

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
19. Juli 2001 (19.07.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/51631 A2

(51) Internationale Patentklassifikation⁷: C12N 15/12,
5/10, 15/11, C07K 14/47, C12Q 1/68, A61K 48/00, 35/14

(74) Anwalt: VOSSIUS & PARTNER; Siebertstrasse 4,
D-81675 München (DE).

(21) Internationales Aktenzeichen: PCT/EP01/00362

(22) Internationales Anmeldedatum:
12. Januar 2001 (12.01.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
100 01 169.1 13. Januar 2000 (13.01.2000) DE
100 10 188.7 2. März 2000 (02.03.2000) DE

(71) Anmelder und

(72) Erfinder: RESKE-KUNZ, Angelika [DE/DE]; Am
Römerberg 22, 55270 Essenheim (DE). ROSS, Xiaolan
[DE/DE]; Bastion Martin 12, 55131 Mainz (DE). ROSS,
Ralf [DE/DE]; Bastion Martin 12, 55131 Mainz (DE).
BROS, Matthias [DE/DE]; Oberer Laubenheimer Weg
24, 55131 Mainz (DE).

(81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches Patent (AT, BE, CH, CY, DE, DK,
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR),
OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML,
MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu
veröffentlichen nach Erhalt des Berichts

[Fortsetzung auf der nächsten Seite]

(54) Title: REGULATORY SEQUENCE FOR THE SPECIFIC EXPRESSION IN DENDRITIC CELLS AND USES THEREOF

(54) Bezeichnung: REGULATORISCHE SEQUENZ ZUR SPEZIFISCHEN EXPRESSION IN DENDRITISCHEN ZELLEN
UND DEREN ANWENDUNGEN

(57) Abstract: The invention relates to regulatory sequences which impart a specific expression in dendritic cells. The regulatory sequences are isolated from the human fascin gene and also comprise, for example, promoter sequences. The invention also relates to recombinant nucleic acid molecules and vectors, which contain the regulatory sequences, and to preferred embodiments of the recombinant nucleic acid molecules and vectors, which code the antigens or immunoregulatory proteins. The invention additionally relates to host cells, which contain the recombinant nucleic acid molecules or vectors, and to methods for the production thereof. Additional embodiments relate to in-vitro methods for stimulating T cells and for producing T cell-stimulating dendritic cells, and to their formulation as medicaments. Additional medicaments are described which essentially relate to DNA vaccines and to gene-therapeutic medicaments, for example, for the immunization against and for the treatment of infectious diseases, tumors, allergies, Creutzfeldt-Jakob plaques or Alzheimer plaques. Additional inventive medicaments can be used for the targeted modulation of immune responses that is imparted by dendritic cells, for example, for treating autoimmune diseases or transplant rejection. Finally, the invention relates to different uses of the regulatory sequences.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft regulatorische Sequenzen, die eine spezifische Expression in dendritischen Zellen vermitteln. Die regulatorischen Sequenzen sind vom menschlichen Fascinogen isoliert und umfassen beispielsweise auch Promotorsequenzen. Beschrieben sind ferner rekombinante Nucleinsäuremoleküle und Vektoren, die die regulatorischen Sequenzen enthalten, sowie bevorzugte Ausführungsformen der rekombinanten Nucleinsäuremoleküle und Vektoren, die Antigene oder immunregulatorische Proteine codieren. Die Erfindung betrifft weiterhin Wirtszellen, die die rekombinanten Nucleinsäuremoleküle oder Vektoren enthalten sowie Verfahren zu deren Herstellung. Weitere Ausführungsformen betreffen in vitro-Verfahren zur Stimulierung von T-Zellen und zur Herstellung T-Zell-stimulierender dendritischer Zellen sowie zu deren Formulierung als Arzneimittel. Weitere Arzneimittel sind beschrieben, die sich im wesentlichen auf DNA-Vakzine und gentherapeutische Arzneimittel beziehen, beispielsweise zur Immunisierung gegen und Behandlung von Infektionskrankheiten, Tumoren, Allergien, Creutzfeldt-Jakob-Plaques oder Alzheimer-Plaques. Weitere erfindungsgemasse Arzneimittel können zur gezielten, durch dendritische Zellen vermittelten Modulation von Immunantworten verwendet werden, beispielsweise zur Behandlung von Autoimmunerkrankungen oder Transplantatabstossung. Es werden ferner verschiedene Verwendungen der regulatorischen Sequenzen beschrieben.

WO 01/51631 A2



Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

Regulatorische Sequenz zur spezifischen Expression in dendritischen Zellen und deren Anwendungen

Die vorliegende Erfindung betrifft regulatorische Sequenzen, die eine spezifische Expression in dendritischen Zellen vermitteln, rekombinante Nucleinsäuremoleküle, Vektoren, Wirtszellen sowie Verfahren zur Herstellung der Wirtszellen. Ferner betrifft die Erfindung Verwendungen der regulatorischen Sequenzen sowie Arzneimittel.

Dendritische Zellen (DC) spielen als Antigen-präsentierende Zellen (APC) eine Schlüsselrolle bei der Mobilisierung der spezifischen Immunabwehr. Als einzige Zellen sind dendritische Zellen in der Lage, ruhende, für die Abwehr in Bereitschaft stehende, sogenannte naive T-Lymphocyten effizient zu aktivieren. Dabei können sie sowohl $CD4^+$ T-Helferzellen als auch $CD8^+$ cytotoxische T-Zellen induzieren. Dendritische Zellen steuern daher gleichermaßen die humorale Antikörperdominierte wie die zelluläre Immunantwort. Für eine effiziente Immunabwehr von bakteriellen, viralen und parasitären Krankheitserregern sowie von Tumorzellen sind dendritische Zellen unabdingbar. Auch an pathologischen Entgleisungen des Immunsystems wie Autoimmunkrankheiten und Allergien sind dendritische Zellen ursächlich beteiligt. Es ist daher ein herausragendes medizinisches Anliegen, die Funktionen von dendritischen Zellen therapeutisch nutzbar zu machen.

Dendritische Zellen in unterschiedlichen Differenzierungsstadien besitzen unterschiedliche Funktionskompetenz. Drei diskrete Reifephasen lassen sich aus funktioneller Sicht definieren. 1.) Eine Überwachungsfunktion wird von jungen, unreifen dendritischen Zellen ausgeübt. Junge dendritische Zellen sind in nahezu allen Organen des Körpers und besonders in den Epithelien, die den Körper nach außen abgrenzen, als Wächterzellen strategisch positioniert. Ihre Aufgabe besteht in der Aufnahme von Fremdstoffen (Antigenen) in löslicher oder partikulärer Form, in der Antigenprozessierung und der Peptidbeladung von MHC-Molekülen. 2.) Als Peptidtransporter wandern die dendritischen Zellen in der sich anschließenden migratorischen Phase in die drainierenden Lymphknoten, um sich dort in den T-Zellarealen anzusiedeln. 3.) In diesem Bereich kommt das immunogene Potential

der dendritischen Zellen zum Tragen. Die dendritischen Zellen präsentieren das prozessierte Peptid zusammen mit MHC-Molekülen T-Lymphocyten mit entsprechender Rezeptorspezifität. Dabei handelt es sich um naive T-Lymphocyten, d.h. um Zellen, die zuvor noch nicht mit dem Antigen in Kontakt gekommen waren und daher eine hohe Aktivierungsschwelle besitzen. Die direkte Zell-Zell-Interaktion zwischen den dendritischen Zellen und den T-Lymphocyten resultiert in der Aktivierung, der Expansion und damit der funktionellen Rekrutierung der peptidspezifischen T-Lymphocyten.

Dendritische Zellen sind aufgrund ihrer zentralen Funktion bei der Vermittlung von Immunantworten ein wesentlicher Ansatzpunkt für Schutzimpfungen, ein Aspekt, der im folgenden etwas ausführlicher diskutiert werden soll.

Die klassische Schutzimpfung, die vor allem auf der Verabreichung von attenuierten Erregern basiert, ist mit einer Reihe von Nachteilen verbunden. Dies betrifft beispielsweise das potentielle Risiko einer Infektion, falls der Erreger seine Human-Pathogenität wiedererlangen sollte (McKee, *Am.J.Trop.Med.Hyg.* 36 (1987), 435-442), Probleme in Zusammenhang mit der Produktion und Aufbewahrung der Vakzine (Rabinovich, *Science* 265 (1994), 1401; Fynan, *Proc. Nat. Acad. Sci. USA* 90 (1993), 11478) sowie eine mögliche Behinderung der Antigenpräsentation durch immunmodulatorische Eigenschaften des Erregers (Levine et al., in *New Generation Vaccine*, Woodrow, G. and Levine, M.M., eds., Marcel Decker, New York, (1990), 269-287). In Anbetracht dessen, kristallisiert sich immer mehr heraus, daß die Technik der DNA-Vakzinierung hierzu eine vielversprechende Alternative darstellt. Bei der DNA-Vakzinierung, die erstmals von Ulmer (*Science* 259 (1993), 1745-1749) gezeigt wurde, wird "nackte" DNA in den Körper injiziert oder anderweitig appliziert, die eine protektive Immunantwort auslöst (für eine Übersicht siehe Lai und Bennet, *Crit. Rev. Immunol.* 18 (1998), 449-484). Die DNA enthält Sequenzen, die für ein Antigen eines Erregers oder Tumors oder für ein Allergen codieren und die unter der Regulation eines ubiquitär funktionierenden Promotors stehen. Die applizierte DNA wird von Zellen des umliegenden Gewebes aufgenommen, das Antigen wird exprimiert und über MHC-Proteine an der Zelloberfläche präsentiert.

Im Verlauf der Entwicklung von DNA-Vakzinierungsstrategien stellte sich zunehmend heraus, daß hierbei eine gezielte Adressierung von dendritischen Zellen vielversprechende Vorteile bieten könnte. Das läßt sich beispielsweise aus

Immunisierungsversuchen ableiten, bei denen dendritische Zellen in vitro mit Antigen beladen und reinjeziert wurden (WO 94/02156). So erzeugten etwa murine dendritische Zellen, die in vitro mit synthetischen Peptiden (Mayordomo et al., *Nat. Med.* 1 (1995), 1297-1302; Celluzzi et al., *J. Exp. Med.* (1996), 283-87), Säure-eluierten Peptiden von Tumorzellen (Zitvogel et al., *J. Exp. Med.* 183 (1996), 87-97) und sogar intakten Tumorproteinen (Paglia et al., *J. Exp. Med.* 183 (1996), 317-22) gepulst wurden, in vivo protektive Antworten cytotoxischer T-Lymphocyten (CTL) gegen den Tumor. Humane dendritische Zellen, die aus peripheren, mononucleären Blutzellen (PBMC) gesunder Donoren gezogen wurden, können ebenfalls CTL-Antworten gegen Tumor-Antigene erzeugen (van Elsas et al., *Eur. J. Immunol.* 26 (1996), 1683-1689; Falo et al., *Nature Med.* 1 (1995), 649-653) und in vitro kultivierte dendritische Zellen wurden bereits erfolgreich in klinischen Studien für die Therapie von Tumoren eingesetzt (Hsu et al., *Nature Med.* 2 (1996), 540-544). Auch dendritische Zellen und Zelllinien dieses Zelltyps, die in vitro mit Antigen-codierenden Expressionsplasmiden transfiziert wurden, können nach adoptivem Transfer in vivo Antigen-spezifische Immunantworten induzieren (Manickan et al., *J. Leucocyte Biol.* 61 (1997), 125-132; Timares-Lebow et al., *J. Invest. Dermatol.* 109 (1997), 266; Tüting et al., *Eur. J. Immunol.* 27 (1997), 2702-2707). Dendritische Zellen, in die mRNA aus Tumorzellen eingeführt wurde, wurden ebenfalls bereits erfolgreich für Vakzinierungen eingesetzt (Boczkowski et al., *J. Exp. Med.* 184 (1996), 465-72). Aus diesen und anderen Beobachtungen lassen sich folgende Vorteile von DNA-Vakzinierungen unter gezielter Antigenexpression in dendritischen Zellen ableiten. So ist beispielsweise bekannt, daß die Präsentation eines Antigens durch nicht professionelle APCs, welche nicht die notwendigen costimulatorischen Signale für eine effiziente T-Zell Stimulation zur Verfügung stellen, zu einer mangelhaften Reaktivität bzw. Anergie der T-Zellen führt. Die Restriktion der Expression des Antigens auf reife dendritische Zellen könnte daher zu einer verstärkten Immunantwort führen. Desweiteren könnte es beispielsweise zusätzlich zur Expression des Antigens möglich sein, durch Co-Transfektion weitere, immunmodulatorische Proteine in DC zu exprimieren und dadurch regulatorisch in die Immunantwort einzugreifen. Die Beschränkung der Expression dieser Mediatoren (z.B. Cytokine, costimulatorische Membranproteine) auf transfizierte dendritische Zellen würde es erlauben, sehr spezifisch die erwünschte

Immunantwort zu beeinflussen, ohne die bei systemischer Gabe dieser Mediatoren zu erwartenden Nebenwirkungen befürchten zu müssen.

Daraus ergibt sich ein Bedarf an Promotoren oder regulatorischen Sequenzen, die eine spezifische Expression in dendritischen Zellen vermitteln. Damit ließen sich beispielsweise DNA-Vakzinierungskonstrukte herstellen, die nach breitgestreuter Verabreichung, wie beispielsweise durch intramuskuläre Injektion oder bioballistischen Transfer in die Haut, nur in dendritischen Zellen, die in den peripheren Geweben relativ schwach repräsentiert sind, das Antigen oder immunmodulatorische Proteine exprimieren. Es besteht zumindest ein Bedarf an Promotoren oder regulatorischen Sequenzen, die in solchen Geweben eine spezifische Expression für dendritische Zellen vermitteln, in denen die Antigenbeladung der dendritischen Zellen stattfindet. Dies sind in erster Linie das Blutssystem, das Hautgewebe, mukosale Gewebe und die Muskeln. Weitere mögliche Anwendungen solcher Promotoren oder regulatorischer Sequenzen sind in vitro-Transfektionen von heterogenen Zellpopulationen, wobei das Antigen exklusiv in reifen dendritischen Zellen exprimiert würde.

Im Stand der Technik ist bereits ein Promotor mit Spezifität für dendritische Zellen beschrieben worden (Brocker, J. Exp. Med. 185 (1997), 541-550). Dieser Promotor wurde aus dem Maus-CD11c-Gen isoliert und konnte bei einem Transgen (MHC-Klasse II-I-E-Protein) in transgenen Mäusen eine spezifische Expression vermitteln. Jedoch erfüllt der in dieser Veröffentlichung präsentierte Promotor in einigen Punkten nicht die Anforderungen des oben skizzierten Promotors zur spezifischen Expression in dendritischen Zellen.

