHSLFLES R YAEIERLVTG NEVLNEINYA
 401 LRKRWENGEV NNGILSMVD GYFTELTFIF YELFRTCKPG AKVAFVNDNV
 451 RYAGEIIPVD FLSTKIAEDI GFKPIKIYTL KQRKGNSSQQ MGKYGKSSVK
 501 KKHNNMGKTI RVSTVSLSFY FYRNLSPYTN DYYGGRGKID DN

[0037] (2) INFORMATIONEN ZUR SEQ ID NR: 6

- (i) SEQUENZCHARAKTERISTIKEN:
- (A) LÄNGE: 323
- (B) TYP: AMINOSÄURE
- (C) STRÄNGIGKEIT: UNBEKANNT
- (D) TOPOLOGIE: UNBEKANNT
- (XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 6:

1 MNTQKPFENH LKSVDDLKTT YEEYRAGFIA FALEKNKRST PYIERARALK
 51 VAASVAKTPK DLLYLEDIQD ALLYASGISD KAKTFLTEDD KKE SINNLIE
 101 NFLGPAGEEF IAELIFRYLL FQGLGGTMRN IAGALAQKQL TRAIISALDI
 151 ANIPYKWLDS RDKKYTNWMD KPEDDYELET FAKGISWTIN GKHRTLMYNI
 201 TVPLVKKNV D ICLFNCEPEI YTPQKVHQQP EKYLLLGELK GGIDPAGADE
 251 HWKTANTALT RIRNKFSEKG LSPKTIFIGA AIEHSMAREI WDQLQSGSLT
 301 NSANLTKT EQVGS LCRWI INI*

[0038] (2) INFORMATIONEN ZUR SEQ ID NR: 7

(i) SEQUENZCHARAKTERISTIKEN:

(A) LÄNGE: 6

(B) TYP: NUCLEINSÄURE

(C) STRÄNGIGKEIT: UNBEKANNT

(D) TOPOLOGIE: UNBEKANNT

(XI) SEQUENZBESCHREIBUNG: SEQ ID NR: 7: 1 CSCGWG

Patentansprüche

1. Isoliertes DNA-Segment, das die BsoBI Restriktionsendonuclease kodiert, wobei das isolierte DNA-Segment vom Vektor pBsoRI (ATCC Nr. 75966) erhältlich ist.

2. Rekombinanter DNA-Vektor, umfassend einen Vektor, in den ein DNA-Segment nach Anspruch 1 eingesetzt wurde.

3. Klonierungsvektor, der das isolierte DNA-Segment aus Anspruch 1 umfasst.

4. Wirtszelle, die durch den Klonierungsvektor aus Anspruch 3 transformiert ist.

5. Verfahren zur Herstellung einer BsoBI Restriktionsendonuclease, umfassend das Kultivieren einer mit dem Vektor aus Anspruch 3 transformierten Wirtszelle unter Bedingungen zur Expression der genannten Endonuclease.

