



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 35 88 252 T2** 2006.02.02

(12)

## Übersetzung der europäischen Patentschrift

(97) **EP 0 518 443 B1**

(21) Deutsches Aktenzeichen: **P 35 88 252.2**

(96) Europäisches Aktenzeichen: **92 201 711.6**

(96) Europäischer Anmeldetag: **30.10.1985**

(97) Erstveröffentlichung durch das EPA: **16.12.1992**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **18.05.2005**

(47) Veröffentlichungstag im Patentblatt: **02.02.2006**

(51) Int Cl.<sup>8</sup>: **C12N 15/10** (2006.01)

(30) Unionspriorität:

**667501**            **31.10.1984**    **US**

**696534**            **30.01.1985**    **US**

**773447**            **06.09.1985**    **US**

(73) Patentinhaber:

**Chiron Corp. (n.d.Ges.d. Staates Delaware),  
Emeryville, Calif., US**

(74) Vertreter:

**BEETZ & PARTNER Patentanwälte, 80538  
München**

(84) Benannte Vertragsstaaten:

**AT, BE, CH, DE, FR, GB, IT, LI, LU, NL, SE**

(72) Erfinder:

**Luciw, Paul A., Emeryville, California 94608, US;  
Dina, Dino, Oakland, CA 94602, US; Steimer,  
Kathelyn, Benicia, California 94510, US; Pescador,  
Ray Sanchez, Oakland, California 94605, US;  
George-Nascimento, Carlos, Danville, California  
94526, US; Parkes, Deborah, No. 104 Oakland,  
California 94610, US; Hallewell, Rob, San  
Francisco, California 94109, US; Barr, Philip J.,  
Orinda, California 94563, US; Truett, Martha,  
Oakland, California 94611, US**

(54) Bezeichnung: **Rekombinante Proteine von mit Lymphadenopathiesyndrom und/oder erworbenem Immunschwächesyndrom assoziierten Viren**

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**

## Technisches Gebiet

**[0001]** Die vorliegende Erfindung stammt aus dem Gebiet des Lymphadenopathie-Syndroms und/oder des erworbenen Immunschwäche-Syndroms.

## Stand der Technik

**[0002]** Mit der Entdeckung des humanen T-Zell-lymphotropen Virus I (HTLV-I) als infektiösem Agens in Menschen wurde nachgewiesen, dass Retroviren Menschen infizieren können und das ätiologische Agens von Erkrankungen sein können. Nachdem HTLV-I nachgewiesen war, wurde ein zweites Retrovirus aus der gleichen Familie, HTLV-II, in einem nachgewiesenen Haarzellenleukämie-Stamm gefunden. Seit dieser Zeit wurden weitere humane Retroviren isoliert, die mit dem Lymphadenopathie-Syndrom (LAS) und/oder der erworbenen Immunschwäche bei erkrankten Menschen assoziiert sind. Verschiedene Retroviren sind aus Menschen mit AIDS (manchmal als HTLV-III bezeichnet) oder LAS (manchmal als LAV bezeichnet) isoliert worden, siehe beispielsweise Barre-Sinoussi et al., Science (1983) 220: 868–871 und Montagnier et al., Cold Spring Harbor Symposium (1984), im Druck; Vilmer et al., Lancet (1984) 1: 753, Popovic et al., Science (1984) 224: 497 und Gallo et al. Science (1984) 224: 500. Ein Vergleich von HTLV-III und LAV kann bei Ferrino et al. (1984), siehe oben, gefunden werden, siehe auch Klatzman et al., Science (1984) 225: 59–62, Montagnier et al., ebenda (1984) 63–66 und die darin als Überblick über das Fachgebiet angegebenen Literaturzitate. Eine allgemeine Diskussion der T-Zell-Leukämie-Viren kann bei Marx, Science (1984) 224: 475–477 gefunden werden. Levy et al. Science (1984) 225: 840–842, berichten über die Isolierung von ARV (AIDS-assoziierten Retroviren).

**[0003]** Für die Zwecke dieser Anmeldung werden diese Viren (HTLV-III, LAV und ARV) generisch als humanes T-Zell-lymphotropes Retrovirus (hTLR) bezeichnet. Auf Grund der Ähnlichkeit ihrer Morphologie, ihrer Serologie, der Optima der Reversen Transkriptase und ihrer Cytopathologie, die in den oben angegebenen Literaturziten beschrieben werden, kann gezeigt werden, dass die hTLRs zur gleichen Klasse gehören. Die Reverse Transkriptase bevorzugt beispielsweise  $Mg^{2+}$  und hat ein pH-Optimum von etwa 7,8.

## Offenbarung der Erfindung

**[0004]** Die Erfindung ist in den Ansprüchen definiert.

## Kurze Beschreibung der Zeichnungen

**[0005]** [Fig. 1](#) zeigt eine Restriktionskarte von provi-

raler DNA (ARV-2).

**[0006]** [Fig. 2](#) zeigt die Nucleotidsequenz von ARV-2(9B). Die Aminosäuresequenzen für die Produkte des gag-, des pol- und des env-Gens werden angegeben. Die U3-, die R- und die U5-Region der LTRs sind ebenfalls gekennzeichnet. Die Cap-Site ist die Position +1. Eine invertierte Sequenzwiederholung von 3 bp an den Enden des LTR, die TATA-Box in der Position -29, die zu dem 3'-Ende der tRNA<sup>lys</sup> komplementäre Sequenz in der Position 183 und das Polyadenylierungssignal in der Position 9174 sind unterstrichen. Die Überstriche zeigen die Aminosäuresequenzen an, die für Virusproteine ermittelt wurden. Die Nucleotide sind am Anfang jeder Zeile nummeriert, die Aminosäuren am Ende jeder Zeile.