So konnte die Spezifität für dendritische Zellen nur für Milz- und Thymusgewebe gezeigt werden. Dagegen exprimierte dieser Promotor das Transgen auch in einem Teil der peritonealen Makrophagen. Desweiteren ist bekannt, daß CD11c nur in einer Subpopulation der dendritischen Zellen exprimiert wird (Rich et al., Poster Nr. D6, 5th International Symposium on Dendritic Cells in Fundamental and Clinical Immunology, Pittsburgh, Penn. U.S.A., 23.-28. Sept. 1998). Schwerwiegender ist jedoch, daß die gezeigte Promotoraktivität sich auf das Maussystem bezieht. Da der gesuchte Promotor für humanmedizinische Anwendungen zur Verfügung stehen soll, bietet der CD11c-Promotor keine Lösung der Problemstellung, da nicht zu

erwarten ist, daß der Promotor im Mensch dieselbe Gewebe- oder Zelltypspezifität besitzt wie in der Maus. Ein Beispiel, das eine unterschiedliche Zelltypspezifität zwischen Maus und Mensch zeigt, ist das Chemokinen DC/B-CK. In der Maus wird dieses Gen nur in dendritischen Zellen und aktivierten B-Zellen aber nicht in Makrophagen exprimiert (Ross, J. Invest. Dermatol. 113 (1999), 991-998). Dagegen ist das homologe menschliche Gen in Makrophagen exprimiert (Godiska, J. Exp. Med. 185 (1997), 1595).

Ferner wurde von Tubb et al. bereits eine ca. 3,5 kb lange genomische 5'-flankierende Sequenz des Maus-Fascingens unter der Genbank/EMBL-Hinterlegungsnummer U90355 veröffentlicht. Zwar deutet einiges darauf hin, daß Fascin vorwiegend in dendritischen Zellen exprimiert wird, jedoch ist fraglich, ob diese Sequenz einen Promotor enthält, der in der Lage ist, einem Transgen eine spezifische Expression in dendritischen Zellen zu vermitteln. Bis auf die rohe Sequenz liegen jedenfalls keinerlei Informationen hierzu vor. Zusammenfassend läßt sich also feststellen, daß bisher kein Promotor vorliegt, der im Menschen eine für dendritische Zellen spezifische Expression vermittelt.

Daher besteht nach wie vor der Bedarf an solchen Promotoren oder regulatorischen Sequenzen, um gezielte Immunisierungen oder Immuntherapien durchführen zu können, die auf dendritische Zellen fokussiert sind.

Der vorliegenden Erfindung liegt somit die Aufgabe zugrunde, eine regulatorische Sequenz bereitzustellen, die in Mensch eine spezifische Expression in dendritischen Zellen vermittelt.

Diese Aufgabe wird erfindungsgemäß durch die Bereitstellung der in den Patentansprüchen charakterisierten Ausführungsformen gelöst.

Somit betrifft die vorliegende Erfindung regulatorische Sequenzen ausgewählt aus der Gruppe bestehend aus:

- (a) regulatorischen Sequenzen, die die unter SEQ ID NO:72 von Position 1 bis 3069 oder die unter SEQ ID NO:1 angegebene Nucleotidsequenz umfassen;

- (b) regulatorischen Sequenzen, die die Nucleotidsequenz umfassen, die in der Insertion des Klons DSM13274 enthalten ist und durch Amplifikation unter Verwendung eines Paares von Oligonucleotiden erhältlich ist, die beispielsweise die unter SEQ ID NO: 36 und 37 angegebenen Sequenzen besitzen;
- (c) regulatorischen Sequenzen, die eine Nucleotidsequenz umfassen aus SEQ ID NO:72 von Position 1136 bis 3069, 1451 bis 3069, 1621 bis 3069, 1830 bis 3069, 2127 bis 3069, 2410 bis 3069 oder 2700 bis 3069 oder ausgewählt aus der Gruppe bestehend aus: SEQ ID NOs:2 bis 8;
- (d) regulatorischen Sequenzen, die eine Nucleotidsequenz umfassen, die in der Insertion des Klons DSM13274 enthalten ist und durch Amplifikation unter Verwendung eines Paares von Oligonucleotiden erhältlich ist, wobei die Sequenzen der Oligonucleotide beispielsweise angegeben sind unter SEQ ID-Nummern, ausgewählt aus der Gruppe von Paaren bestehend aus: 38 und 37; 39 und 37; 40 und 37; 41 und 37; 42 und 37; 43 und 37; und 44 und 37;
- (e) regulatorischen Sequenzen, die zumindest einen funktionalen Teil einer unter (a) bis (d) genannten Sequenz umfassen und eine für dendritische Zellen spezifische Expression bewirken; und
- (f) regulatorischen Sequenzen, die eine Nucleotidsequenz umfassen, die mit einer unter (a) bis (e) angegebenen regulatorischen Sequenz hybridisiert, und eine für dendritische Zellen spezifische Expression bewirken.

Die regulatorischen Sequenzen der vorliegenden Erfindung verleihen Nucleotidsequenzen, die von ihnen kontrolliert werden, eine für dendritische Zellen spezifische Expression.

Der Begriff "regulatorische Sequenz" bezieht sich auf Nucleotidsequenzen, die das Expressionslevel eines Gens beeinflussen, beispielsweise indem sie der Expression Gewebe- oder Zellspezifität verleihen. In diesem Sinne werden beispielsweise unter regulatorischen Sequenzen Elemente, im folgenden auch als regulatorische Elemente bezeichnet, verstanden, die einem Minimalpromotor zusätzliche Expressionseigenschaften verleihen, die über die Minimalpromotoren charakterisierende basale, konstitutive Expression hinausgehen. Der Begriff "Minimalpromotor" bezieht sich im Zusammenhang der Erfindung auf

Nucleotidsequenzen, die für die Initiation der Transkription, d.h. für das Binden der RNA-Polymerase, notwendig sind, und beispielsweise die TATA-Box beinhalten. Ferner fallen unter den Begriff "regulatorische Sequenz" auch Sequenzen, die außerhalb des 5'-flankierenden Promotorbereichs liegen. Solche Sequenzen sind in beiden Orientierungen funktional und sind in der Position weniger festgelegt als Promotoren, vorzugsweise liegen sie innerhalb des Bereichs der im Rahmen der vorliegenden Erfindung offenbarten nicht-translatierten Sequenzen des humanen Fascingens (SEQ ID NO: 1 bis 20 oder die nicht-translatierten Sequenzen von SEQ ID NO:72). Zu solchen Sequenzelementen zählen Enhancer und Silencer, die die Expression eines Gens hoch- bzw. herunterregulieren können. Enhancer oder Silencer sind häufig in Introns oder im 3'-flankierenden Bereich eines Gens lokalisiert. Die regulatorische Sequenz kann auch ein Promotor sein, der im Sinne der Erfindung dadurch charakterisiert ist, daß er sämtliche Funktionen eines Promotors ausübt, d.h. Initiation der RNA-Polymerisation, Vermittlung einer bestimmten Expressionstärke und Regulation der Expression, vorzugsweise in Abhängigkeit des Zelltyps, besonders bevorzugt spezifisch für dendritische Zellen. Die in den SEQ ID NO: 1 bis 8 oder den oben genannten Abschnitten von SEQ ID NO:72 dargestellten Sequenzen sind regulatorische Sequenzen, die gleichzeitig der Definition eines Promotors entsprechen.

"Für dendritische Zellen spezifische Expression" bedeutet im Rahmen der vorliegenden Erfindung eine Expression ausschließlich in dendritischen Zellen, sofern Zellen des Hautgewebes, vorzugsweise der Epidermis, der mukosalen Gewebe, des Lymph- und Blutsystems und des Muskels betrachtet werden, vorzugsweise sofern solche Zellen mit einem Expressionskonstrukt transfiziert werden, das eine erfindungsgemäße regulatorische Sequenz enthält. Für Blutzellen aus Personen, die mit dem Epstein-Barr-Virus (EBV) infiziert sind, bezieht sich die für dendritische Zellen spezifische Expression zusätzlich auf B-Lymphocyten. Dabei können die erfindungsgemäßen regulatorischen Sequenzen sehr wohl auch in Zellen anderer Gewebetypen, wie z.B. Neuronen, Gliazellen oder Fibroblasten, eine Expression vermitteln, sofern sie in Haut, vorzugsweise in der Epidermis, in den mukosalen Geweben, im Blut und im Muskel spezifisch eine Expression nur in dendritischen Zellen gewährleisten. Vorzugsweise erfolgen Applikationen der

erfindungsgemäßen regulatorischen Sequenzen so, daß eine Expression nur in dendritischen Zellen gewährleistet ist. Dies ist beispielsweise für die Haut, vorzugsweise für die Epidermis, für die muskosalen Gewebe, das Blut oder den Muskel der Fall.

„Dendritische Zellen“ sind Antigen-präsentierende Zellen, die als einzige in der Lage sind, naive T-Zellen effizient zu aktivieren. Zu den dendritischen Zellen zählen auch die Langerhanszellen des Hautgewebes, welche unreife dendritische Zellen repräsentieren.

Die Spezifität der erfindungsgemäßen regulatorischen Sequenzen ist aus dem Expressionsverhalten des Fascingens, von dem die regulatorischen Sequenzen abgeleitet wurden, ersichtlich. So ist beispielsweise in der Haut, vorzugsweise der Epidermis, keine Fascinexpression detektiert worden, auch nicht in Langerhans-Zellen, da diese Zellen in der Haut unreif sind. Fascinexpression wurde allerdings in fortgeschritteneren Reifungsstadien der Langerhans-Zellen nachgewiesen (Ross, J. Immunol. 160 (1998), 3776-3782). In Blutzellen, zumindest von Personen, die nicht mit dem Epstein-Barr-Virus (EBV) infiziert sind, läßt sich Fascinexpression ausschließlich in dendritischen Zellen nachweisen (Mosialos, Am. J. Pathol. 148 (1996), 593-600). Zusätzlich wurde Fascinexpression in EBV-transfizierten B-Lymphocyten detektiert (Mosialos, J. Virol. 68 (1994), 7320). Im Muskel ist auf RNA-Ebene eine sehr geringe Fascin-Expression nachzuweisen, die auf die Präsenz von dendritischen Zellen zurückzuführen sein könnte (Mosialos, J. Virol. 68 (1994), 7320).

Die erfindungsgemäßen regulatorischen Sequenzen sind durch die unter SEQ ID NO: 1 bis 8 angegebenen Nucleotidsequenzen oder die oben genannten entsprechenden Abschnitte von SEQ ID NO:72 offenbart.

Die erfindungsgemäßen regulatorischen Sequenzen wurden durch DNA-Sequenzierung des humanen Fascin-Gens bereitgestellt, dessen Gesamtsequenz unter SEQ ID NO:72 sowie in Figur 9 dargestellt ist. Die Ergebnisse einer vorläufigen DNA-Sequenzierung dieses Gens erbrachten die unter SEQ ID NOs:10 bis 15 und 33 bis 35 sowie in Figur 2 angegebenen partiellen Nucleotidsequenzen. Die im Rahmen der Erfindung offenbarten regulatorischen Sequenzen SEQ ID

NOs:1 bis 20 sind diesen vorläufigen Sequenzen entnommen. Unterschiede zwischen der neueren Gesamtsequenz und den älteren Teilsequenzen sollten im allgemeinen auf Ungenauigkeit bei der vorläufigen DNA-Sequenzierung zurückzuführen sein.

Gleichwohl sind die regulatorischen Sequenzen offenbart durch den hinterlegten Klon DSM13274 und die Angabe von Oligonucleotiden, mit denen die jeweiligen Teilsequenzen der Insertion, beispielsweise durch PCR, amplifiziert werden können. Der Klon DSM13274 ist der PAC-Klon RPCIP704C24766Q3/4, der unter der Hinterlegungsnummer DSM13274 bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, am 1. Februar 2000 hinterlegt wurde. Mit Hilfe von Oligonucleotiden, beispielsweise mit solchen, deren Sequenzen unter SEQ ID NO: 36 und 37 angegeben sind, läßt sich der vollständige Promotor des humanen Fascingens amplifizieren. Mit Hilfe von Oligonucleotidpaaren, beispielsweise mit solchen, deren Sequenzen unter SEQ ID NO: 38 und 37; 39 und 37; 40 und 37; 41 und 37; 42 und 37; 43 und 37; oder 44 und 37 angegeben sind, lassen sich 5'-deletierte Fragmente des Fascinpromotors (s. Fig. 4) amplifizieren, die eine für dendritische Zellen spezifische Expression erlauben. Auf der Basis der unter SEQ ID NOs:1 bis 8 sowie 72 angegebenen Nucleotidsequenzen lassen sich weitere Oligonucleotidpaare ableiten, die es dem Fachmann erlauben, aus dem hinterlegten Klon die entsprechenden Promotorfragmente zu amplifizieren.

Bei der Bereitstellung der erfindungsgemäßen regulatorischen Sequenzen aus dem Fascinpromoter durch Amplifikation aus dem Klon DSM13274 kann die Spezifität der PCR-Reaktion erhöht werden, indem man ihr eine zusätzliche PCR-Reaktion voranstellt. Bei einer solchen "nested PCR" wird zunächst ein größerer Bereich des Fascingenlocus amplifiziert, der die gewünschten regulatorischen Sequenzen enthält, beispielsweise unter Verwendung von Oligonucleotiden, deren Sequenzen unter SEQ ID NO: 69 und 70 angegeben sind. Das Produkt der ersten PCR-Reaktion kann dann als Matrize für eine zweite PCR-Reaktion dienen, bei der mit Hilfe eines der oben angegebenen Oligonucleotidpaare eine der erfindungsgemäßen regulatorischen Sequenzen gewonnen werden kann.

Des weiteren kann die Sequenz der Promotorfragmente durch direkte Sequenzierung mit dem hinterlegten Klon als Matrize ermittelt werden. Der

Fachmann kann hierzu aus den unter SEQ ID NOs:1 bis 8 sowie 72 angegebenen Nucleotidsequenzen Sequenzierungsprimer ableiten.

Die Erfindung basiert auf der Beobachtung, daß ein etwa 3.0 kb umfassendes Promotorfragment des menschlichen Fascingens (pFascin-3.0, Fig. 4, SEQ ID NO: 1 bzw. Position 1 bis 3069 in SEQ ID NO:72), das vor die codierende Region eines Reportergens (Photinus-Luciferase) cloniert wurde, in transfizierten, kultivierten dendritischen Zellen zur Expression des Reportergens führte (siehe Beispiel 2). Desweiteren konnte gezeigt werden, daß diese Expression spezifisch ist für dendritische Zellen, da die Expression in THP-1-Zellen (menschliche Monocytenzelllinie, die endogen kein Fascin exprimiert, siehe Fig. 5) um etwa das 8-fache niedriger war. Weitere Expressionsstudien zeigten, daß der Fascinpromoter pFascin-3.0 sowie ein Subfragment davon (pFascin-1.6) in der Lage ist, Spezifität für die ausgereiften dendritischen Zellen (CD83⁺-Fraktion) zu vermitteln (Beispiel 2 und Figur 7). Die Stärke des Fascinpromoters wurde durch Cotransfektionsversuche mit einem Reportergenkonstrukt, das den unspezifischen, stark exprimierenden Promotor des Haushaltsgens EF1 α enthielt, bestimmt (Beispiel 2 und Figur 8). Bemerkenswerterweise übertraf der Fascinpromoter die Genexpression des EF1 α -Konstrukts in reifen dendritischen Zellen um etwa das 1½-fache.