Es folgen 5 Blatt Zeichnungen

Anhängende Zeichnungen

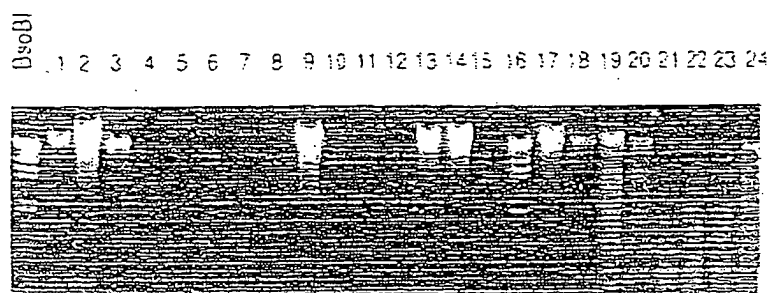


FIG. 3

```

1  GTGGAACCG TACAATGAG CTTGGATTTA CAACCAGAAG ACACAAAAGC
51  TAATATTGAG TGTTCCTCTC CTGATTACGA AATAAAGTGT TTCTCCCAT
101 ATCAATTTGC TCCTCAAATT GAGCGTTTAA AAATTGAAAA AAAATACATT
151 CATATTGTAG AAGAGACAGA AACATTTAAC AGAAAATTAG TGAGTTACCA
201 AGCCAACAAG AACCAATCAA TCCACAATTG GATTAGATAT AAAGAGGGGT
251 TCTCCTCCGA GTTGGTTCCA AATCTAATTG AAGAGTTTGG ATTATCAAAA
301 GGTGACACTA TATTAGACCC TTTTCTAGGC TCAGGAACTA CAAGTCTGAC
351 TCCGAGAATA CTCGGAATAG ATAGTATTGG AATAGACATT TTGCCAATTT
401 CACACATAGC ATTTGAACCC AAATCCTTTA TCTTTGAATA TAATCTCGAA
451 GAACTAGATA GGGCCTATAG AGAAATATGC GAAATATCGC CAACGAAGAT
501 AGACCAAAAAG TTTAATCATT TAAGTATAGC TGAGGGCGCT TTTCCAGAAG
551 AAACAGAAAA TGACCTTTTA TTTTTTACCC ATTGGGATAA TAACTCACAG
601 TATAGTTATC AAACCAAAAC GTTAATAAAA TTAATCTTGG TAAGCATACT
651 AGAAGAGATT AGCTATACTA GAAAGATGG GCAATATTTA AGATGGGATT
701 ATAGATCACA GAAGGTTATC GAAACTAATA AAAAAAGACT TGAACAAGGA
751 APAGAGCCAA TTAAAPCCAT ATTAGATAAG GGCGAAGCTC CTAAGTGTGA
801 AGAGTCCCTT TTAAACACTC TTTTAACAAT AAAAGAAGAT ATAAAAGAAA
851 TACAACAAAA ATGTTTACCT AATGAGTCCG TTCATGAACT TATAAAGAC
901 AGTGCACTTA ATGCACTACC TAAATTAAT GATAATACTT TTGACGCTGT
951 TATTACTTCC CCCCCATATT GCAATAGATA CGACTACACT AGAACATATG
1001 CACTTGAACT TGCTTACTTA GGGGTAGATG AAAAGAAGAT AAGAGAATTA
1051 AGACAAGCAC AACTTTCGTG TACAGTAGAA AATAAATCTA AACTTAAACA
1101 ACTAAAGGAC TACTACCAT CACTTTTCCT AGAAAGTAGA TACGCAGAAA
1151 TTGAACGACT TGTAACAGGA AATGAGGTAT TAAATGAAAT TAATTACGCT
1201 TTAAGGAAGC GCTGGGAAA CGGTGAGGT AATAACAAGG GTATACTTTC
1251 AATGGTTGAC GGATATTTTA CAGAGTTAAC CTTTATATTT TATGAACAT
1301 TTAGAACGTG TAAGCCAGGA GCTAAAGTTG CCTTTGTTAA TGATAATGTT
1351 CGGTATGCAG GTGAAATTAT CCTGTGAT TTCTTATCAA CGAAAATTGC

```

FIG. 4A

1401 AGAAGATATC GGTTTTAAAC CAATTAAAT CTATACACTA AAACAAAGAA
 1451 AAGGCAATAG CAGTCAACAA ATEGGTAAGT ATGGCAAGAG TAGCGTTAAG
 1501 AAAAAGCATC ACAATATGGG AAAAACCATA AGAGTGAGCA CAGTCTCACT
 1551 CTCTTTTAT TTTTATCGCA ACTTATCCCC CACATATAAC GATTACTATG
 1601 GCGGAAGGGG GAAGAAAGAT GATAATTGA

 1 TTGAACACTC AAAACCTTT TGAATCAT CTAAAAAGTG TAGACGATTT
 51 AAAACGACA TATGAAGAT ACAGAGCAGG ATTTATTGCT TTTGCTTTAG
 101 AGAAAAATAA ACGCTCTACC CCATATATTG AGAGAGCAGC AGCATTAAAA
 151 GTTGCTGCAT CTGTTGCAA AACACCAAA GATTACTTT ACTTGGAAGA
 201 TATCAAGAC GCTCTCCTTT ATGCGTCGGG TTTTCTGAT AAAGCAAAGA
 251 CATCTTAAC GGAAGATGAT AAAAGGAGT CTATAATAA TTTAATTGAG
 301 AACTTTTGG GACCTGCCGG AGAGGAGTTT ATAGCCGAGT TAATATTCCG
 351 CTATTTATTA TTCCAAGGCC TAGGAGGAAC AATGAGAAC ATAGCAGGTG
 401 CACTCGCTCA ACAAAAATTA ACXAGAGCAA TAATCTCTGC CTCGGATATT
 451 GCTAATATCC CTTATAAATG GTTAGACTCA CGAGATAAAA AGTACACTAA
 501 TTGGATGGAC AAGCCTGAGG ATGATTACGA ATTAGAAACA TTTGCTAAAG
 551 GTATTAGCTG GACTATTAAAC GGAAACACA GAACTTTAAT GTATAATATT
 601 ACTGTTCCAC TTGTAACAAA GAATGTTGAT ATATGTTTAT TTAAGTGTGA
 651 ACCTGAAATA TATACGCCAC AAAAAGTACA CCAACAGCCT GAAAAGTATC
 701 TACTGCTTGG AGAGTTAAAG GGCGGAATAG ACCCTGCTGG AGCGGATGAA
 751 CATTGGAATA CACCAATAC TGCTCTTACT AGGATTAGAA ATAAATTCTC
 801 CGAAXAGGGA TTATCACCTA AAACAATTTT TATAGGCGCT GCTATTGAGC
 851 ATAGTATGGC AGAAGAAATA TGGGACCAAC TTCAATCCGG TTCATTAACA
 901 AACTCAGCAA ATCTTACTAA AACTGAACAA GTAGGCTCTC TTTGTAGATG
 951 GATAATTAAT ATATAA