## Ausführungsformen der Erfindung

**[0007]** Die hTLR-DNA-Sequenzen werden mindestens teilweise synthetisiert, und können für die Expression von Polypeptiden verwendet werden, die ein Vorläuferprotein, das einer weiteren Prozessierung durch Spaltung unterzogen wird, oder ein vollständiges reifes Protein oder ein Fragment davon sein können. Die kleinste interessierende Sequenz, die eine Sequenz liefert, die eine Aminosäuresequenz codiert, die imstande ist, spezifisch an einen Rezeptor zu binden, z.B. ein Immunglobulin, umfasst 21 bp, üblicherweise mindestens 45 bp, abgesehen vom Startcodon. Die Sequenz kann einen beliebigen größeren Teil eines Polypeptids oder das vollständige Polypeptid codieren, oder kann flankierende Regionen eines Vorläuferpolypeptids einschließen, sie kann aber Teile von Sequenzen oder vollständige Sequenzen einschließen, die zwei oder mehrere verschiedene reife Polypeptide codieren. Die Sequenz umfasst üblicherweise weniger als etwa 5 kbp, noch üblicher umfasst sie weniger als etwa 3 kbp.

**[0008]** Sequenzen von besonderem Interesse, die offene Leserahmen aufweisen ([Fig. 5](#)), definieren die Strukturgene für die gag-Proteine (p16 und p25) und das env-Protein. Es wird darauf hingewiesen, dass die obigen Sequenzen unter Erhalt von anderen Sequenzen gespleißt werden können, die in dem Retrovirus vorhanden sind, so dass das 5'-Ende der Sequenz nicht die N-terminale Aminosäure des Expressionsproduktes codieren muss. Die Spleißstelle kann sich am 5'-Terminus des offenen Leserahmens oder innerhalb des offenen Leserahmens befinden. Das Startcodon für das Protein muss nicht das erste Codon für Methionin sein, sondern kann das zweite oder das dritte Methionin sein, so dass der Einsatz der oben angegebenen vollständigen Sequenz zu einem verlängerten Protein führen kann. Im Fall des gag-Gens und des env-Gens findet jedoch eine proteolytische Prozessierung in den Zellen von Säugtieren statt, die eine Entfernung von zusätzlichen

Aminosäuren einschließen kann.

**[0009]** Für die Isolierung der verschiedenen Domänen kann das Provirus mit Restriktionsendonucleasen verdaut werden und können die Fragmente einer Elektrophorese unterzogen werden, und Fragmente, die eine geeignete Größe haben und mit einer Sonde, sofern verfügbar, einen Doppelstrang bilden, werden isoliert, in einen einen Doppelstrang bilden, werden isoliert, in einen Klonierungsvektor kloniert und aus dem Vektor herausgeschnitten. Die Fragmente können dann für die Expression prozessiert werden. Überflüssige Nucleotide können von einem oder beiden Enden unter Durchführung eines Verdau mit Bal31 entfernt werden. Durch Restriktionskartierung können geeignete Restriktionsstellen außerhalb oder innerhalb der codierenden Region lokalisiert werden. Primer-Reparatur oder in vitro-Mutagenese können eingesetzt werden, um einen Terminus festzulegen, für Insertionen, Deletionen, Punktmutationen oder Mehrfachmutationen oder dergleichen, wo Codons geändert werden können, entweder kryptisch oder unter Änderung der Aminosäure, Restriktionsstellen eingefügt oder entfernt werden können oder dergleichen. Wo Gene gekappt worden sind, können die verloren gegangenen Nucleotide unter Verwendung eines Adaptors ersetzt werden. Adaptoren sind besonders gut brauchbar für die Verbindung codierender Regionen, um den korrekten Leserahmen zu gewährleisten.

**[0010]** Die env-Domäne des hTLR-Genoms kann durch Verdau des Provirus mit EcoRI und KpnI und Reinigung eines Fragments aus 3300 Basenpaaren (bp) erhalten werden, wobei dieses Fragment etwa 400 bp der 5'-nicht-codierenden Region und etwa 200 bp der 3'-nicht-codierenden Region enthält. Drei verschiedene Methionine, die durch die Sequenz am 5'-Ende des offenen Leserahmens codiert werden, können als Startstelle für die Translation verwendet werden.

**[0011]** Der Verdau von proviralen Sequenzen mit SacI und EcoRV liefert ein Fragment von etwa 2300 bp, das die gag-Domäne und einen zweiten kleinen offenen Leserahmen in Richtung des 3'-Endes der gag-Region enthält. Die gag-Domäne hat eine Länge von etwa 1500 bp und codiert ein großes Vorläuferprotein, das prozessiert wird, wodurch Proteine von etwa 25000 (p25), 16000 (p16) und 12000 (p12) Dalton erhalten werden. Der Verdau mit SacI und BglII kann ebenfalls verwendet werden, um ausschließlich die gag-Domäne mit der p12-, der p25- und einem Teil der p16-Region zu erhalten.

**[0012]** Die hTLR-DNA-Sequenzen können mit isotopischen oder nicht-isotopischen Labels oder Markern markiert werden und können als DNA-Sonden für den Nachweis des Vorhandenseins von nativen hTLR-Nucleotidsequenzen in Proben verwendet wer-

den, von denen angenommen wird, dass sie diese Sequenzen enthalten.