Zur weiteren funktionellen Charakterisierung des Promotors wurden 5'-Deletionen des 3,0 kb-Fragments erzeugt und diese ebenfalls in Reportergen-Assays funktionell analysiert (siehe Beispiel 2, Fig. 4 und 6). Diese Analysen zeigten überraschenderweise, daß ein 211 bp-Promotorfragment (pFascin-0.11, SEQ ID NO: 21 bzw. Position 2859 bis 3069 in SEQ ID NO:72) in beiden untersuchten Zelltypen die gleiche Expressionsstärke aufwies. Ein um 56 bp weiter 5'-deletiertes Fragment (pFascin 0.05, SEQ ID NO: 22 bzw. Position 2915 bis 3069 in SEQ ID NO:72), das noch die TATA-Box enthält, zeigte dagegen keine Expression, die nennenswert über die Negativkontrolle hinausging. pFascin-0.11 ist demnach ein Minimalpromotor, der eine basale Grundtranskription vermittelt, die keine Spezifität für dendritische Zellen gegenüber Monocyten ausübt.

Überraschenderweise zeigte sich, daß offensichtlich in weiter distal gelegenen Regionen Elemente lokalisiert sind, die einerseits die Transkription in dendritischen Zellen erhöhen (maximal 3,3-fach gegenüber dem Minimalpromotor) und

andererseits Elemente vorhanden sind, die die Transkription in THP-1-Zellen vermindern (maximal um das 3-fache). Das Promotorkonstrukt pFascin-1.4 (SEQ ID NO. 4 bzw. Position 1621 bis 3069 in SEQ ID NO:72) zeigte die höchste Spezifität mit einer ca. 10-fach höheren Expression in dendritischen Zellen als in THP-1-Zellen. Diese Beobachtungen lassen den Schluß zu, daß die Promotorsequenz regulatorische Elemente enthält, die in der Lage sind, einem Minimalpromotor, wie z.B. dem in pFascin-0.11 enthaltenen Promotorfragment, Zellspezifität zu verleihen.

Somit betrifft die Erfindung auch funktionale Teile der Promotorsequenzen SEQ ID NO: 1 bis 8 bzw. Position 1 bis 3069 in SEQ ID NO:72, die, beispielsweise in Kombination mit einem Minimalpromotor, eine für dendritische Zellen spezifische Expression vermitteln. Beispiele für Minimalpromotoren sind der SV40- oder der Thymidinkinase-Minimalpromotor.

Darüberhinaus können die erfindungsgemäßen regulatorischen Sequenzen, die funktionale Teile des Fascinpromotors darstellen, dazu eingesetzt werden, das Expressionsverhalten bestehender heterologer Promotoren zu modifizieren. Beispielsweise kann ein Promotor, der eine von ihm kontrollierte Nucleotidsequenz konstitutiv exprimiert oder eine bestimmte Spezifität, wie z.B. Induzierbarkeit oder Entwicklungsspezifität aufweist, durch Integration eines oder mehrerer regulatorischer Sequenzen der Erfindung zusätzlich Spezifität für dendritische Zellen verliehen bekommen.

Es war bereits bekannt, daß das Protein Fascin, das Actin-Filamente kreuzvernetzt und in dendritischen Zellen auf diese Weise an der Ausbildung der Dendritenstruktur beteiligt ist, von dendritischen Zellen sowohl des Menschen (Ross, J. Invest. Dermatol. 115 (2000), 658-663) als auch der Maus (Ross, J. Immunol. 160 (1998), 3776-3782) stark exprimiert wird. Ferner war bekannt, daß Epidermalzellen in der unbehandelten Haut (zumindest bei Mäusen) keine Expression zeigen. Erst nach Aktivierung exprimieren die ausgereiften Langerhanszellen, die dendritischen Zellen der Epidermis, Fascin. Es wurde auch gezeigt, daß in humanen Blutzellen Fascin außer in dendritischen Zellen ebenfalls nicht exprimiert wird (Mosialos, Am. J. Pathol. 148 (1996), 593-600). Bisher war es jedoch nicht gelungen, die Sequenzen bereitzustellen, die die beschriebene Spezifität der Fascin-Expression bewirken.

Der Nachweis der Spezifität für dendritische Zellen dieser funktionalen Teile kann unter anderem durch das in Beispiel 2 beschriebene Verfahren erfolgen. Die Isolierung von Teilsequenzen aus einer der oben beschriebenen Promotorsequenzen kann nach den, dem Fachmann bekannten molekularbiologischen Standardmethoden erfolgen, beispielsweise nach Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2. Ausgabe, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY (1989)). Diese Quelle kann auch für alle weiteren in der vorliegenden Beschreibung aufgeführten molekularbiologischen Techniken zu Rate gezogen werden. Zur Testung der isolierten Fragmente auf Spezifität für dendritische Zellen kann beispielsweise das unter Beispiel 2 beschriebene Verfahren verwendet werden. Hierzu werden die Fragmente stromaufwärts eines Minimalpromotors, wie oben definiert, cloniert und anschließend in transienten Assays die Expression eines Reportergens in dendritischen Zellen und in nicht-Fascin-exprimierenden Zellen (z.B. THP-1) gemessen. Spezifische Expression in dendritischen Zellen im Sinne der Erfindung liegt vor, wenn der Expressionslevel im Vergleich zu den nicht-Fascin-exprimierenden Zellen mindestens um das 5-fache, vorzugsweise mindestens um das 8-fache, besonders bevorzugt mindestens um das 10-fache, insbesondere bevorzugt mindestens um das 15-fache und am meisten bevorzugt um mindestens das 20-fache erhöht ist. Ein weiterer Aspekt der Erfindung betrifft regulatorische Sequenzen, die mit einer der oben definierten erfindungsgemäßen regulatorischen Sequenzen, vorzugsweise dem komplementären Strang davon, hybridisieren, und eine für dendritische Zellen spezifische Expression einer von ihnen kontrollierten Nucleotidsequenz bewirken.

Bei diesen hybridisierenden Sequenzen kann es sich sowohl um Promotoren, wie weiter oben definiert, als auch um regulatorische Elemente handeln, die Minimalpromotoren Spezifität für dendritische Zellen verleihen können.

Der verwendete Begriff "hybridisieren" bezieht sich auf konventionelle Hybridisierungsbedingungen, vorzugsweise auf Hybridisierungsbedingungen, bei denen als Lösung 5xSSPE, 1% SDS, 1xDenhardts-Lösung verwendet wird und/oder die Hybridisierungstemperaturen zwischen 35°C und 70°C, vorzugsweise bei 65°C liegen. Nach der Hybridisierung wird vorzugsweise zuerst mit 2xSSC, 1% SDS und

danach mit 0,2XSSC bei Temperaturen zwischen 35°C und 70°C, vorzugsweise bei 65°C gewaschen (zur Definition von SSPE, SSC und Denhardts-Lösung siehe Sambrook et al., a.a.O.). Besonders bevorzugt sind stringente Hybridisierungsbedingungen, wie sie beispielsweise in Sambrook et al., s.o. beschrieben sind. Besonders bevorzugte stringente Hybridisierungsbedingungen liegen beispielsweise vor, wenn Hybridisierung und Waschen wie oben angegeben bei 65°C erfolgt. Weniger bevorzugt sind nicht-stringente Hybridisierungsbedingungen wie beispielsweise mit Hybridisierung und Waschen bei 45°C, noch weniger bevorzugt bei 35°C.

Vorzugsweise weisen derartige regulatorische Sequenzen eine Homologie, bestimmt durch Sequenzidentität, vorzugsweise über die gesamte Länge der verglichenen Sequenzen, von mindestens 50%, bevorzugt von mindestens 60%, besonders bevorzugt von mindestens 70%, vorteilhafterweise von mindestens 80%, vorzugsweise von mindestens 90% und insbesondere bevorzugt mindestens 95% zu der unter SEQ ID NO. 1 oder der von Position 1 bis 3069 in SEQ ID NO:72 gezeigten Sequenz auf. Vorzugsweise sind die hybridisierenden Sequenzen Fragmente mit einer Länge von mindestens 500 Nucleotiden, die zu der unter SEQ ID NO: 1 oder der von Position 1 bis 3069 in SEQ ID NO:72 gezeigten Sequenz eine Identität von mindestens 70%, vorzugsweise von mindestens 80%, besonders bevorzugt von mindestens 90% und insbesondere bevorzugt von mindestens 95% aufweisen. Wenn zwei zu vergleichende Sequenzen eine unterschiedliche Länge aufweisen, bezieht sich die Sequenzidentität vorzugsweise auf den prozentualen Anteil der Nucleotidreste der kürzeren Sequenz, die identisch sind mit den Nucleotidresten der längeren Sequenz. Die Sequenzidentität kann konventionell durch Verwendung von Computerprogrammen wie z. B. das Bestfit-Programm (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive Madison, WI 53711) bestimmt werden. Bestfit nützt den lokalen Homologiealgorithmus von Smith und Waterman, *Advances in Applied Mathematics* 2 (1981), 482-489, um das Segment mit höchster Sequenzidentität zwischen zwei Sequenzen zu finden. Bei der Anwendung von Bestfit oder einem anderen Sequenz-Alignment-Programm zur Bestimmung, ob eine bestimmte Sequenz beispielsweise zu 95% identisch ist mit einer Referenzsequenz der vorliegenden Erfindung, werden die Parameter vorzugsweise so eingestellt, daß

der Prozentanteil der Identität über die gesamte Länge der Referenzsequenz berechnet wird und daß Homologielücken ("gaps") von bis zu 5% der Gesamtzahl der Nucleotide in der Referenzsequenz erlaubt sind. Bei der Verwendung von Bestfit werden die sogenannten optionalen Parameter vorzugsweise bei ihren voreingestellten ("default") Werten belassen. Die Abweichungen, die bei dem Vergleich einer gegebenen Sequenz mit den oben beschriebenen Sequenzen der Erfindung auftreten, können beispielsweise durch Addition, Deletion, Substitution, Insertion oder Rekombination verursacht sein. Vorzugsweise kann ein derartiger Sequenzvergleich auch mit dem Programm DNASIS (Version 6.0, Hitachi Software Engineering Co. Ltd., 1984, 1990) durchgeführt werden. Hierbei sollten ebenfalls die "default"-Parametereinstellungen (Cut off Score: 16, Ktup: 6) verwendet werden. Ob hybridisierende Sequenzen eine für dendritische Zellen spezifische Expression vermitteln, kann beispielsweise ebenfalls nach den in Beispiel 2 beschriebenen Techniken festgestellt werden.

Die erfindungsgemäßen regulatorischen Sequenzen ermöglichen eine für dendritische Zellen spezifische Expression von Nucleotidsequenzen, die von Ihnen kontrolliert werden. Durch ihre Eigenschaft, als einzige Subpopulation der Antigen-präsentierenden Zellen (APC) naive T-Zellen effizient aktivieren zu können, nehmen dendritische Zellen eine zentrale Stellung bei der Vermittlung und Modulation von Immunantworten ein. Die Bereitstellung der erfindungsgemäßen regulatorischen Sequenzen eröffnet nun die Möglichkeit, dendritische Zellen gezielter als bisher zu nutzen und dabei Nachteile von im Stand der Technik beschriebenen Methoden zu überwinden. Dies bezieht sich im wesentlichen auf die Techniken der DNA-Vakzinierung und der in vitro-Transformation von dendritischen Zellen, unter anderem zum Zwecke der Gentherapie.

Daß sich dendritische Zellen als Ansatzpunkte für DNA-Vakzinierungen eignen, ist bekannt. So zeigten beispielsweise kinetische Studien, bei denen das immunisierte Gewebe zu unterschiedlichen Zeiten nach der Immunisierung entfernt wurde, daß direkt transfizierte dendritische Zellen, die kurze Zeit nach Immunisierung in die Lymphknoten wanderten, von ausschlaggebenderer Bedeutung für eine erfolgreiche Immunisierung sind (Torres et al., J. Immunol. 158 (1997), 4529-4532; Klinman et al., J. Immunol. 160 (1998), 2388-2392). Ferner konnte gezeigt werden, daß

dendritische Zellen, die nach Injektion von DNA aus dem transfizierten Gewebe auswanderten, das Antigen in immunogener Form trugen und in der Lage waren, Immunantworten gegen das Plasmid-codierte Protein hervorzurufen (Casares et al., J. Exp. Med. 186 (1997), 1481-1486).

DNA-Vakzinierungen werden primär durch Injektion, beispielsweise durch intramuskuläre oder intradermale Injektion vorgenommen, können aber z.B. auch durch Beschuß der Haut mit Partikeln, an die Expressionsplasmide gebunden sind, verabreicht werden. Ebenso können Expressionsplasmide oral oder sublingual gegeben werden oder auf die Schleimhäute des Respirationstraktes durch nasale oder intratracheale Applikation aufgebracht werden. Die Induktion von humoralen und zellulären Immunantworten durch Applikation von Antigen-codierender DNA in die Haut wird nach sehr gutem Erfolg im Tiermodell bereits am Menschen erprobt. Bisher werden Promotoren verwendet, die zur Expression des codierten Proteins in allen Hautzellen also auch beispielsweise in Keratinocyten führen. Dabei ist die epidermale Subpopulation der dendritischen Zellen, die Langerhans-Zellen, der eigentliche Adressat solcher Vakzinierungen. So zeigten Arbeiten an Knochenmarks-chimären Mäusen, daß die MHC-Restriktion der T-Zellantwort nach DNA-Immunisierung allein vom MHC-Haplotyp der Knochenmarkszellen, also den professionellen APCs (dendritischen Zellen) abhängig ist und nicht vom MHC-Haplotyp der Myocyten oder Keratinocyten (Fu et al., Mol. Med. 3 (1997), 362-371). Diese Zellen sind folglich selbst nicht in der Lage, eine Immunantwort zu induzieren. Dennoch könnte die Expression des Antigens in diesen Zellen zu einer durch dendritische Zellen vermittelten Immunisierung beitragen, wenn das sekretierte Antigen von dendritischen Zellen aufgenommen wird. Das Antigen könnte dann von der dendritischen Zelle prozessiert und über MHC-Moleküle präsentiert werden. Solche exogen aufgenommenen Antigene werden über MHC-Klasse II-Proteine präsentiert. Im Sinne einer effektiven DNA-Vakzinierungsstrategie ist jedoch eine Antigenpräsentation über MHC-Klasse I-Proteine wünschenswert, da diese die zelluläre Immunabwehr induziert, während MHC-Klasse II-Präsentation eine humorale Immunantwort hervorruft. Dies ist im Gegensatz zu herkömmlichen DNA-Vakzinierungsmethoden durch die regulatorischen Sequenzen der vorliegenden Erfindung nun besser realisierbar. Im Gegensatz zur Aufnahme exogener Antigene führt die intracelluläre Expression eines Antigens in dendritischen Zellen zur

Präsentation antigener Peptide über MHC-Klasse I-Moleküle. Durch die gezielte Adressierung von Antigen-exprimierenden Vektoren an dendritische Zellen wird die gestreute Expression von mittransfizierten umliegenden Zellen, wie z.B. Keratinocyten, und damit auch die exogene Aufnahme von exprimiertem Antigen via Endocytose durch dendritische Zellen weitgehend ausgeschlossen. Somit ist zu erwarten, daß durch Vakzinierung mit DNA, die spezifisch in dendritischen Zellen exprimiert wird, eine Stärkung der zellulären Immunantwort unter Beteiligung cytotoxischer T-Zellen erreicht wird.

Ferner ist bekannt, daß die Präsentation eines Antigens durch nicht-professionelle APCs, welche nicht die notwendigen co-stimulatorischen Signale für eine effiziente T-Zell-Stimulation zur Verfügung stellen, zu einer mangelhaften Reaktivität bzw. Anergie der T-Zellen führt. Mit Hilfe der erfindungsgemäßen regulatorischen Sequenzen ist es nun möglich, derartige Effekte weitgehend auszuschalten, indem die Expression des Antigens nach DNA-Vakzinierung auf reife dendritische Zellen beschränkt wird. Dies führt zu einer in der Summe verstärkten Immunantwort.