FIG. 4B

1 METVQMSLDL QPEDTKANIE CSSPDYEIKC PSHNQFAPQI ERLKIEKKYI
 51 HIVEETETFN RKLVSQANK NQSIHNWIRY KEGPSSSELVP NLIEEFGLSK
 101 GDTILDPFLG SGTTSITPRI LGIDSIGIDI LPISHIAFEP KSFIFEYNLE
 151 ELDRAYREIC EISPTKIDQK FNHLSIAEGA FPEETENDLL FFTHWDNNSQ
 201 YSYQTKTLIK LILVSILEEI SYTRKDGQYL RWDYRSQKVI ETNKKRLEQG
 251 KEPIKTILDK GELPTVKESL LNTLLTIKED IKEIQQKCLP NESVHELKID
 301 SALNALPKIN DNTFDVITS PPYCNRYDYT RTYALELAYL GVDEKKIREL
 351 RQAQLSCTVE NKSCLKQLKD YYHSLFLESY YAEIERLVG NEVLNEINYA
 401 LRKRWENGEV NNKGILSMVD GYFTELTFFI YELFRTCKPG AKVAFVNDNV
 451 RYAGEIIPVD FLSTKIAEDI GFKPIKIYTL KQRKGNSSQQ MGKYGKSSVK
 501 KKHHNMGKTI RVSTVSLSFY FYRNLSPYD DYYGGRGKKD DN*

FIG. 4C

1 MINTQKPFENH LKSVDDLKTT YEEYRAGFIA FALERKNKRST PYIERARALK
 51 VAASVAKTPK DLLYLEDIQD ALLYASGISD KARTFLTEDD KRESINNLE
 101 NFLGPAGEEF IAEILFRYLL FQGLGGTMRN IAGALAQQL TRAIISADI
 151 ANIPYKWLDS RDKRYTNWMD KPEDDYELET FAKGISWTIN GKHRTLMYNI
 201 TVPLVKKNVD ICLFNCEPEI YTPQKVHQQP EKYLLLGELK GGIDPAGADE
 251 HWKTPNTALT RIRNKFSEKG LSPKTIFFGA AIEHSMAEEI WDQLQSGSLT
 301 NSANLTRTEQ VGSLCRWIIN I*

FIG. 4D

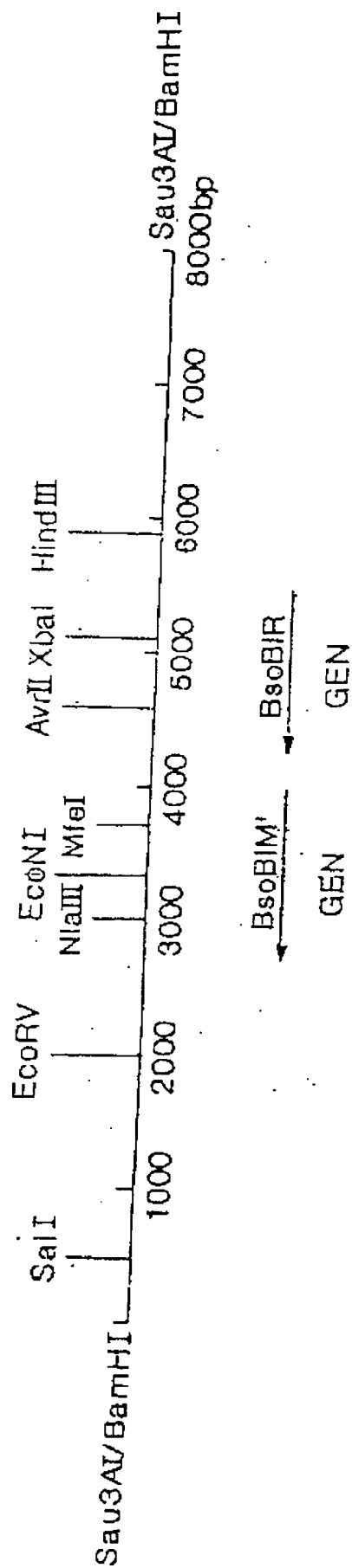


FIG. 5