**[0013]** Die folgenden Beispiele dienen der Veranschaulichung und stellen keine Beschränkung dar.

#### EXPERIMENTELLER TEIL

##### 1. Reinigung des AIDS-bezogenen Virus-2 (ARV-2) und Gewinnung von viraler RNA

**[0014]** HUT-78-Zellen, die mit ARV-2 infiziert sind (ATCC-Zugangsnummer CRL 8597, hinterlegt am 7. August 1984), wurden von Dr. Jay Levy, University of California, San Francisco, erhalten. Kulturen wurden zwei Wochen in RPMI-Medium mit 10% fetalem Kälberserum kultiviert. Die Kulturen wurden 1 h bei 4°C und 2000 U/min unter Verwendung eines Rotors SW-28 zentrifugiert. Das Pellet, das das Virus enthält, wurden in 10 mM Tris-HCl, pH 7,5, auf Eis erneut in Suspension gebracht. Das suspendierte Pellet wurde mit 10 µg DNase (Boehringer Mannheim) behandelt und auf einen linearen Saccharosegradienten geschichtet (15–50% in 10 mM Tris-HCl, pH 7,5, 1 mM EDTA, 20 mM NaCl). Der Gradient wurde 4 h bei 4°C und bei 34000 U/min in einem Rotor SW-41 geschleudert. Fünf 2,5 ml-Fractionen wurden gesammelt, und ein Aliquot jeder Fraktion wurde einer Elektrophorese in einem 1%-Agarose-5-mM-Methylquecksilberhydroxid-Gel unterzogen (Bailey und Davidson, Anal. Biochem. (1976) 70: 75–85), um zu ermitteln, welche Fraktion die 9 kb umfassende virale RNA enthielt. Die Fraktion, die die virale RNA enthielt, wurde in 10 mM Tris-HCl, pH 7,5, 1 mM EDTA auf 10 ml verdünnt und 2 h bei 4°C und 34000 zentrifugiert. Das Pellet wurde dann wieder in 20 mM Tris-HCl, pH 7,6, 10 mM EDTA, 0,1% SDS und 200 µg/ml Proteinase K suspendiert. Die Inkubation wurde 15 min bei Raumtemperatur durchgeführt. Das Gemisch wurde mit Phenol extrahiert, und die wässrige Phase wurde auf 400 mM NaCl eingestellt, wonach mit Ethanol gefällt wurde. Das Pellet wurde in Wasser suspendiert und bei –70°C aufbewahrt.

**[0015]** Für die Reinigung der viralen DNA aus dem Nucleinsäurepellet, das wie oben beschrieben erhalten wurde, wurde eine Probe einer Elektrophorese in einem niedrig schmelzenden 1%-Agarosegel, das 5 mM Methylquecksilberhydroxid enthielt, unterzogen. Nach der Elektrophorese wurde das Gel mit 0,1% Ethidiumbromid gefärbt, und die Nucleinsäurebanden wurde mit UV-Licht sichtbar gemacht. Der Bereich, der 9 kb entsprach, wurde aus dem Gel herausgeschnitten, wonach die Agarose 2 bis 3 min bei 70°C in drei Volumina 0,3 M NaCl, 10 mM Tris, pH 7,5, 1 mM EDTA geschmolzen wurde. Das Gemisch wurde mit einem gleich großen Volumen Phenol extrahiert. Die wässrige Phase wurde erneut mit Phenol extrahiert und mit Ethanol gefällt. Das Pellet wurde mit kaltem 95-%igem Ethanol gewaschen, an der Luft

getrocknet, erneut in Wasser suspendiert und bis zur Verwendung bei  $-70^{\circ}\text{C}$  aufbewahrt. 100 ml Kulturmedium ergaben 0,5 bis 1  $\mu\text{g}$  gereinigte DNA.

## 2. Synthese einer markierten homologen viralen Sonde

**[0016]** Eine  $^{32}\text{P}$ -markierte cDNA zu der durch Reinigung aus dem Gel erhaltenen viralen RNA wurde unter Verwendung von Primern mit Zufallssequenz (Kalbsthymusprimer) erzeugt, die wie in Maniatis et al., A Laboratory Manual, Cold Spring Harbor, NY, 1982, beschrieben, hergestellt wurde(n). Das Reaktionsgemisch enthielt 2  $\mu\text{l}$  0,5 M  $\text{MgCl}_2$ , 5  $\mu\text{l}$  0,1 M Dithiothreitol, jeweils 2,5  $\mu\text{l}$  10 mM dATP, 10 mM dGTP und 10 mM dTTP, 2,5  $\mu\text{l}$  Kälberthymusprimer (100  $\text{A}_{260}/\text{ml}$ ), 0,5  $\mu\text{l}$  virale DNA, 5  $\mu\text{l}$  Actinomycin D (200  $\mu\text{g}/\text{ml}$ ), 10  $\mu\text{l}$   $^{32}\text{P}$ -dCTP ( $> 3000 \text{ Ci}/\text{mmol}$ , 1  $\text{mCi}/\text{ml}$ ) und 1  $\mu\text{l}$  AMV-Reverse-Transkriptase (17 Einheiten/ $\mu\text{l}$ ) in einem 50- $\mu\text{l}$ -Reaktionsvolumen. Das Reaktionsgemisch wurde 1 h bei  $37^{\circ}\text{C}$  inkubiert. Die Sonde wurde von den freien Nucleotiden durch Gelfiltration unter Verwendung einer Sephadex-G50-Säule getrennt. Das Lückenvolumen wurde aufgefangen, NaCl wurde bis zum Erhalt einer Endkonzentration von 400 mM, einzelsträngige Träger-DNA bis zu einer Endkonzentration von 100  $\mu\text{g}/\text{ml}$  zugegeben, und die cDNA wurde mit Ethanol gefällt. Das Pellet wurde dann wieder suspendiert, und die Zählrate auf Grund von eingebautem  $^{32}\text{P}$  wurde ermittelt.

## 3. Nachweis von ARV-Sequenzen in Poly(A)<sup>+</sup>-RNA, hergestellt aus infizierten HUT-78-Zellen.

**[0017]** Poly(A)<sup>+</sup>-RNA wurde aus HUT-78-Zellen, die mit ARV-2, ARV-3 oder ARV-4 (drei verschiedene Isolate von drei verschiedenen AIDS-Patienten) infiziert waren, und aus nicht infizierten HUT-78-Zellen gewonnen. Die Poly(A)<sup>+</sup>-RNA wurde einer Elektrophorese auf 1%-Agarosegelen, die 5 mM Methylquecksilberhydroxid (Bailey und Davidson, siehe oben) enthielten, unterzogen und anschließend auf Nitrocellulosefilter übertragen und mit der homologen Sonde hybridisiert, die wie in Abschnitt 2 beschrieben hergestellt wurde. Die Hybridisierungen wurden in 50% Formamid,  $3 \times \text{SSC}$  bei  $42^{\circ}\text{C}$  durchgeführt. Waschschritte wurde bei  $50^{\circ}\text{C}$  in  $0,2 \times \text{SSC}$  durchgeführt. Für alle drei Proben infizierter HUT-78-Zellen wurde eine 9 kbp-Bande gefunden. Diese Bande fehlte in der Poly(A)<sup>+</sup>-RNA der nicht infizierten Zellen.