Eine weitere wichtige Anwendung der erfindungsgemäßen regulatorischen Sequenzen liegt in der DNA-Vakzinierung mit Konstrukten, die immunmodulatorische Proteine (z.B. Cytokine, co-stimulierende Membranproteine) exprimieren. Zusätzlich zur Expression von Antigenen kann durch Co-Transfektion von Konstrukten, die immunmodulatorische Proteine codieren, regulatorisch in die Immunantwort eingegriffen werden. Die Beschränkung dieser Mediatoren auf transfizierte dendritische Zellen erlaubt es, sehr spezifisch die erwünschte Immunantwort zu beeinflussen, ohne die bei systemischer Gabe der Mediatoren zu erwartenden Nebenwirkungen befürchten zu müssen.

Auch die Transfektion von dendritischen Zellen *in vitro*, unter anderem zu anschließenden therapeutischen Anwendungen, kann durch die regulatorischen Sequenzen der Erfindung entscheidend verbessert werden. Wie oben beschrieben, existieren im Stand der Technik bereits vielfältige Methoden zur Verwendung von *in vitro*-gereiften dendritischen Zellen zu therapeutischen Zwecken, die inzwischen bis hin zu klinischen Tests reichen. Für viele dieser Methoden und Anwendungen ist es jedoch notwendig, durch aufwendige Aufreinigungsprozeduren zu präparierende, möglichst homogene Populationen von dendritischen Zellen einzusetzen, um den bekannten Nachteilen der Transfizierung von Mischpopulationen zu entgehen.

Dendritische Zellen sind nur zu 1% oder weniger in der peripheren Blutleukocytenpopulation vertreten. Zu diesen Nachteilen zählt beispielsweise der oben im Zusammenhang mit DNA-Vakzinierung diskutierte Effekt der mangelnden Reaktivität oder Anergie von T-Zellen in Folge von Antigenpräsentation durch nicht-professionelle APCs.

Ein Beispiel für Therapieansätze mit in vitro-transfizierten dendritischen Zellen bezieht sich auf die Expression von Tumor- oder Pathogen-abgeleiteten Antigenen in dendritischen Zellen, die zur anschließenden Therapie in Patienten eingesetzt werden können.

Ein weiteres Beispiel für eine Therapieform durch gezielte Expression in dendritischen Zellen bezieht sich auf dendritische Zellen, die ein relevantes Antigen (z.B. Autoantigen oder Transplantationsantigen) präsentieren und mit dem IL-10 Gen oder einem anderen immunsupprimierenden Gen effizient transfiziert werden. Dies kann zur Induktion einer gezielten Anergie in den korrespondierenden T-Zellen eingesetzt werden. Auf diese Weise läßt sich beispielsweise Autoimmunität oder Transplantatabstoßung therapieren.

Ebenso ermöglichen die erfindungsgemäßen regulatorischen Sequenzen, Gene für regulatorisch eingreifende Proteine (z.B. Cytokine, co-stimulatorische Moleküle) spezifisch in dendritischen Zellen zur Expression zu bringen und somit die Qualität der Immunantwort durch Selektion der jeweiligen Klassen von T-Helferzellen bzw. cytotoxischen T-Zellen (Th1/Th2, Tc1/Tc2) und die Stärke der Immunantwort zu beeinflussen.

Der Vorteil der erfindungsgemäßen regulatorischen Sequenzen gegenüber z.B. dem konstitutiv exprimierenden CMV-Promotor liegt darin, daß die gesamte Leukocytenpopulation (bei EBV-negativen Personen), bzw. die B-Zell-depletierte Leukocytenpopulation (bei EBV-infizierten Personen) oder angereicherte frisch isolierte oder kultivierte dendritische Zellen in geeigneten Reifestadien eingesetzt werden können, und somit die zeitaufwendige und kostenintensive Aufreinigung von dendritischen Zellen umgangen werden kann.

In einer weiteren Ausführungsform der Erfindung sind die oben beschriebenen regulatorischen Sequenzen kombiniert mit mindestens einer Nucleotidsequenz, die

- (a) ausgewählt ist aus der Gruppe bestehend aus: den Abschnitten von SEQ ID NO: 72 von Position 3911 bis 13398, 13556 bis 13637, 13760 bis 14004, 14173 bis 15414 und 16791 bis 16951 und SEQ ID NOs: 9 bis 20, oder Teilen davon; oder
- (b) in der Insertion des Klons DSM13274 enthalten ist und durch Amplifikation unter Verwendung eines Paares von Oligonucleotiden erhältlich ist, wobei die Sequenzen der Oligonucleotide beispielsweise angegeben sind unter den SEQ ID-Nummern, ausgewählt aus der Gruppe von Paaren bestehend aus: 45 und 46; 47 und 48; 49 und 50; 51 und 52; 53 und 54; 55 und 56; 57 und 58; 59 und 60; 61 und 62; 63 und 64; 65 und 66; 67 und 68; und 45 und 60.

Nach den Ergebnissen aus Beispiel 2 ist die Reporterexpression in THP-1 durch die getesteten regulatorischen Sequenzen zwar deutlich vermindert gegenüber der basalen Expression, vor allem maximal ca. 10-fach niedriger als in dendritischen Zellen. In Anbetracht dessen, daß in THP-1 endogen keine Fascinexpression nachzuweisen ist (Fig. 5), ergibt sich die Frage, warum die Expression in THP-1 in transienten Expressionsassays nicht weit niedriger ausfällt. Offensichtlich reicht die Kontrolle des Fascinpromotors allein nicht aus für die völlige Repression der Fascinexpression. Demnach ist zu erwarten, daß ein weiteres kontrollierendes Element beispielsweise ein Silencer, im Fascingen die Transkription in nicht-dendritischen Zellen reprimiert. Bevorzugte Aufenthaltsorte für Silencer sind Introns oder die 3'-genflankierende Region. Somit umfaßt die Erfindung auch solche regulatorischen Sequenzen, die neben den oben beschriebenen Sequenzen zusätzlich mit einer oder mehreren der folgenden Sequenzen kombiniert ist: Intronsequenzen des Fascingens (die Abschnitte von SEQ ID NO: 72 von Position 3911 bis 13398, 13556 bis 13637, 13760 bis 14004 und 14173 bis 15414 bzw. SEQ ID NOs: 9 bis 19), die 161 bp umfassende 3'-genflankierende Region (Position 16791 bis 16951 in SEQ ID NO:72 bzw. SEQ ID NO: 20) und jeweils Teile davon. Vorzugsweise vermitteln die in dieser Ausführungsform gekennzeichneten regulatorischen Sequenzen Nucleotidsequenzen, die von ihnen kontrolliert werden, eine Expression, die noch spezifischer ist als die der regulatorischen Sequenzen, die nicht mit Intronsequenzen und/oder 3'-flankierenden Sequenzen kombiniert sind. "Noch spezifischer" bedeutet, daß die Expression durch die Kombination mit

regulatorischen Sequenzen der vorliegenden Ausführungsform zwischen dendritischen Zellen und nicht-Fascin-exprimierenden Zellen um einen Faktor divergiert, der größer ist als mit denselben regulatorischen Sequenzen ohne eine Sequenz aus den Intron- und 3'-flankierenden Sequenzen in SEQ ID NO:72 bzw. SEQ ID NOs:9 bis 20. Dieser Faktor ist vorzugsweise größer als 10, besonders bevorzugt größer als 15, insbesondere bevorzugt größer als 20, am meisten bevorzugt größer als 30. Vorzugsweise liegt die Expression einer von diesen regulatorischen Sequenzen kontrollierten Nucleotidsequenzen in nicht-Fascin-exprimierenden Zellen unter der Nachweisgrenze.

"Kombiniert" bedeutet in diesem Zusammenhang, daß eine oder mehrere Sequenzen aus den Intron- und 3'-flankierenden Sequenzen in SEQ ID NO:72 bzw. SEQ ID NOs:9 bis 20 oder Teile davon nach den gängigen molekularbiologischen Techniken in die Nähe einer unter SEQ ID NOs: 1 bis 8 gezeigten Sequenz bzw. der oben genannten entsprechenden Abschnitte von SEQ ID NO:72, einem funktionalen Teil davon oder einer Sequenz, die eine der vorgenannten Sequenzen hybridisiert, cloniert wird. "In die Nähe" heißt, daß die Sequenz(en) aus den Intron- und 3'-flankierenden Sequenzen in SEQ ID NO:72 bzw. SEQ ID NOs: 9 bis 20 oder Teilen davon direkt oder in einem gewissen Abstand zu den vorgenannten regulatorischen Sequenzen stromabwärts oder stromaufwärts oder intermittierend cloniert wird, vorzugsweise erfolgt die Clonierung jedoch stromabwärts, da die vorgenannten regulatorischen Sequenzen (Promotoren oder funktionale Teile davon) bekanntermaßen für ihre Funktionsweise, d.h. zum Binden von RNA-Polymerase oder Transkriptionsfaktoren, relativ definierte Distanzen zum Transkriptionsstartpunkt bzw. zur TATA-Box benötigen.

Unter "in einem gewissen Abstand" wird eine Distanz verstanden, die dazu geeignet ist, daß Silencer oder Enhancer ihre Funktion ausüben können. Die Formulierung "Teil davon" ist zu verstehen als eine Sequenz innerhalb einer in SEQ ID NOs: 9 bis 20 gezeigten Sequenz bzw. einer Intron- oder 3'-flankierenden Sequenz in SEQ ID NO: 72, die geeignet ist, eine Silencer- oder Enhancer-Funktion auszuüben. Derartige Sequenzen können durch eine experimentelle Vorgehensweise, die zu der in Beispiel 2 analog ist, d.h. durch funktionelle Analysen von Reporter-Gen-Konstrukten, eingegrenzt werden. Hierzu kann man beispielsweise eine isolierte Teilsequenz einer Sequenz aus SEQ ID NOs:9 bis 20 bzw. einer Intron- oder 3'-

flankierenden Sequenz in SEQ ID NO: 72 in ein funktionales Promotor-Reportergen-Konstrukt, wie beispielsweise pFascin-3.0 klonieren, vorzugsweise direkt an das 5'-Ende des Promotorfragments angrenzend, oder in ein Intron des Reportergens, und anschließend dieses Konstrukt und zum Vergleich dasselbe Konstrukt ohne diese Teilsequenz in einem transienten Expressionsassay in dendritischen Zellen, vorzugsweise reifen dendritischen Zellen, sowie in nicht-Fascin-exprimierenden Zellen analysieren. Ergibt sich hierbei für das Konstrukt mit dem Teilstück eine größere Differenz der Expressionslevel zwischen dendritischen Zellen und nicht-Fascin-exprimierenden Zellen als für das Konstrukt ohne das Teilstück, so umfaßt dieses Teilstück ein funktionales Silencer- oder Enhancer-Element. Weitere Techniken zur Eingrenzung solcher Elemente stehen dem Fachmann zur Verfügung. Die Ausführungsform umfaßt die oben beschriebenen regulatorischen Sequenzen, die kombiniert sind mit mindestens einer Nucleotidsequenz, die durch Amplifikation, beispielsweise mittels PCR, aus der Insertion des hinterlegten Klon DSM13274 bereitgestellt werden kann, oder Teilen davon. Zur Amplifikation können beispielsweise Paare von Oligonucleotiden verwendet werden, deren Sequenzen unter folgenden SEQ ID-Nummern angegeben sind: 45 und 46; 47 und 48; 49 und 50; 51 und 52; 53 und 54; 55 und 56; 57 und 58; 59 und 60; 61 und 62; 63 und 64; 65 und 66 (Intronsequenzen); sowie 67 und 68 (3'-genflankierende Region). Weitere Oligonucleotidpaare lassen sich zu diesem Zweck von den unter SEQ ID NO: 9 bis 20 angegebenen Nucleotidsequenzen ableiten.

Des weiteren kann die Sequenz dieser Fragmente durch direkte Sequenzierung mit dem hinterlegten Klon als Matrize ermittelt werden. Der Fachmann kann hierzu aus den unter SEQ ID NO: 9 bis 20 angegebenen Nucleotidsequenzen bzw. den in SEQ ID NO: 72 enthaltenen Intron- oder 3'-flankierenden Sequenzen Sequenzierungsprimer ableiten.

Ebenso betrifft die vorliegende Ausführungsform die Kombination der oben definierten regulatorischen Sequenzen mit dem gesamten Intron 1 des Fascingens oder Teilen davon. Die Nucleotidsequenz des Introns 1 ist in dem hinterlegten Klon DSM13274 enthalten und kann durch Amplifikation bereitgestellt werden, wobei hierzu ein Oligonucleotidpaar mit Sequenzen, die beispielsweise unter SEQ ID NO: 45 und 60 angegeben sind, verwendet werden kann.

Auch hier kann die Spezifität von PCR-Reaktionen, die der Bereitstellung von Sequenzen aus dem Klon DSM13274 dienen, die im Rahmen der vorliegenden Ausführungsform mit einer der oben beschriebenen regulatorischen Sequenzen kombiniert werden können, dadurch erhöht werden, daß man ihr eine zusätzliche PCR-Reaktion voranstellt. Bei einer solchen "nested PCR" wird zunächst ein größerer Bereich des Fascingenlocus amplifiziert, der die gewünschten Sequenzen enthält, beispielsweise unter Verwendung von Oligonucleotiden, deren Sequenzen unter SEQ ID NO: 71 und 68 angegeben sind. Das Produkt der ersten PCR-Reaktion kann dann als Matrize für eine zweite PCR-Reaktion dienen, bei der mit Hilfe eines der oben angegebenen Oligonucleotidpaare eine Intron- oder 3'-genflankierte Sequenz zur Kombination mit den oben beschriebenen regulatorischen Sequenzen gewonnen werden kann.

Die Kombination der durch Amplifikation aus dem Klon DSM13274 gewonnenen Sequenzen mit den oben definierten regulatorischen (Promoter-)sequenzen kann entsprechend der für die im Sequenzprotokoll offenbarten Intron- und 3'-flankierenden Sequenzen angegebenen Anleitung erfolgen.

In einer besonders bevorzugten Ausführungsform stammen die regulatorischen Sequenzen aus Mensch.

Vorzugsweise handelt sich bei den erfindungsgemäßen regulatorischen Sequenzen um DNA- oder RNA-Moleküle, wobei es sich bei den DNA-Molekülen vorzugsweise um genomische DNA handelt.

Eine weitere bevorzugte Ausführungsform der Erfindung bezieht sich auf rekombinante Nucleinsäuremoleküle, die die erfindungsgemäßen regulatorischen Sequenzen enthalten.

Der Begriff „rekombinantes Nucleinsäuremolekül“ bezieht sich auf Nucleinsäuremoleküle, die einem unterschiedlichen genetischen Kontext entstammen und durch molekularbiologische Methoden kombiniert wurden. „Unterschiedlicher genetischer Kontext“ bezieht sich hierbei auf Genome aus unterschiedlichen Spezies, Varietäten oder Individuen oder unterschiedlichen Positionen innerhalb eines Genoms. Neben natürlichen Sequenzen können

rekombinante Nucleinsäuremoleküle auch Sequenzen enthalten, die im Vergleich zu ihrer natürlichen Vorlage mutiert oder chemisch modifiziert sind, oder es handelt sich um gänzlich neu synthetisierte Sequenzen.

An den erfindungsgemäßen rekombinanten Nucleinsäuremolekülen liegen eine oder mehrere der oben beschriebenen regulatorischen Sequenzen in Kombination mit Sequenzen aus einem anderen genetischen Kontext vor. Ein Beispiel für ein rekombinantes Nucleinsäuremolekül enthält eine oder mehrere erfindungsgemäße regulatorische Sequenzen in Kombination mit einem Minimalpromotor, der einem anderen Gen als dem humanen Fascingen entnommen wurde. Die regulatorischen Sequenzen sind hierbei regulatorische Promotorelemente, die dem Gesamtpromotor eine für dendritische Zellen spezifische Expression verleihen.

Desweiteren können die rekombinanten Nucleinsäuremoleküle neben einem Promotor, der eine oder mehrere erfindungsgemäße regulatorische Sequenzen enthält, stromabwärts davon eine Polylinkersequenz mit einer oder mehreren Restriktionserkennungsstellen enthalten, in die durch dem Fachmann bekannte Verfahren Nucleotidsequenzen cloniert werden können, die dadurch unter die Expressionskontrolle des Promotors gelangen. Vorzugsweise liegt dieser Polylinker in einer Region, die direkt hinter dem durch den Promotor definierten Transkriptionsstartpunkt liegt.

Ferner kann das erfindungsgemäße rekombinante Nucleinsäuremolekül stromabwärts des Polylinkers ein Transkriptions-Terminationssignal enthalten. Beispiele für geeignete Terminationssignale sind im Stand der Technik beschrieben. Dabei kann es sich beispielsweise um das Thymidinkinase-Polyadenylierungssignal handeln. Die hier beschriebenen rekombinanten Nucleinsäuremoleküle, die vorzugsweise eine zu exprimierende Nucleotidsequenz enthalten, können direkt für Anwendungen im Sinne der Erfindung, wie beispielsweise DNA-Vakzinierungen eingesetzt werden. Die erfindungsgemäßen rekombinanten Nucleinsäuremoleküle können beispielsweise durch gängige in vitro-Amplifikationstechniken, wie zum Beispiel PCR, vermehrt werden. Sie können allerdings auch konventionell in einem Vektor in vivo vermehrt werden und nach Nucleinsäurepräparation und anschließender Heraustrennung aus dem Vektor, wie beispielsweise durch Restriktionsspaltung, für Anwendungen, in denen beispielsweise linearisierte Expressionseinheiten benötigt werden, zur Verfügung gestellt werden.

Rekombinante Nucleinsäuremoleküle, die vorzugsweise eine zu exprimierende Nucleotidsequenz enthalten, können auch Expressionseinheiten darstellen, die oft auch als Expressionskassetten bezeichnet werden, die leicht in verschiedene Standardvektoren cloniert werden können und so je nach Vektor verschiedene Funktionen ausüben können.

Für Anwendungen der erfindungsgemäßen regulatorischen Sequenzen bei EBV-infizierten Personen kann die Expression in B-Zellen durch die Kombination der regulatorischen Sequenzen mit bekannten Silencer-Elementen für B-Zellen unterdrückt werden. Andererseits kann es auch durchaus von Vorteil sein, wenn EBV-infizierte B-Zellen ebenfalls die von den regulatorischen Sequenzen kontrollierte Nucleotidsequenz exprimieren, da EBV-transfizierte B-Zellen gute Antigen-präsentierende Zellen für aktivierte T-Zellen darstellen und somit eine Amplifikation der Immunantwort bewirken können.

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft Vektoren, die die erfindungsgemäße regulatorische Sequenz oder das erfindungsgemäße rekombinante Nucleinsäuremolekül enthalten.

Der Begriff „Vektor“ betrifft zirkuläre oder lineare Nucleinsäuremoleküle, die sich autonom in Wirtszellen, in die sie eingeführt werden, replizieren können. Die Vektoren können die oben gekennzeichneten rekombinanten Nucleinsäuremoleküle in ihrer vollen Länge enthalten oder können auch neben den erfindungsgemäßen regulatorischen Sequenzen, die für die rekombinanten Nucleinsäuremoleküle beschriebenen Bestandteile, wie beispielsweise Minimalpromotor, Polylinker und/oder Terminationssignal, enthalten.

Die erfindungsgemäßen Vektoren können zur Replikation in prokaryontischen und/oder eukaryontischen Wirtszellen geeignet sein und enthalten einen entsprechenden Replikationsursprung (origin of replication). Vorzugsweise eignen sich die Vektoren zur Replikation in Säugerzellen, besonders bevorzugt in menschlichen Zellen.

Vorzugsweise enthalten die erfindungsgemäßen Vektoren einen Selektionsmarker. Beispiele für Selektionsmarkergene sind dem Fachmann bekannt. Selektionsmarkergene, die sich zur Selektion in eukaryontischen Wirtszellen eignen,

sind beispielsweise Gene für die Dihydrofolat-Reductase, G418 oder Neomycin-Resistenz.

Vorzugsweise handelt es sich bei den erfindungsgemäßen Vektoren um Expressionsvektoren zur Expression in eukaryontischen Zellen. Solche Vektoren lassen sich auch ausgehend von bekannten Expressionsvektoren konstruieren, indem dessen Promotor bzw. die nicht zu einem Minimalpromotor gehörenden Sequenzen davon durch die regulatorischen Sequenzen der Erfindung ersetzt oder um regulatorische Sequenzen (regulatorische Elemente) ergänzt werden. Beispiele für Expressionsvektoren, die sich auf diese Weise modifizieren lassen, sind pcDV1 (Pharmacia), pRC/CMV, pcDNA1 oder pcDNA3 (Invitrogen).

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft die oben beschriebenen rekombinanten Nucleinsäuremoleküle oder Vektoren, die zusätzlich eine zu exprimierende Nucleotidsequenz enthalten, wobei die Expression der Nucleotidsequenz von der regulatorischen Sequenz oder von einem Promoter, der die regulatorische Sequenz enthält, kontrolliert wird.

Die „zu exprimierenden Nucleotidsequenzen“ codieren entweder ein Protein oder (Poly)peptid oder RNA-Moleküle, die ihre Funktion auf der RNA-Ebene entfalten. Nucleotidsequenzen, die ein Protein, Polypeptid oder Peptid codieren, umfassen eine codierende Region, die gekennzeichnet ist durch ein Startcodon (ATG), eine Folge von Aminosäure-codierenden Basentriplets und ein Stopcodon (TGA, TAG oder TAA), wenn es sich um DNA handelt. Bei RNAs ist das Thymidin (T) durch Uracil (U) ersetzt. Bei degenerierten Aminosäurecodons können die Basentriplets entsprechend der Codon-Usage der Zielzellen angepaßt werden, mit Techniken, die dem Stand der Technik entnommen werden können. Beispiele für Nucleotidsequenzen, die RNA-Moleküle exprimieren, sind Antisense-RNA oder Ribozyme.

Weiterhin betrifft die Erfindung die oben genannten rekombinanten Nucleinsäuremoleküle oder Vektoren, wobei die zu exprimierende Nucleotidsequenz ein Antigen codiert.

Der Begriff „Antigen“ betrifft im Zusammenhang mit der vorliegenden Erfindung Moleküle, die von einem Organismus als fremd erkannt werden, und dadurch eine

spezifische Immunantwort auslösen. Antigene werden natürlicherweise von Antigen-präsentierenden Zellen (APC) endocytisiert und zusammen mit Histokompatibilitätsantigenen (MHC) der Klasse II an der Zelloberfläche präsentiert. Bei den Antigenen, die von der vorliegenden Ausführungsform betroffen sind, handelt es sich um Proteine, Polypeptide oder Peptide. Antigene, die von einem Expressionsvektor in einer dendritischen Zelle exprimiert werden, werden durch MHC-I-Proteine präsentiert. Dabei werden die synthetisierten Proteine durch Proteasomen in Peptide gespalten. Die Peptide werden bevorzugt durch TAP1/TAP2-Komplexe in das endoplasmatische Retikulum transloziert und binden dort an MHC-Klasse-I-Moleküle. Ein Teil der Peptide bindet aber auch auf unbekanntem Weg an MHC-Klasse-II-Moleküle.

In einer bevorzugten Ausführungsform der rekombinanten Nucleinsäuremoleküle oder Vektoren sind die Antigene tumor- oder pathogenspezifisch.

Vorzugsweise sind die Antigene, die von den rekombinanten Nucleinsäuremolekülen oder Vektoren spezifisch in dendritischen Zellen exprimiert werden, Proteine von Pathogenen, wie zum Beispiel Viren (z. B. HIV), Bakterien, Pilzen oder Parasiten, die in Patienten eine Immunantwort auslösen. Unter pathogenspezifischen Antigenen sind allerdings auch Proteine oder (Poly)Peptide eingeschlossen, die zumindest eine antigene Determinante (Epitop) eines Pathogens enthalten.

Tumorspezifische Antigene sind Proteine oder (Poly)Peptide von Tumorzellen, die eine spezifische Immunantwort auslösen können. Auch hierbei sind (Poly)Peptide eingeschlossen, die mindestens ein Epitop eines tumorspezifischen Antigens enthalten.

Zusätzlich kann es sich bei dem Antigen um Proteine der Alzheimer-Plaques (oder zumindest Epitope daraus) handeln, die an der Alzheimer-Krankheit ursächlich beteiligt sind.

In einer weiteren Ausführungsform der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren ist das Antigen ein Autoantigen oder ein Transplantationsantigen.

Der Begriff „Autoantigen“ betrifft körpereigene Antigene eines Patienten, die beispielsweise in Folge einer Störung der Selbsterkennung oder von

Regulationsmechanismen des Immunsystems unter Bildung von Autoantikörpern eine Autoimmunerkrankung verursachen. Geeignete Autoantigene können beispielsweise bei einer vorliegenden Autoimmunerkrankung durch Identifizierung von Autoantikörpern im Patienten bestimmt werden.

Bei „Transplantationsantigenen“ handelt es sich um Histokompatibilitätsantigene (MHC) der Klasse I und II, die durch allogene Transplantate in einen Organismus eingeführt werden und dort eine Immunreaktion (Transplantatabstoßung) auslösen.

Die erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren, die spezifisch in Zellen Autoantigene oder Transplantationsantigene exprimieren können, können dazu verwendet werden, die gegen diese Antigene gerichtete Immunreaktion zu hemmen. Hierzu können beispielsweise dendritische Zellen, die diese Antigene exprimieren, in vitro mit einem weiteren Expressionsvektor transfiziert werden, der ein immunregulatorisches Molekül (z. B. IL-10), exprimiert, das in der Lage ist, Immunreaktionen zu hemmen. Durch Verabreichung derart transformierter dendritischer Zellen kann gezielt eine Anergie von T-Zellen, die Autoantigen- bzw. Transplantationsantigen-spezifisch sind, erzeugt werden, und so die pathologische Immunantwort behandelt werden.

In einer weiteren bevorzugten Ausführungsform der rekombinanten Nucleinsäuremoleküle oder Vektoren sind die Antigene Allergene.

Im Zusammenhang mit der vorliegenden Ausführungsform handelt es sich bei "Allergenen" um Proteine oder (Poly)Peptide, die in Organismen eine allergische Reaktion auslösen können, oder zumindest um ein Epitop aus einem solchen Protein oder (Poly)Peptid.

Dendritische Zellen, die intrazellulär ein Allergen exprimieren, können beispielsweise eine Allergen-spezifische cytotoxische T-Zellantwort induzieren, woraufhin cytotoxische T-Zellen Allergen-präsentierende Zellen töten können. Auf diese Weise ist es möglich, die Aktivierung von Allergen-spezifischen T-Helferzellen und somit die Aktivierung von IgE-Antikörper-produzierenden B-Zellen zu verhindern.

Desweiteren können dendritische Zellen mit Expressionsvektoren cotransfiziert werden, die ein Allergen und ein immunregulatorisches Molekül, das in der Lage ist, eine Immunantwort zu hemmen (z.B. IL-10), exprimieren. Alternativ können dendritische Zellen mit Expressionsvektoren cotransfiziert werden, die ein Allergen

und eine Antisense-RNA exprimieren, wobei letztere die Expression eines co-stimulatorischen Moleküls (wie z.B. der B7-Familie) inhibiert. In beiden Fällen kann dadurch eine Anergie in den Allergen-spezifischen Zellen induziert werden.

In einer weiteren bevorzugten Ausführungsform der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren codieren die zu exprimierenden Nucleotidsequenzen ein Protein, das eine Immunantwort reguliert.

Durch die spezifische Adressierbarkeit von dendritischen Zellen durch die erfindungsgemäßen regulatorischen Sequenzen, besteht zusätzlich zur Antigen-Präsentation die Möglichkeit, eine Immunantwort zielgerichtet zu steuern, indem man Proteine in dendritischen Zellen exprimiert, die eine Immunantwort regulieren. Verabreichungen von rekombinanten Nucleinsäuren oder Vektoren, die solche Proteine exprimieren, können erfolgen, indem in vitro dendritische Zellen transfiziert werden, und die Zellen anschließend in eine Person eingebracht werden oder indem Expressionskonstrukte für diese Proteine direkt verabreicht werden, beispielsweise durch Injektion. Immunregulatorische Proteine sollten vorzugsweise in dendritischen Zellen exprimiert werden, die mit einem Antigen beladen sind, vorzugsweise indem die Zellen das Antigen selbst exprimieren, beispielsweise durch einen erfindungsgemäßen Vektor, damit zielgerichtet die Immunantwort reguliert wird, die durch das Antigen hervorgerufen wird.

Beispiele für immunregulatorische Proteine sind Cytokine, zu denen die Lymphokine, Monokine, Interleukine, Chemokine, koloniestimulierenden Faktoren, Interferone und transformierende Wachstumsfaktoren, wie zum Beispiel TGF- β , gerechnet werden. Weitere immunregulatorische Proteine sind co-stimulierende Moleküle. Nucleotidsequenzen, die die immunregulatorischen Proteine codieren, sind im Stand der Technik beschrieben und können der Literatur entnommen werden.

Besonders bevorzugt ist dabei eine Ausführungsform der rekombinanten Nucleinsäuremoleküle oder Vektoren, wobei das Protein, das eine Immunantwort reguliert, ein Cytokin oder ein co-stimulierendes Molekül ist.

„Cytokine“ sind definiert als von verschiedenen Zellarten gebildete und sezernierte Substanzen, die als interzelluläre Mediatoren zur Steuerung der Aktivität von

anderen Zellen beitragen. Für die Anwendungen, die sich durch die erfindungsgemäßen regulatorischen Sequenzen eröffnen, sind vor allem von Bedeutung die Interleukine, Interferone, Chemokine, koloniestimulierenden Faktoren und transformierenden Wachstumsfaktoren.

Unter „co-stimulierenden Faktoren“ sind Moleküle zu verstehen, die von professionellen Antigen-präsentierenden Zellen wie den dendritischen Zellen membrangebunden exprimiert werden oder von diesen Zellen sekretiert werden und die für die effiziente Aktivierung von T-Lymphocyten benötigt werden.