## 4. Nachweis von ARV-Sequenzen in infizierten und nicht infizierten HUT-78-Zellen.

**[0018]** DNA mit hohem Molekulargewicht (chromosomale DNA) wurde aus Kulturen von HUT-78-Zellen, die mit ARV-2 infiziert waren, und aus Kulturen von nicht infizierten HUT-78-Zellen unter Befolgung des Verfahrens von Luciw et al., Molec. and Cell Biol. (1984) 4: 1260–1269 gewonnen. Die DNA wurde mit

einem oder mehreren Restriktionsenzymen verdaut, einer Elektrophorese in 1%-Agarosegelen unterzogen und unter Befolgung der von Southern beschriebenen Methode (1975), siehe oben, auf Nitrocellulose durch Blotting übertragen. Die Blots wurden mit der  $^{32}\text{P}$ -markierten Sonde 36 h bei  $42^{\circ}\text{C}$  in einem Gemisch hybridisiert ( $10^6 \text{ cpm}/\text{Blot}$ ), das 50% Formamid,  $3 \times \text{SSC}$ , 10 mM Hepes, pH 7,0, 100  $\mu\text{g}/\text{ml}$  denaturierte Träger-DNA, 100  $\mu\text{g}/\text{ml}$  Hefe-RNA und  $1 \times \text{Denhardt's}$  enthielt. Die Filter wurden einmal bei Raumtemperatur in  $2 \times \text{SSC}$  und zweimal bei  $42^{\circ}\text{C}$  in  $0,2 \times \text{SSC}$ , 0,1% SDS gewaschen. Die Filter wurden an der Luft getrocknet und einem X-Omat-Film unter Verwendung einer Verstärkerfolie ausgesetzt.

**[0019]** Die zu ARV-2 homologe  $^{32}\text{P}$ -Sonde hybridisierte spezifisch mit zwei Banden der DNA von infizierten Zellen, die mit ScaI gespalten wurde. Diese Banden fehlten, wenn die DNA von nicht infizierten Zellen verwendet wurde, was darauf hinweist, dass die Sonde spezifisch mit infizierten Zellen hybridisiert, wahrscheinlich mit dem Provirus, das in die chromosomale DNA integriert ist. Das Molekulargewicht der Banden beträgt etwa 5 kb und 3 kb.

**[0020]** Um zu ermitteln, ob andere Enzyme die provirale Sequenz spalten, wurden weitere Spaltungen mit Restriktionsenzymen unter Verwendung von EcoRI, SphI oder KpnI oder doppelte Spaltungen, für die zwei der hier genannten Enzyme verwendet wurden, durchgeführt.

**[0021]** Die Southern-Blots zeigen spezifische Banden, die hybridisieren, wenn die DNA von infizierten Zellen verwendet wird. [Fig. 1](#) zeigt eine schematische Karte der Positionen von Restriktionsenzymstellen in der proviralen Sequenz und gibt Fragmentbereiche an.

## 5. Klonierung von proviraler ARV-2-DNA

**[0022]** Zell-DNA mit hohem Molekulargewicht aus infizierten HUT-78-Zellen wurde unter Befolgung des Verfahrens von Luciw et al., siehe oben, gewonnen. Die DNA wurde mit EcoRI verdaut, die einmal in dem Provirus schneidet, in einem Saccharosegradienten zentrifugiert, und Fraktionen, die 8–15 kb entsprechen, wurden vereinigt, einer Dialyse unterzogen und durch Ethanol fällung aufkonzentriert. Der vom  $\lambda$ -Bakteriophagen abgeleitete Klonierungsvektor EMBL-4 (Karn et al., Methods Enzymol (1983) 101: 3–19) wurde vollständig mit einem Gemisch der Restriktionsenzyme EcoRI, BamHI und Sall verdaut, und dann wurden durch Phenol-Chloroform-Extraktion die Proteine von der DNA entfernt, wonach die DNA mit kaltem Ethanol gefällt und in einem Ligationspuffer erneut suspendiert wurde. Die EMBL-4-Phagen-DNA und das Produkt des Verdauens der zellulären DNA mit EcoRI wurden gemischt und ligiert, und die resultierenden rekombinanten Phagen Genome wurden in vitro

verpackt. Nach der Phageninfektion von  $\lambda$ -sensitivem *E. coli* (DP50supF) wurden etwa 500 000 Phagenplaques auf Nitrocellulosefilter übertragen, die DNA wurde fixiert und die Filter wurden einem Screening mit einer homologen  $^{32}\text{P}$ -Sonde unterzogen, die wie in Abschnitt 2 beschrieben hergestellt wurde. Elf rekombinante Phagen von 500 000 Phagen reassozierten in dem zunächst durchgeführten "Double-lift"-Screening-Verfahren (Maniatis et al. Molecular Cloning, A Laboratory Manual, NY, 1982) mit der viralen cDNA-Sonde, diese Phagen wurden Plaque-gereinigt und in großen flüssigen Kulturen für die Gewinnung der rekombinanten DNA vermehrt. Plaque-gereinigte Phagen, die ARV-DNA enthalten, halten, wurden in flüssiger Kultur in *E. coli* DP50supF vermehrt; die Phagenpartikel wurden geerntet und in CsCl-Gradienten in Form von Banden erhalten, und die rekombinante Phagen-DNA wurde durch Phenol-extraktion und anschließende Ethanol-fällung gewonnen (Maniatis et al., siehe oben). 1  $\mu\text{g}$  gereinigte Phagen-DNA wurde mit Restriktionsenzymen verdaut, einer Elektrophorese auf 1%-Agarosegelen unterzogen und mit Ethidiumbromid unter Ultraviolettlicht sichtbar gemacht. Die DNA dieser Gele wurde auf Nitrocellulose übertragen, wo man sie mit der viralen cDNA-Sonde reassoziieren lässt.