Nucleotidsequenzen, die Cytokine oder co-stimulierende Moleküle codieren, sind dem Fachmann bekannt. Beispielsweise sind Aminosäuresequenzen von Cytokinen in "The Cytokine Handbook" (A.W. Thomson, ed. Academic Press, San Diego, CA, 1998) und von co-stimulatorischen Molekülen in "The Leucocyte Antigen Facts Book" (A.N. Barclay, M.H. Brown, S.K.A. Law, A.J. McKnight, M.G. Tomlinson, P.A. van der Merwe, eds., Academic Press, San Diego, CA, 1997) veröffentlicht. Entsprechende codierende Nucleotidsequenzen können den öffentlich zugänglichen Datenbanken entnommen werden. Vorzugsweise werden zur Konstruktion der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle und Vektoren dieser und der nachfolgenden Ausführungsformen Nucleotidsequenzen verwendet, die humanen Ursprungs sind. Durch die Wahl von geeigneten Cytokinen oder co-stimulierenden Molekülen zur spezifischen Expression in dendritischen Zellen, kann die Immunantwort entweder gesteigert oder gehemmt werden.

In einer besonders bevorzugten Ausführungsform exprimieren die erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren Proteine, die eine Immunantwort regulieren, wobei die Regulation eine Hemmung ist.

Diese Ausführungsform der Erfindung kann dazu verwendet werden, dendritische Zellen, die ein Antigen tragen, das eine unerwünschte Immunantwort hervorruft (z.B. Autoantigene oder Transplantationsantigene), mit rekombinanten Nucleinsäuremolekülen oder Vektoren zu transfizieren, die ein Protein exprimieren, das die Immunantwort hemmt. Dies kann zur Induktion einer gezielten Anergie in den korrespondierenden T-Zellen eingesetzt werden.

Vorzugsweise tragen die dendritischen Zellen das Antigen, weil sie mit einem erfindungsgemäßen Nucleinsäuremolekül oder Vektor, das/der das Antigen

exprimiert, transfiziert wurden. Cytokine mit einer derart hemmenden Aktivität sind in der Literatur beschrieben. Hierzu zählt beispielsweise das Interleukin IL-10 oder der transformierende Wachstumsfaktor TGF- β .

Somit betrifft eine besonders bevorzugte Ausführungsform rekombinante Nucleinsäuremoleküle oder Vektoren, die das Protein IL-10 oder TGF- β exprimieren.

In einer weiteren bevorzugten Ausführungsform exprimieren die erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren Proteine, die eine Immunantwort regulieren, wobei die Regulation eine Steigerung der Immunantwort ist.

Eine gezielte Steigerung der Immunantwort kann durch Verabreichung von rekombinanten Nucleinsäuremolekülen oder Vektoren erreicht werden, die geeignete Cytokine oder co-stimulierende Moleküle exprimieren. Unter Cytokinen sind beispielsweise für die Interleukine IL-2, IL-4, IL-12, IL-15, IL-18, für die Interferone IFN-gamma und IFN-alpha, für die Chemokine DC-CK1 und MDC sowie für den Granulocyten/Monocyten koloniestimulierenden Faktor (GM-CSF) immunstimulierende Wirkungen beschrieben worden. Nucleinsäuremoleküle, die diese Faktoren codieren, sind dem Stand der Technik zu entnehmen, wie beispielsweise in den oben genannten Quellen.

Somit betrifft eine besonders bevorzugte Ausführungsform rekombinante Nucleinsäuremoleküle oder Vektoren, die die Proteine IL-2, IL-4, IL-12, IL-15, IL-18, IFN-gamma, IFN-alpha, DC-CK1, MDC oder GM-CSF exprimieren.

Eine weitere besonders bevorzugte Ausführungsform betrifft rekombinante Nucleinsäuremoleküle oder Vektoren, die ein co-stimulatorisches Molekül, vorzugsweise ein Mitglied der B7-Familie, ICOS-Ligand oder CD40, exprimieren.

Proteine der B7-Familie, ICOS-Ligand und CD40 sind co-stimulatorische Moleküle, die sich zu einer Steigerung der Immunantwort in dem oben ausgeführten Sinne eignen. Nucleotidsequenzen, die diese co-stimulatorischen Faktoren codieren sind ebenfalls in der Literatur beschrieben (siehe oben).

In einer weiteren besonders bevorzugten Ausführungsform der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren codiert die zu exprimierende Nucleotidsequenz ein Apoptose-induzierendes Molekül.

Der Begriff „Apoptose“ bezeichnet programmierten Zelltod, der durch exogene Signale induziert werden kann. Die rekombinanten Nucleinsäuremoleküle oder Vektoren der vorliegenden Ausführungsform können dazu verwendet werden, in Antigen-beladene dendritische Zellen transfiziert zu werden, um T-Zellen, die gegen das Antigen spezifisch sind, zum Absterben zu bringen. Auf diese Weise kann die Zahl solcher T-Zellen herabgesetzt werden und eine unerwünschte Immunreaktion, beispielsweise gegen Autoantigene oder Transplantationsantigene, abgeschwächt werden.

Im Stand der Technik sind einige Proteine beschrieben, die membrangebunden sind, aber auch teilweise sezerniert werden können und in nächster Nachbarschaft in Zellen das Selbstmordprogramm auslösen können. Darunter fallen beispielsweise die Proteine der TNF-Superfamilie.

Somit betrifft eine besonders bevorzugte Ausführungsform der Erfindung rekombinante Nucleinsäuremoleküle oder Vektoren, die ein Apoptose-induzierendes Molekül exprimieren, wobei das Apoptose-induzierende Molekül zur TNF-Superfamilie gehört.

In einer weiteren besonders bevorzugten Ausführungsform der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren ist die zu exprimierende Nucleotidsequenz eine Antisense-Sequenz oder exprimiert ein Ribozym.

Die Antisense-Sequenzen und Ribozyme sind Moleküle, deren Expression auf der Ebene der RNA stattfindet. Bei „Antisense-Sequenzen“ handelt es sich um Sequenzen, die komplementär sind zu einer in der Zielzelle vorhandenen mRNA oder einem Teil davon, wobei der Teil die codierende Region, 5'- und/oder 3'-nicht-translatierte Region umfassen kann. Antisense-RNAs, d.h. die Transkripte der Antisense-Sequenz, sind in der Lage, an die komplementäre mRNA in vivo zu hybridisieren und auf diese Weise deren Translation zu inhibieren.

„Ribozyme“ sind katalytische RNA-Moleküle. Vorzugsweise handelt es sich im Zusammenhang mit der vorliegenden Erfindung um Ribozyme, die spezifisch an

eine mRNA binden können, um diese durch eine katalytische Aktivität für eine erfolgreiche Translation unzugänglich zu machen, vorzugsweise durch hydrolytische Spaltung. Anleitungen zur Auswahl von geeigneten Antisense-Sequenzen sowie zur Konstruktion von Ribozymen mit gewünschter Sequenzspezifität sind in der Literatur beschrieben und können beispielsweise Antisense: "From Technology to Therapy" (Schlingensiepen, R., Brysch, W., Schlingensiepen, K.-H., eds., Blackwell Science Ltd. Oxford, 1997) bzw. Rossi (AIDS Research and Human Retroviruses 8 (1992), 183) entnommen werden.

Eine besonders bevorzugte Ausführungsform betrifft die vorgenannten rekombinanten Nucleinsäuremoleküle oder Vektoren, wobei die Antisense-Sequenz oder das Ribozym spezifisch ist für eine mRNA, die ein Cytokin oder ein co-stimulatorisches Molekül codiert.

Die rekombinanten Nucleinsäuremoleküle oder Vektoren der vorliegenden Ausführungsform können dazu verwendet werden, in Antigen-beladenen dendritischen Zellen gezielt die Expressionen eines Cytokins oder co-stimulatorischen Moleküls zu inhibieren. Hierzu können Antigen-beladene dendritische Zellen in vitro transfiziert werden, wobei die Antigen-Beladung vorzugsweise durch Co-Transfektion mit einem erfindungsgemäßen rekombinanten Nucleinsäuremolekül oder Vektor, der ein Antigen codiert, durchgeführt wird. Durch die gezielte Inhibition eines Cytokins oder co-stimulatorischen Moleküls läßt sich die Immunantwort gegen das Antigen modulieren. Weiterhin können die rekombinanten Nucleinsäuremoleküle oder Vektoren der vorliegenden Ausführungsform auch in vivo appliziert werden, um eine allgemeine Expression spezifisch für dendritische Zellen auszulösen. Durch eine solche Vorgehensweise ist es möglich, die allgemeine Immunsituation eines Patienten zu modulieren.

In einer weiteren besonders bevorzugten Ausführungsform der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren codiert die Nucleotidsequenz einen Transkriptionsfaktor.

Vorzugsweise handelt es sich bei den Transkriptionsfaktoren um solche, die die endogene Expression von Cytokinen oder co-stimulatorischen Molekülen in dendritischen Zellen induzieren. Die Expression eines Transkriptionsfaktors kann

dabei die Induktion von mehreren Genen für Cytokine oder co-stimulatorische Faktoren gleichzeitig bewirken.

Die vorliegende Ausführungsform der rekombinanten Nucleinsäuremoleküle oder Vektoren kann allerdings auch dazu verwendet werden, die endogene Expression von Cytokinen oder co-stimulatorischen Molekülen in dendritischen Zellen zu inhibieren. Es ist bekannt, daß Transkriptionsfaktoren in Kombination mit anderen (endogenen) Transkriptionsfaktoren eine Repressoraktivität besitzen können.

In einer besonders bevorzugten Ausführungsform der oben beschriebenen rekombinanten Nucleinsäuremoleküle oder Vektoren enthalten die rekombinanten Nucleinsäuremoleküle oder Vektoren neben der einen zu exprimieren Nucleotidsequenz eine zweite zu exprimierende Nucleotidsequenz, wobei eine Nucleotidsequenz ein wie oben definiertes Antigen codiert und die zweite Nucleotidsequenz ein Protein, das eine Immunantwort reguliert. Die Immunantwort kann dadurch reguliert werden, daß die zweite Nucleotidsequenz ein Cytokin, ein co-stimulatorisches Molekül, ein Apoptose-induzierendes Molekül, einen Transkriptionsfaktor, eine Antisense-Sequenz oder ein Ribozym exprimiert. Dabei können diese zwei Nucleotidsequenzen in einem Leserahmen hintereinander geschaltet, d. h. translational fusioniert, sein (falls beide Nucleotidsequenzen Protein-codierend sind). Dabei können diese codierenden Regionen direkt aneinander grenzen oder durch einen Spacer getrennt sein. Durch einen Spacer wird die Tertiärstruktur der beiden Proteine voneinander räumlich getrennt, um zu verhindern, daß ihre Tertiärstrukturen negativ interagieren. Vorzugsweise hat der Spacer jedoch die Funktion, als Angriffsstelle für eine Protease zu fungieren, vorzugsweise für eine endogene Protease der transfizierten Zelle, so daß die exprimierten Proteine in vivo getrennt werden. Alternativ kann der Spacer eine IRES-Sequenz (internal ribosomal entry site) enthalten. Dies erlaubt die Transkription beider Gene unter der Kontrolle eines einzigen Promotors, wobei deren Translation getrennt erfolgt.

Andererseits können die zwei Nucleotidsequenzen auch unabhängig voneinander transkriptionell codiert werden. Hierzu steht jede Nucleotidsequenz unter der Kontrolle eines eigenen Promotors, wobei mindestens ein Promotor, vorzugsweise

beide Promotoren die regulatorischen Sequenzen der vorliegenden Erfindung umfaßt.

Mit dieser Ausführungsform ließen sich die besonderen Vorteile, die eine Co-Transfektion mit Expressionskonstrukten, die ein Antigen bzw. ein immunregulatorisches Protein codieren, auf die in-vivo-Situation übertragen. Nach den oben beschriebenen Ausführungsformen ist solch eine Co-Transfektion von dendritischen Zellen nur in vitro möglich.

Durch die Anwendung von rekombinanten Nucleinsäuremolekülen oder Vektoren der vorliegenden Ausführungsform ist es möglich, durch direkte Applikation in den Körper einerseits gezielt eine Antigen-Präsentation durch dendritische Zellen hervorzurufen, andererseits die induzierte Immunantwort durch die Aktivität eines Mediators zu regulieren, wobei Nebenwirkungen durch unspezifische Wirkungen beispielsweise des Mediators weitgehend ausgeschlossen sind. Diese Ausführungsform beinhaltet sämtliche Antigene und immunregulatorischen Proteine, die in den vorgenannten Ausführungsformen definiert wurden.

In einer besonders bevorzugten Ausführungsform sind die oben beschriebenen Vektoren Viren.

Im Stand der Technik ist eine Vielzahl von viralen Vektoren zur Transfektion von Säugerzellen ex vivo oder in vivo beschrieben. Dabei handelt es sich stets um Abkömmlinge von Säuger- bzw. human-pathogenen Viren, die durch genetische Modifizierung ihrer pathogenen Eigenschaften beraubt wurden. Zur Transfektion werden virale Vektoren nach dem Fachmann bekannten Methoden in vitro verpackt, d. h. mit viralen Hüllproteinen versehen. Es können DNA- sowie auch RNA-Viren verwendet werden. Beispiele für Viren zur Transfektion von Säuger-, vorzugsweise Humanzellen sind Herpesvirus, Retroviren, Adenoviren und Adeno-assoziierte Viren.

In einer weiteren Ausführungsform der Erfindung sind die oben beschriebenen Vektoren zur Gentherapie oder DNA-Vakzinierung geeignet.

Gentherapie und DNA-Vakzinierung basieren auf der Einführung von therapeutischen bzw. immunisierenden Genen in Zellen ex vivo oder in vivo. Geeignete Vektoren oder Vektorsysteme und Methoden zu deren Einsatz für Gentherapie oder DNA-Vakzinierung sind in der Literatur beschrieben und dem

Fachmann bekannt, siehe beispielsweise Giordano, *Nature Medicine* 2 (1996), 534-539; Schaper, *Circ. Res.* 79 (1996), 911-919; Anderson, *Science* 256 (1992), 808-813; Isner, *Lancet* 348 (1996), 370-374; Muhlhauser, *Circ. Res.* 77 (1995), 1077-1086; Wang, *Nature Medicine* 2 (1996), 714-716; WO 94/29469; WO 97/00957; Schaper, *Current Opinion in Biotechnology* 7 (1996), 635-640; oder Verma, *Nature* 389 (1997), 239-242 und darin zitierte Referenzen. Die oben beschriebenen rekombinanten Nucleinsäuremoleküle oder Vektoren können zur direkten Einführung oder zur Einführung über Liposomen oder virale Vektoren, z. B. adenoviral oder retroviral, ausgelegt werden.

In einer weiteren bevorzugten Ausführungsform der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren ist die zu exprimierende Nucleotidsequenz ein Reportergen.

Beispiele für Reportergene, mit denen man vorzugsweise in eukaryontischen Zellen, die Expressionsaktivität von regulatorischen Sequenzen, vorzugsweise von Promotoren detektieren kann, sind in der Literatur beschrieben. Beispiele für Reportergene codieren Luciferase, grün-fluoreszierendes Protein, β -Galactosidase oder Chloramphenicol-Acetyltransferase.

Eine weitere Ausführungsform der vorliegenden Erfindung betrifft ein Verfahren zur Herstellung von genetisch modifizierten Wirtszellen, dadurch gekennzeichnet, daß man Wirtszellen mit einem der oben beschriebenen Vektoren transfiziert und die transfizierte Wirtszelle in einem Kulturmedium kultiviert.