**[0023]** Einer der 11 Phagen, der als  $\lambda$  ARV-2(9B) bezeichnet wird, wurde bei der ATCC am 25. Januar 1985 hinterlegt und hat die Zugangsnummer 40158 erhalten.  $\lambda$  ARV-2(9B) enthält eine Insertion, die aus der proviralen DNA voller Länge zusammen mit flankierenden Zellsequenzen besteht. Der Verdau von  $\lambda$  ARV-2(9B)-DNA mit *SacI* lieferte Fragmente der viralen DNA von 3,8 kb und 5,7 kb. Der *EcoRI*-Verdau von  $\lambda$  ARV-2(9B) führte zu virushaltigen DNA-Stücken bei 6,4 kb und 8,0 kb; der doppelte Verdau mit *SacI* und *EcoRI* ergab Fragmente der viralen DNA bei 3,8 kb und 5,4 kb. Dieses Muster stimmt mit dem Muster eines Provirus überein, das an zelluläre DNA gebunden ist.

**[0024]** Zusätzlich zu  $\lambda$  ARV-2(9B) wurden erhalten: (1) ein Phage ( $\lambda$  ARV-2(8A)), der die linke Hälfte des viralen Genoms von der *EcoRI*-Stelle in der viralen DNA sich erstreckend bis in die flankierende Zell-DNA aufweist, und (2) ein Phage ( $\lambda$  ARV-2(7D)), der die rechte Hälfte des viralen Genoms von der *EcoRI*-Stelle in der viralen DNA sich erstreckend in die flankierende Zell-DNA aufweist. Die Bakteriophagen  $\lambda$  ARV-2(7D) (rechts) und  $\lambda$  ARV-2(8A) (links) wurden bei der ATCC am 26. Oktober 1984 hinterlegt und haben die Zugangsnummern 40143 bzw. 40144 erhalten.

## 6. Polymorphismus

**[0025]** Zur Messung des Grades der Verwandtschaft unabhängiger ARV-Isolate wurden die mit Restriktionsenzymen erhaltenen Spaltprodukte der

DNA von HUT-78-Zellen, die mit ARV-3 und ARV-4 infiziert waren, mit der Sonde analysiert, die aus klonierter ARV-2-DNA hergestellt wurde. Der *SacI*-Verdau von ARV-3 DNA ähnelte dem Verdau von ARV-2, wohingegen die mit *HindIII* erhaltenen Spaltprodukte voneinander abweichende Muster zeigten. Der *SacI*-Verdau und der *PstI*-Verdau von ARV-4-DNA unterschied sich vom entsprechenden Verdau von ARV-2 DNA. Die Intensität der infolge der Reassoziations erhaltenen Signale, die mit der ARV-3- und der ARV-4-Probe erhalten wurden, war wesentlich schwächer (in etwa 10-fach schwächer) als das Signal für ARV-2 DNA, was wahrscheinlich daran lag, dass in den ARV-3- und ARV-4-Kulturen weniger Zellen infiziert wurden. Die virusspezifischen DNA-Fragmente, die durch die Behandlung von ARV-3- und ARV-4-DNA mit *SacI* erhalten wurden, hatten ein Gesamtgewicht von 9,0–9,5 kbp, was einen Wert darstellt, der dem Wert von ARV-2 ähnelt und mit der Größe der RNA-Genome im Einklang steht.

## 7. Sequenzierung von proviraler DNA

**[0026]** Fragmente oder Subfragmente von ARV-2-DNA aus dem  $\lambda$ -Phagen 9B wurden hergestellt und unter Anwendung herkömmlicher Verfahren (Maniatis et al, siehe oben) in M13 kloniert. Die Sequenzierung wurde gemäß Sanger et al, Proc. Natl. Acad. Sci. USA (1977) 74: 5463, unter Verwendung des Universalprimers für M13 oder von chemisch synthetisierten Primern, die komplementär zur ARV-2-Sequenz sind, durchgeführt. Die Sequenz wird in Fig. 2 gezeigt.

## Patentansprüche

1. Verfahren zur Herstellung eines isolierten DNA-Polynucleotids, das ein Fragment von mindestens 21 bp aus der gag-Region oder der env-Region der ARV-2-Sequenz gemäß Fig. 2 umfaßt, wobei das DNA-Polynucleotid mindestens teilweise chemisch synthetisiert wird.

2. Verfahren nach Anspruch 1, wobei das DNA-Polynucleotid ein Fragment von mindestens 21 bp aus der gag-Region der ARV-2-Sequenz gemäß Fig. 2 umfaßt.

3. Verfahren nach Anspruch 1, wobei das DNA-Polynucleotid ein Fragment von mindestens 21 bp aus der env-Region der ARV-2-Sequenz gemäß Fig. 2 umfaßt.