Der Begriff „genetisch modifiziert“ bedeutet, daß die Wirtszelle oder der Wirt zusätzlich zum natürlichen Genom ein Nucleinsäuremolekül oder einen Vektor der vorliegenden Erfindung enthält, der in die Wirtszelle oder den Wirt oder einen Vorgänger eingeführt wurde. Das Nucleinsäuremolekül oder der Vektor kann in der genetisch modifizierten Wirtszelle/Wirt entweder als ein unabhängiges Molekül außerhalb des Genoms, vorzugsweise als replizierbares Molekül, oder stabil in das Genom der Wirtszelle oder des Wirts integriert vorliegen.

Die Einführung eines Vektors in Wirtszellen kann durchgeführt werden nach bekannten Standardverfahren, wie zum Beispiel in Sambrook et al. (a.a.O.) beschrieben. Beispiele für anwendbare Transfektionstechniken sind

Calciumphosphat-Transfektion, DEAE-Dextran-vermittelte Transfektion, Elektroporation, Transduktion, Infektion, Lipofektion oder biolistischer Transfer. Zur anschließenden Kultivierung kann man ebenfalls auf Standardverfahren zurückgreifen oder, bei der genetischen Modifizierung von dendritischen Zellen, vorzugsweise nach dem in den Beispielen und den darin zitierten Referenzen beschriebenen Verfahren.

In einer weiteren bevorzugten Ausführungsform betrifft die Erfindung Wirtszellen, die genetisch modifiziert sind mit einer regulatorischen Sequenz, einem rekombinanten Nucleinsäuremolekül oder einem Vektor der vorliegenden Erfindung oder erhältlich sind nach dem oben beschriebenen Verfahren.

Die Wirtszelle der vorliegenden Erfindung kann prinzipiell jede prokaryontische oder eukaryontische Zelle sein, und umfaßt unter anderem Säugerzellen, Pilzzellen, Pflanzenzellen, Insektenzellen oder Bakterienzellen. Geeignete bakterielle Zellen sind solche, die allgemeine Verwendung finden für Clonierungen, wie z. B. *E.coli* oder *Bacillus subtilis*.

Beispiele für Pilzzellen sind Hefezellen, vorzugsweise solche der Gattungen *Saccharomyces* oder *Pichia*, besonders bevorzugt von *Saccharomyces cerevisiae* oder *Pichia pastoris*. Unter geeigneten tierischen Zellen sind beispielsweise Insektenzellen, Vertebratenzellen, vorzugsweise Säugerzellen, wie z. B. CHO, Hela, NIH3T3, MOLT-4, Jurkat, K562, HepG2 oder PC12 zu zählen. Weitere geeignete Zelllinien sind im Stand der Technik beschrieben und lassen sich beispielsweise von der Deutschen Sammlung für Mikroorganismen und Zellkulturen (DSMZ, Braunschweig) beziehen.

Besonders bevorzugt ist die Ausführungsform der Wirtszellen, die dendritische Zellen sind.

Dendritische Zellen sind der primäre Anwendungsort der erfindungsgemäßen regulatorischen Sequenzen, rekombinanten Nucleinsäuremoleküle oder Vektoren. Dendritische Zellen lassen sich beispielsweise aus peripheren Blutleukocyten und Langerhanszellen aus epidermalen Präparationen gewinnen, wobei es sich bei den isolierten Zellen auch um Vorläuferzellen handeln kann, die durch eine geeignete in

vitro-Kultivierung zu dendritischen Zellen umgewandelt werden können. Methoden hierzu sind im Stand der Technik beschrieben und lassen sich beispielsweise auch aus den Beispielen wie auch aus Ross (J. Invest. Dermatol. 115 (2000), 658-663), Ross (J. Immunol. 160 (1998), 3776-3782) oder darin zitierten Referenzen entnehmen. Diese Quellen liefern ebenfalls Methoden zur Kultivierung von dendritischen Zellen.

Besonders bevorzugt sind im Rahmen der Erfindung Wirtszellen, die aus Mensch stammen.

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft Nucleotidsequenzen umfassend ein Fragment mit einer Länge von mindestens 15 Nucleotiden, das spezifisch mit einem Strang einer erfindungsgemäßen regulatorischen Sequenz unter stringenten Bedingungen hybridisiert.

Hybridisierende Nucleotidsequenzen gemäß der vorliegenden Ausführungsform können beispielsweise als Sonden dienen, die beispielsweise zur Identifikation homologer Promotoren, vorzugsweise regulatorischer Sequenzen anderer Gene, die aufgrund bestimmter übereinstimmender Sequenzelemente ein den erfindungsgemäßen regulatorischen Sequenzen vergleichbares Expressionsmuster induzieren, beitragen. Des weiteren können diese Sequenzen zum Design geeigneter Oligonucleotide, beispielsweise als PCR-Primer, verwendet werden.

Der Begriff „Hybridisierung“ ist weiter oben bereits definiert worden. Vorzugsweise hybridisieren die Nucleotidsequenzen unter stringenten Bedingungen. Die Fragmente haben eine Länge von mindestens 15 Nucleotiden, vorzugsweise von mindestens 20 Nucleotiden, besonders bevorzugt von mindestens 50 Nucleotiden, insbesondere bevorzugt von mindestens 100 Nucleotiden, vorteilhafterweise von mindestens 200 Nucleotiden und am meisten bevorzugt von mindestens 500 Nucleotiden.

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft ein Verfahren zur Antigen-spezifischen Stimulierung von T-Zellen in vitro, enthaltend die Schritte:

- (a) Transfektion von dendritischen Zellen mit einem Vektor, der eine Nucleotidsequenz enthält, die ein Antigen codiert, allein oder in Kombination

mit einem Vektor, der eine Nucleotidsequenz enthält, die ein immunregulatorisches Protein codiert, wobei die Regulation vorzugsweise eine Steigerung der Immunantwort ist, und wobei die Nucleotidsequenz vorzugsweise ein Cytokin oder ein co-stimulatorisches Molekül, eine Antisense-Sequenz oder ein Ribozym exprimiert, oder mit einem Vektor, der sowohl ein Antigen als auch ein immunregulatorisches Protein codiert;

- (b) Cokultivierung der durch Schritt (a) gewonnenen transfizierten dendritischen Zellen mit T-Zellen; und
- (c) Detektion der Aktivierung der T-Zellen aus Schritt (b).

Die Verfahrensschritte sind im Prinzip im Stand der Technik beschrieben und lassen sich beispielsweise WO94/02156 entnehmen. Auf die Bereitstellung von dendritischen Zellen ist bereits weiter oben eingegangen worden. Ein Verfahren hierzu läßt sich den Beispielen im Anhang entnehmen. Die Präparation von T-Zellen ist nach Verfahren aus dem Stand der Technik vorzunehmen, wie beispielsweise in "Current Protocols in Immunology" (Coligan, J.E., Kruisbeek, A.M., Margulies, D.H., Shevach, E.M., Strober, W., eds., Greene Publishing Associates and Wiley-Interscience, New York, 1991, Vol. 1) ausgeführt. Der Vorteil des hier vorgestellten Verfahrens gegenüber dem Stand der Technik betrifft die Verwendung der erfindungsgemäßen Vektoren. Da diese die Expression in dendritischen Zellen in spezifischer Weise erlauben, lassen sich nun Zellpopulationen verarbeiten, die einen weitaus niedrigeren Aufreinigungsgrad für dendritische Zellen benötigen als für herkömmlichen Verfahren notwendig ist, ohne Nebenwirkungen oder Risiken durch die Transfektion von nicht-dendritischen Zellen befürchten zu müssen.

Bei den in Schritt (b) zur Co-Kultivierung eingesetzten T-Zellen kann es sich um naive oder aktivierte T-Zellen handeln. Die Detektion der Aktivierung der T-Zellen in Schritt (c) kann nach einem oder mehreren der folgenden Methoden erfolgen: Messung der Proliferation, Detektion der cytotoxischen Aktivität, Detektion der Cytokinproduktion, Detektion der metabolischen Aktivität und Detektion von Aktivierungsmarkern.

Eine weitere bevorzugte Ausführungsform betrifft ein Verfahren zur Herstellung eines Arzneimittels umfassend die Schritte (a) bis (c) des Verfahrens zur Antigen-spezifischen Stimulierung von T-Zellen in vitro und zusätzlich den Schritt

- (d) Formulierung eines Arzneimittels durch Versetzen der durch den Schritt (c) gewonnenen stimulierten T-Zellen mit einem pharmazeutisch verträglichen Träger.

Zur Formulierung von Zellen zur Verabreichung als Arzneimittel werden die Zellen in einem pharmazeutisch verträglichen Trägermaterial suspendiert. Dies gilt für die stimulierten T-Zellen wie auch für die im folgenden beschriebenen dendritischen Zellen. Beispiele für Trägermaterial sind Wasser, Kochsalzlösung, Dextrose, Glycerin, etc. oder Kombinationen davon. Zusätzlich kann die zu verabreichende Zellsuspension weitere Stoffe wie Emulsionsmittel, pH-Puffer, Adjuvanzen, oder auch immunregulatorische Faktoren, wie Cytokine, enthalten.

T-Zellen, die nach dem oben beschriebenen Verfahren stimuliert wurden, können beispielsweise zur Bekämpfung von schweren Virusinfektionen, wie z.B. gezeigt für CMV bei immunschwachen Patienten (Walter, New Engl. J. Med. 333 (1995), 1038-1044), oder zur Induktion der Immunabwehr von Metastasen (Nestle, Nature Med. 4 (1998), 328-332) eingesetzt werden.

Eine weitere Ausführungsform der Erfindung betrifft ein Verfahren zur Herstellung von T-Zell-stimulierenden dendritischen Zellen in vitro, enthaltend die Schritte:

- (a) Transfektion von dendritischen Zellen mit einem Vektor, der eine Nucleotidsequenz enthält, die ein Antigen codiert, allein oder in Kombination mit einem Vektor, der eine Nucleotidsequenz enthält, die ein immunregulatorisches Protein, ein Apoptose-induzierendes Molekül, vorzugsweise zur TNF-Superfamilie gehörend, codiert oder eine Antisense-Sequenz oder ein Ribozym exprimiert, oder mit einem Vektor, der sowohl ein Antigen als auch ein immunregulatorisches Protein codiert; und
- (b) Kultivierung der transfizierten dendritischen Zellen in einem geeigneten Medium und/oder Detektion der T-Zell-stimulierenden Aktivität.

Eine weitere bevorzugte Ausführungsform betrifft ein Verfahren zur Herstellung eines Arzneimittels umfassend die Schritte (a) und (b) des Verfahrens zur Herstellung von T-Zellen stimulierenden dendritischen Zellen und zusätzlich den Schritt

- (c) Formulierung eines Arzneimittels durch Versetzen der durch den Schritt (b) gewonnenen T-Zell-stimulierenden dendritischen Zellen mit einem pharmazeutisch verträglichen Träger.

Die durch dieses Verfahren gewonnenen dendritischen Zellen können als Arzneimittel Patienten verabreicht werden, um durch T-Zell-Aktivierung gezielte Immunantworten auszulösen. Die dendritischen Zellen können intradermal, subcutan, intravenös oder, im Falle einer Tumorbehandlung, in Regionen des Tumorwachstums oder in Lymphgefäße, die diese Regionen drainieren, injiziert werden.

In einer besonders bevorzugten Ausführungsform des oben beschriebenen Verfahrens stammen die dendritischen Zellen aus Mensch.

Eine weitere bevorzugte Ausführungsform der vorliegenden Erfindung betrifft Arzneimittel umfassend die rekombinanten Nucleinsäuremoleküle und die Vektoren der vorliegenden Erfindung, die Wirtszellen, Antigen-spezifisch stimulierte T-Zellen, erhältlich nach dem oben beschriebenen Verfahren, oder T-Zell-stimulierende dendritische Zellen, erhältlich nach dem oben beschriebenen Verfahren, und optional einen pharmazeutisch verträglichen Träger.

In einer besonders bevorzugten Ausführungsform der Erfindung ist das Arzneimittel ein Impfstoff.

Die regulatorischen Sequenzen, rekombinanten Nucleinsäuremoleküle, Vektoren oder Wirtszellen der Erfindung können verwendet werden zur Herstellung eines Impfstoffs, d. h. im Sinne der Erfindung eines DNA-Vakzins. Modi der Verabreichung von DNA-Vakzinen sind im Stand der Technik beschrieben und DNA-Vakzinierung wurde bereits erfolgreich eingesetzt, um Antitumorimmunantworten auszulösen (Tighe M. et al., Immunology Today 19 (1998), 89-97). Darüber hinaus konnte

bereits gegen verschiedene Krankheitsformen ein Immunschutz durch Verabreichung von DNA-Nucleinsäuremolekülen erzielt werden (Fynan, Proc. Natl. Acad. Sci. U.S.A. 90 (1993), 11478-11482; Boyer, Nat. Med. 3 (1997), 526-532; Webster, Vaccine 12 (1994), 1495-1498; Montgomery et al., DNA Cell Biol. 12 (1993), 777-783; Barry, Nature 311 (1995), 632-635; Xu and Liew, Immunology 84 (1995), 173-176; Zhoug, Eur. J. Immunol. 26 (1996), 2749-2757; Luke, J. Inf. Dis. 175 (1997), 91-97; Mor, Biochem. Pharmacology 55 (1998), 1151-1153; Donnelly, Annu. Rev. Immun. 15 (1997), 617-648; MacGregor, J. Infect. Dis. 178 (1998), 92-100).

Die erfindungsgemäßen Nucleinsäuremoleküle zur Verwendung als Impfstoff können in einer neutralen Form oder als Salz formuliert werden. Pharmazeutisch wirksame Salze sind dem Fachmann geläufig. Die Impfstoffe der Erfindung können, unter anderem, verwendet werden zur Behandlung und/oder Prävention von Infektionen und werden in Dosierungen verabreicht, die pharmakologisch wirksam sind zur Prophylaxe oder Behandlung.

Die im Sinne der Erfindung anzuwendenden Vakzinierungsprotokolle beinhalten aktive Immunisierung, wobei die Verabreichung von Nucleinsäuremolekülen, die gezielt in den dendritischen Zellen einer Person Antigene oder Allergene exprimieren, eine protektive Immunantwort auslösen sollen. Zusätzlich beinhalten die Vakzinierungsprotokolle die Kombination der Antigenexpression mit immunmodulatorischen Effekten durch zusätzliche Verabreichung von Nucleinsäuren, die entsprechende Mediatoren ebenfalls gezielt in dendritischen Zellen exprimieren oder auch weitere unterstützende medikamentöse Behandlung.

Weitere Strategien zur Erreichung eines Impfschutzes beinhalten die Verabreichung in vitro-transfizierter dendritischer Zellen, in vitro-aktivierter T-Zellen, jeweils nach den oben beschriebenen Verfahren. Methoden und Prinzipien sind dem Stand der Technik zu entnehmen und zum Beispiel beschrieben in Paul, "Fundamental Immunology" Raven Press, New York (1989) oder Morein, "Concepts in Vaccine Development", ed: S.H.E. Kaufmann, Walter de Gruyter, Berlin, New York (1996), 243-264.