4. Verfahren nach einem der vorhergehenden Ansprüche, wobei das DNA-Polynucleotid markiert ist.

5. Verfahren nach Anspruch 4, wobei das DNA-Polynucleotid mit einem isotopischen Marker markiert ist.

6. Verfahren nach Anspruch 4, wobei das DNA-Polynucleotid mit einem nicht-isotopischen Marker markiert ist.

Es folgen 10 Blatt Zeichnungen

Fig. 1

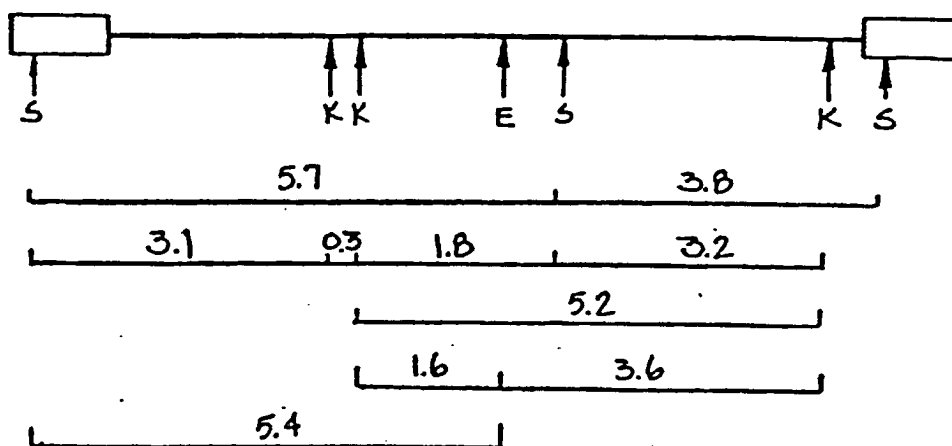








Fig. 3

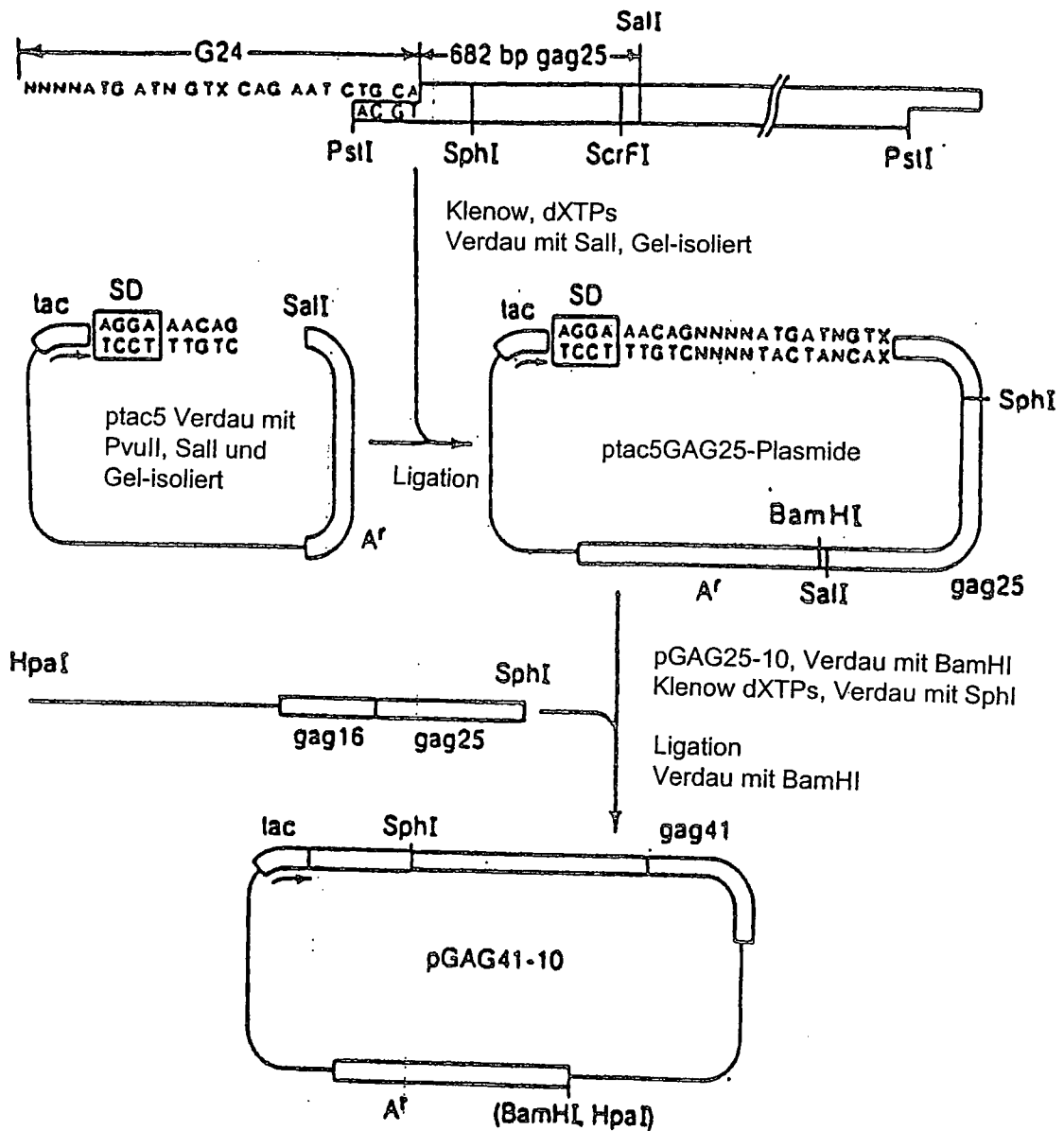


Fig. 4

\* Met Pro Leu Val 141  
ATG CCT A T C G T A

ptac 5 Promotor

740 G I A S n L e u G I n G I y G I n M e t V a I M I S G I n A I e I I e S e r P r o A r g T h r L e u A s n A I e T r p V a I L y s V a I Y e I G I u G I L y s A I o P h e S e r P r o G I u V a I I I e P r o M e t P h e S e r A l e L e u 181  
C A S A A T C T G C A G G G C A A A T G G T A C A T C A G G C C A T A T C A C C T A G A C T T T A A A T G C A T G G T A A A G T A G T A G A G A A A G G C T T C A G C C C A G A A G T A T A C C C A T G T T T T C A G C A T T A

868 S e r G I u G I y A I e T h r P r o G I n A s p L e u A s n T h r M e t L e u A s n T h r V a I G I y G I y H I S G I n A I e A I e M e t G I n M e t L e u L y s G I u T h r I I e A s n G I u G I u A I e A I e I u T r p A s p A r g V a I 221  
T C A G A G G A G C C A C C C A C A A S A T T A A C A C C A T G C T A A C A C A G T G G S G G A C A T C A A G C A C C A T G C A A T G T T A A A G A G A C T A T C A A T S A S G A A G C T E C A G A A T G G G A T A G A G T G

988 M I s P r o V a I M I s A I e G I y P r o I I e A I e P r o G I y G I n M e t A r g G I u P r o A r g G I y S e r A s p I I e A I e G I y T h r S e r T h r L e u G I n G I u G I n I I e G I y T C M e t T h r A s n A s n P r o P r o 261  
C A T C C A G T G C A T G C A G G C C T A T T G C A C C A G G C C A A A T S A G A S A C C A A G G G A A C T G A C A T A G C A S A A C T A C T A G T A C C C T T C A G G A A C A A A T A G G A T G C A T G C A C A A T A A T C C A C C T

1108 I I e P r o V a I G I y G I u I I e T y r L y s A r g T r p I I e I I e L e u G I y L e u A s n L y s I I e V a I A r g M e t T y p S e r P r o T h r S e r I I e L e u A s p I I e A r g S I n G I y P r o L y s G I u P r o P h e A r g A s p 301  
A T C C C A G T A G G A G A A T C T A T A A A A G A T G G A T A A T C C T G G A T T A A A T A A A T A G T A A G A A T G T A G C C C T A C C A G C A T T C T G G A C A T A A G A C A R A G G A C C A A A G G A A C C C T T T A G A G A I

1228 T y r V a I A s p A r g P h e T y r L y s T h r L e u A r g A I o G I u G I n A I e S e r G I n A s p V a I L y s A s n T r p M e t T h r G I u T h r L e u L e u V a I G I n A s n A I e A s n P r o A s p C y s L y s T h r I I e L e u L y s 341  
T A T G T A G A C C G G T T C T A T A A A A C T C T A A G A G C C G A C A A G C T T C A C A G S A T G T A A A A A A T T G G A T G A C A G A A C C T T G T T G G T C C A A A T G C A A C C C A G A T T G T A R A G A C T A T T T A A A A

1348 A I e L e u G I y P r o A I e A I e T h r L e u G I u M e t M e t T h r A I a C y s G I n G I y V e I G I y G I y P r o G I y H I s L y s A I e A r g V a I L e u S t o p S t o p  
G C A T T G G G A C C A G C A C T A C A C T A S A A G A A T G A T G A C A G C A T G I C A G G A G T G G G G G A C C C G G C A T A A R C A A G A G T T T T G T A G T A G

ptac 5



## ARV gag p16 – Synthetische Teile A und B

```

5' MetGlnAArgLYAsnPhaArg3'AsnGlnAArgLYsThrcValLYsCYsPheAsnCYsGlyLYs
TATTATGCAAAGAGGTAACCTTCAGGAAICAAAGAAGACCCTTAAGTGTTCACACTGTGGTAAG
ATAATACGTTTCTCCATTGAAGTCC5'TAGTTCTTTCTGGCAATTCACAAAGTTGACACCCATTC
3' 10 mnl1, 23 hinf1,
63 6LUGLYHISILEALALYsAsnCYsAArgAlaPrcAArgLYsLYsAlaCYsIcPcAArgCYsGly
GAAGGTCACATCGCTAAGAACTGTAGAGCTCCAAAGAAAGATAGGCTTGTGGAGATGTGGT
CTTCCAGGTAGCGGATCTTGACATCTCGAGGTCTTCTCCGAAACAACCTCTACACCCA
76 dde1, 88 ban2 hglA hglJ11 sao1 sdu1, 89 alu1,
123 ArgGLUGLYHISGlnMetLYsAspCYsThrcGluAArgGlnAlaAsnPhaLeuGLYsIle
AGAGAAGGTCACCAAATGAAGGACTGTACCGAAAGACAAGCTTCTTTGGGTAAGATC
TCTCTTCCAGTGGT5'TACTTCCCTGA3'CATGGCTTCTGTTCGATTGAAGAACCCTTCTAG
129 bstE2, 131 hph, 148 rsa1, 161 alu1, 178 bg111 xho2, 179
sau3a,
183 TrpPrcSerTYrLYsGlyAArgPrcGlyAsnPhaLeuGlnSerAArgPrcGluPrcThrcAla
TGC3'CCATCTTACAAAGGGTAGACCAGGTAACCTTCTTGCAATCCAGACCAGAACCAACCGCT
A5'CCGGTAGAAATGTTCC3'ATCTGGTCCATTGAAGAACGTTAGGCTCTGGTCTTGGTTGGCGA
183 ball ofr1 hae1, 184 hae111, 199 acc1, 204 apy1 eoor11 so
rf1,
243 PrcPrcGluGluSerPheAArgPheGlyGluGluLYsThrcThrcPrcSerGlnLYsGlnGlu
CCACCTGAAGAAAGTTCAGGTTCCGGTGAAGAAAAGACCACCCCATCTCAAAGCAAGAA
GGTGGACTTCTTTCAAAGTCCAAGCCA3'CTTCTTTCTGGTGGGGTAGAGTTTTCGTTCTT
249 mbo11, 267 hph, 270 mbo11,
303 PrcIleAspLYsGluLeuTYcPrcLeuThrcSerLeuAArgSerLeuPheGlyAsnAspPrc
CCAATCGACAAGGAAATGTACCCATTGACCTCTTTGAGATCCTTGTTCGGTAACGATCCC
GGTTAGCTGTTCCCTTAA3'CTAGGGTA5'ACTGGAGAAACTCTAGGAACAAGCCCATTTGCTAGGG
307 taq1, 320 rsa1, 331 mnl1, 339 xho2, 340 sau3a, 357 sau3a
, 361 mnl1, 362 avat1 xho1,
363 SerSerGlnOP AM
TCGAGCCAATGATAG
AGCTCGGTTACTATCAGCT
363 taq1, 377 acc1 hind11 sal1

```

Fig. 6



Fig. 8-1

Nucleotid- Positionen, bezogen auf Figur 2	1	<p>MetIleAspLysAlaGlnGluGluHisGluLysTyrHisSerAsnTrp                  AGGXAAACAG:::ATGAT:GA:AAGGCACAAGAAAGACATGAGAAATATCACAGTAATTG                  TCCXTTGTG:::TACTA:CT:TTCCGTGTTCTTCTTGTACTCTTTATAGTGTCAATTAAC</p> <p>32 mbo11, 38 nla111,</p>
3820	62	<p>ArgAlaMetAlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSer                  AGAGCCATGGCTAGTGAITTTAACCTGCCACCTGTAGTAGCAAAAAGAAATAGTAGCCAGC                  TCTCGGTACCGATCACTAAAATTGGACGGTGGACATCATCGTTTTCTTTATCATCGGTC</p> <p>66 nco1, 67 nla111, 118 nspBII pvu11, 119 alu1,</p>
3880	122	<p>CysAspLysCysGlnLeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGly                  TGTGATAAATGTCAAGCTAAAAGGAGAAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGA                  ACACTATTTACAGT_CGATTTTCTCTTTCGGTACGTACCTGTT_CATCTGACATCAGGTCTCT</p> <p>135 alu1, 151 nla111, 152 ns11 ava3, 155 nla111, 164 acc1, 176 apy1 bstXI ecor11 scrF1,</p>
3940	182	<p>IleTrpGlnLeuAspCysThrHisLeuGluGlyLysIleIleLeuValAlaValHisVal                  ATATGGCAACTAGATTGTACACATCTAGAAGGAAAAATTATCCTGSTAGCAGTTCATGTA                  TATACCGTTGATCTAACATGIGTAGATCTTCTTTTTTAATAGGACCATCGTCAAGTACAT</p> <p>198 rsa1, 205 xba1, 223 apy1 ecor11 scrF1, 236 nla111,</p>
4000	242	<p>AlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyr                  GCCAGTGGATATATAGAAGCAGAAGTATTCCAGCAGAGACAGGGCAGGAAAACAGCATAT                  CGGTACCTATATCTTCGTCTTCAATAAGGTCGTCTCTGTCCCCTCTTTGTCTGATATA</p> <p>263 xmn1</p>
4060	302	<p>PheLeuLeuLysLeuAlaGlyArgTrpProValLysThrIleHisThrAspAsnGlySer                  TTTCTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACAGCAATGGCAGC                  AAAGAGAATTTAATCGTCCTTCTACCGGTCAITTTGTTATGTATGTCTGTTACCGTCCG</p> <p>321 mbo11, 326 bal1 cfr1 hae1, 327 hae111, 357 bbv fnu4h1,</p>
4120	362	<p>AsnPheThrSerThrThrValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPhe                  AATTTCACCACTACTACGGTTAAGGCCCGCTGTTGGTGGCAGGGATCAAGCAGGAATTT                  TTAAGTGGTCAATGATGCCAATTCGGCGGACAACCAACCCGTCCCTAGTTCTGTCCTTAAA</p> <p>366 hph, 371 sca1, 372 rsa1, 385 hae111, 386 fnu4h1 nsb11, 405 bin1, 406 dpn1 sau3a,</p>
4180	422	<p>GlyIleProTyrAsnProGlnSerGlnGlyValValGluSerMetAsnAsnGluLeuLys                  GGCATTCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAATGAATTAAG                  CCGTAAGGGATGTTAGGGGTTTCAGTTCCTCATCATCTTAGATACTTATTACTTAAATTC</p> <p>423 bsm1, 458 hinf1,</p>
4240	482	<p>LysIleIleGlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAla                  AAAATTATAGGACAGGTAAGAGATCAGGCTGAACACCTTAAGACAGCAGTACAAATGGCA                  TTTAATATCCTGTCCATTCTTAGTCCGACTTGTGGAAATCTGTCTCATGTTTACCGT</p> <p>503 dpn1 sau3a, 518 afl111, 530 rsa1,</p>
4300	542	<p>ValPheIleHisAsnPheLysArgLysGlyGlyIleGlyGlyTyrSerAlaGlyGluArg                  GTATTCATCCACAATTTTAAAAGAAAAGGGGGATTGGGGGATACAGTGCAGGGGAAAGA                  CATAAGTAGGTGTTAAAATTTTCTTTTCCCCCTAACCCCTATGTCACGTCGCCCTTTCT</p> <p>547 fok1, 557 aha111,</p>
4360	602	<p>IleValAspIleIleAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleThrLys                  ATAGTAGACATAATAGCAACAGACATACAAACTAAAGAACTACAAAAGCAAATTACAAAA                  TATCATCTGTATTATCGTTGTCTGTATGTTGATTTCTTGATGTTTTCGTTAATGTTTT</p> <p>605 acc1,</p>
4420	662	<p>IleGlnAsnPheArgValTyrTyrArgAspAsnLysAspProLeuTrpLysGlyProAla                  ATCAAATTTTTCGGGTTTATTACAGGGACAACAAAGATCCCTTTTGAAGGACCAGCA                  TAAGTTTAAAAGCCCAATAATGTCCTGTTGTTCTAGGGGAAACCTTCTCTGGTCTGT</p> <p>697 xho2, 698 dpn1 sau3a, 713 asu1 ava2,</p>

Fig. 8-2

4480 722 LysLeuLeuTrpLysGlyGluGlyAlaValValIleGlnAspAsnSerAspIleLysVal  
AAGCTTCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTA  
TTCGAAGAGACCTTCCACTTCCCGTCATCATTATGTTCTATTATCACTGTAT.TTTCAT  
722 hind111, 723 alu1, 737 hph,

4540 782 ValProArgArgLysAlaLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAsp  
GTGCCAAGAAGAAAAGCAAAAATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGAT  
CACGGTTCCTTTTCGTTTTAGTAATCCCTAATACCTTTTGTCTACCGTCCACTACTA  
789 mbo11, 833 hph,

4600 842 CysValAlaSerArgGlnAspGluAspAM  
TGTGTGGCAAGTAGACAGGATGAGGATTAGTCGACGGAATTCTTTAGTAAACACC  
ACACACCCTTCATCTGTCCTACTCCTAATCAGCTGCCCTTAAGAAATCATTGTGG  
852 acc1, 859 fok1, 863 mn11, 871 acc1 hind11 sal1, 872 taq1  
, 878 ecor1,