Typischerweise werden Impfstoffe zur Injektion hergestellt als flüssige Lösung oder Suspension. Die Präparationen können emulgiert sein oder der Wirkstoff kann in Liposomen eingekapselt werden. Die Wirkstoffe werden oftmals mit Trägermaterial

gemischt, die mit dem Wirkstoff kompatibel sind. Beispiele für Trägermaterial sind Wasser, Kochsalzlösung, Dextrose, Glycerin, Ethanol etc. oder Kombinationen davon. Der Impfstoff kann auch Hilfssubstanzen wie Emulsionsmittel, pH-Puffer und/oder Adjuvanzen enthalten.

Statt durch Injektion kann DNA durch biolistischen Transfer (US-Patent No. 5,100,702; Kalkbrenner, Meth. Mol. Biol. 83, 1996, 203-216) zur Vakzinierung verabreicht werden. Hierzu wird DNA, d. h. rekombinante Nucleinsäuremoleküle oder Vektoren der vorliegenden Erfindung an kleine Partikel, beispielsweise Goldpartikel oder Partikel aus biokompatiblen Material gebunden und beschleunigt durch Gasdruck in die Epidermis oder Dermis eingebracht. DNA kann ebenfalls oral oder sublingual verabreicht werden oder auf die Schleimhäute des Respirationstraktes durch nasale oder intratracheale Applikation aufgetragen werden (Beispiele hierzu in Etchart, J. Gen. Virol. 78 (1997), 1577-1580 oder McCluskie, Antisense and Nucleic Acid Drug Development 8 (1998), 401-414).

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft die Verwendung der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren, die eine Nucleotidsequenz exprimieren, die vorzugsweise ein Antigen codiert, das besonders bevorzugt tumor- oder pathogenspezifisch oder ein Allergen ist, wobei der Vektor vorzugsweise ein Virus ist oder zu Gentherapie oder DNA-Vakzinierung geeignet ist, allein oder in Kombination mit den erfindungsgemäßen rekombinanten Nucleinsäuremolekülen oder Vektoren, die ein immunmodulatorisches Protein exprimieren, das vorzugsweise ein Cytokin oder co-stimulatorisches Molekül ist, dessen Regulation vorzugsweise eine Steigerung der Immunantwort ist, oder die einen Transkriptionsfaktor exprimieren, sowie die Verwendung der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren, die sowohl ein Antigen als auch ein immunregulatorisches Protein codieren, der erfindungsgemäßen Wirtszellen, Antigen-spezifisch stimulierten T-Zellen, erhältlich nach dem oben beschriebenen Verfahren, oder T-Zell-stimulierenden dendritischen Zellen, erhältlich nach dem oben beschriebenen Verfahren, zur Herstellung eines Arzneimittels zur Impfung gegen Viren, Bakterien, Pilze, Parasiten, Tumoren, Allergene, Creutzfeldt-Jakob-Plaques oder Alzheimer-Plaques oder zur

gentherapeutischen Behandlung von Tumoren oder viralen, bakteriellen oder parasitischen Infektionen oder von Allergien.

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft die Verwendung der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren, die eine Nucleotidsequenz exprimieren, die vorzugsweise ein Antigen codiert, das besonders bevorzugt ein Autoantigen, Transplantationsantigen oder Allergen ist, wobei der Vektor vorzugsweise ein Virus ist oder zu Gentherapie oder DNA-Vakzinierung geeignet ist, allein oder in Kombination mit den erfindungsgemäßen rekombinanten Nucleinsäuremolekülen oder Vektoren, die ein immunregulatorisches Protein exprimieren, wobei die Regulation vorzugsweise eine Hemmung ist, oder die ein Apoptose-induzierendes Molekül, einen Transkriptionsfaktor, oder eine Antisense-Sequenz oder ein Ribozym exprimieren, sowie die Verwendung der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren, die sowohl ein Antigen als auch ein immunregulatorisches Protein codieren, oder der erfindungsgemäßen Wirtszellen zur Herstellung eines Arzneimittels zur Behandlung von Autoimmunerkrankungen, Transplantatabstoßung oder Allergien.

Eine weitere bevorzugte Ausführungsform der Erfindung betrifft die Verwendung der erfindungsgemäßen rekombinanten Nucleinsäuremoleküle oder Vektoren, die ein Apoptose-induzierendes Molekül codieren, das vorzugsweise zur TNF-Superfamilie gehört, wobei der Vektor vorzugsweise ein Virus ist oder zu DNA-Vakzinierung oder Gentherapie geeignet ist, zur Herstellung eines Arzneimittels zur Vermeidung der Abstoßung von Transplantaten und von Autoimmunreaktionen. Dendritische Zellen, welche in einer Blutprobe des Spenders eines Transplantates enthalten sind, exprimieren nach Aufnahme des Expressionsvektors das codierte Apoptose-induzierende Molekül. Nach Injektion in den Empfänger des Transplantates bringen diese dendritischen Zellen T-Zellen, die die Transplantationsantigene spezifisch erkennen, zum Absterben. Bei Autoimmunreaktionen bringen Autoantigen-beladene dendritische Zellen, welche nach Aufnahme des Expressionsvektors das codierte Apoptose-induzierende Molekül exprimieren, T-Zellen, die das Autoantigen spezifisch erkennen, zum Absterben.

Ferner betrifft die vorliegende Erfindung die Verwendung der erfindungsgemäßen regulatorischen Sequenzen, der rekombinanten Nucleinsäuremoleküle oder Vektoren zur spezifischen Expression von Antigenen oder immunregulatorischen Proteinen in dendritischen Zellen.

In einer weiteren bevorzugten Ausführungsform betrifft die vorliegende Erfindung die Verwendung der erfindungsgemäßen regulatorischen Sequenzen oder der rekombinanten Nucleinsäuremoleküle oder Vektoren, die vorzugsweise ein Reportergen exprimieren, zur Identifizierung und Isolierung von cis-Elementen aus der regulatorischen Sequenz, die eine für dendritische Zellen spezifische Expression vermitteln.

In einer weiteren bevorzugten Ausführungsform betrifft die vorliegende Erfindung die Verwendung der erfindungsgemäßen regulatorischen Sequenzen oder der rekombinanten Nucleinsäuremoleküle oder Vektoren, die vorzugsweise ein Reportergen exprimieren, zur Bestimmung des Reifegrads von dendritischen Zellen. Diese Ausführungsform läßt sich beispielsweise zur Bestimmung des Reifegrads von in vitro-kultivierten dendritischen Zellen, die in klinischen Studien eingesetzt werden sollen, nutzen.

Eine weitere Ausführungsform der Erfindung betrifft die Verwendung der erfindungsgemäßen regulatorischen Sequenzen oder der rekombinanten Nucleinsäuremoleküle oder Vektoren zur Identifizierung und Isolierung von Faktoren, die eine Expression spezifisch für dendritische Zellen vermitteln.

Ferner betrifft die vorliegende Erfindung die Verwendung der erfindungsgemäßen regulatorischen Sequenzen, die eine Nucleotidsequenz aus einer der unter SEQ ID NO: 1 bis 8 angegebenen Sequenzen bzw. einer entsprechenden Promotorsequenz aus SEQ ID NO:72 umfassen oder die eine Nucleotidsequenz umfassen, die in der Insertion des Klons DSM13274 enthalten ist und durch Amplifikation unter Verwendung eines Paares von Oligonucleotiden erhältlich ist, deren Sequenzen unter einem der folgenden Paare von SEQ ID-Nummern angegeben ist: 36 und 37; 38 und 37; 39 und 37; 40 und 37; 41 und 37; 42 und 37; 43 und 37; oder 44 und 37;

Teilen davon oder Sequenzen, die mit den vorgenannten spezifisch hybridisieren, zur Blockierung von Transkriptionsfaktoren durch Bereitstellung von Transkriptionsfaktor-Bindungsstellen in dendritischen Zellen. Da die erfindungsgemäßen regulatorischen Sequenzen in ihrem Ursprungsgen eine Stadien-spezifische Expression in dendritischen Zellen vermitteln (Fascin wird nicht in unreifen dendritischen Zellen, wohl aber zunehmend in reiferen Stadien exprimiert), ist es möglich, durch Blockierung von Transkriptionsfaktoren, die die Stadienspezifität vermitteln, den Übergang von unreifen dendritischen Zellen in reifere Stadien zu inhibieren. Hierzu können die regulatorischen Sequenzen der Erfindung bzw. Teile davon, vorzugsweise in Form von Oligonucleotiden verwendet werden. Durch die Inhibition der Ausreifung der dendritischen Zellen wird indirekt die Primärstimulation von T-Zellen inhibiert. Dies kann beispielsweise bei Gewebetransplantationen zur Verhinderung von Abstoßungsreaktionen genutzt werden.

Diese und andere Ausführungsformen sind dem Fachmann offenbart und offensichtlich und umfaßt durch die Beschreibung und die Beispiele der vorliegenden Erfindung. Weiterführende Literatur kann zu einer der oben angeführten Methoden, Mittel und Verwendungen, die im Sinne der vorliegenden Erfindung angewendet werden können, dem Stand der Technik entnommen werden, z. B. aus öffentlichen Bibliotheken unter z. B. der Benutzung von elektronischen Hilfsmitteln. Zu diesem Zweck bieten sich unter anderem öffentliche Datenbanken an wie die "Medline", die über Internet zur Verfügung stehen, z. B. unter der Adresse <http://www.ncbi.nlm.nih.gov/PubMed/medline.html>. Weitere Datenbanken und Adressen sind dem Fachmann geläufig und können aus dem Internet entnommen werden, z. B. unter der Adresse <http://www.lycos.com>. Eine Übersicht über Quellen und Informationen zu Patenten bzw. Patentanmeldungen in der Biotechnologie ist in Berks, TIBTECH 12 (1994), 352-364 gegeben.

Die Figuren zeigen:

Figur 1: Genomische Organisation des humanen Fascingens. *Oben:* Schematische Wiedergabe des Genlocus. Das Gen erstreckt sich über

etwa 13 kb und besteht aus fünf Exons (als Kästchen hervorgehoben, grau: untranslatierte, schwarz: proteincodierende Bereiche). Unten: Lage und Größe der aus dem PAC-Klon RPCIP704C24766Q3/4 subklonierten genomischen Fragmente, die für detailliertere Studien verwendet worden sind. Die für die jeweilige Subklonierung genutzten Restriktionsschnittstellen sind *oben* eingezeichnet. H: *Hind III*, Hi: *Hinc II*, P: *Pst I*, S: *Sac I*, E: *Eco RI*.

Figur 2: Nucleotidsequenz des humanen Fascingens. Exonsequenzen sind fett wiedergegeben. Das Start- und Stopcodon sowie das putative Polyadenylierungssignal sind jeweils doppelt unterstrichen. Exon-Intron-Spleißstellen sind kursiv und unterstrichen. (Die Länge der noch nicht sequenzierten Genabschnitte ist in eckigen Klammern angegeben. Für zwei Teilsequenzen in Intron 1 ist die korrekte Orientierung wegen der flankierenden Sequenzierlücken noch nicht bekannt.)

Figur 3: Expressionsstärke des humanen Fascinpromotors (pFascin-3.0) in dendritischen Zellen (DC), die aus CD14⁺ Vorläuferzellen des peripheren Blutes generiert wurden, im Vergleich zur Negativkontrolle (pGL3-Basic). Angegeben ist die normalisierte Expressionsstärke des getesteten Promoterfragments als Quotient aus den Luciferaseaktivitäten für das Testkonstrukt pFascin-3.0 (Photinus-Luciferase) und den konstitutiv exprimierten Coreporter pRL-CMV (Renilla-Luciferase). Angegeben ist der Mittelwert $\pm$ SEM (Standardfehler) für die in Triplikatansätzen getesteten Konstrukte.

Figur 4: Lage und Größe der getesteten Deletionskonstrukte des humanen Fascinpromotors im Vergleich zum Fascin-Gen (oben). Dargestellt sind der 5'-genflankierende genomische Bereich (ungefüllter Abschnitt), der transkribierte nichttranslatierte 5'-Genbereich (5'-UTR, grau hinterlegt) und der translatierte Abschnitt (schwarz hinterlegt) von Exon 1 sowie der flankierende Anteil von Intron 1. Der putative Promotorbereich wurde in den Vektor pGL3-Basic vor das promotorlose Reportergen Photinus-

Luciferase cloniert (pFascin-3.0). 5'-verkürzte Deletionskonstrukte wurden durch gerichtete „nested deletion“ bzw. Religationen hergestellt. Die jeweilige Größe der Promotorkonstrukte (abzüglich des 5'-UTR-Anteils) ist als Klonbezeichnung (in Kb) angegeben.

Figur 5: Nachweis der Abwesenheit endogener Fascinexpression in der humanen Monocytenlinie THP-1 mittels RT-PCR. **A:** Amplifiziert wurde ein 266 bp großes Fragment aus dem 3'-UTR der Fascin-cDNA (hFascin-UTR). *Spur 1:* Positivkontrolle (in pUC18 kloniertes PCR-Produkt hFascin-UTR); *2:* Negativkontrolle (H₂O); *3:* SHSY5Y (Neuroblastomlinie; exprimiert konstitutiv Fascin); *4:* THP-1. **B:** Für die in A dargestellte Fascin-PCR wurden die einzusetzenden cDNA-Mengen mittels HPRT-PCR standardisiert. *Spur 1:* THP-1; *2:* SHSY5Y; *3:* Negativkontrolle. In **A** und **B** ist in der jeweils ersten Spur der Molekulargewichtsmarker (ϕ X174, *Hae* III-restringiert) aufgetragen.

Figur 6: Expressionsstärken der Deletionskonstrukte des humanen Fascinpromoters in DC und der Monocytenlinie THP-1. Die normalisierten Werte für die Reportergenexpression der einzelnen Promotorkonstrukte sind auf die Expressionsstärke des Minimalpromoters (pFascin-0.11) mit der Einheit eins normiert. Angegeben ist der Mittelwert $\pm$ SEM für die in Triplikatansätzen getesteten Konstrukte.

Figur 7: Vergleich der Aktivität des Fascinpromoters in CD83⁺ und CD83⁻ Zellen einer DC-Kultur. Die relative Expressionsstärke dreier Humanfascin-Promotorkonstrukte zunehmender Länge (*pFascin-0.11*, *pFascin-1.6*, *pFascin-3.0*) wurde vergleichend in der CD83-positiven und -negativen Zellfraktion einer DC-Kultur getestet. Angegeben ist die relative Expressionsstärke jedes Promotorkonstrukts im Vergleich zum kürzesten Konstrukt mit Basalaktivität (*pFascin-0.11*), dessen relative Luciferaseaktivität auf 1 normiert ist. Als Negativkontrolle diente der promotorlose Vektor pGL3-Basic (*pGL3-Basic*). Angegeben sind die Mittelwerte $\pm$ SEM von zwei Experimenten, die jeweils in Triplikaten

durchgeführt wurden. Unterschiede in der Transfektionseffizienz wurden durch Cotransfektion mit pRL-CMV normalisiert.

Figur 8: Zelltypspezifische Aktivität des humanen Fascinpromotors. Die zelltypspezifische Aktivität des humanen Fascinpromotors (*pFascin-3.0*) ist relativ zur Expressionsstärke einer Positivkontrolle angegeben. Als Positivkontrolle wurde ein Luciferasekonstrukt eingesetzt, bei dem die Expression des Reporters Luciferase unter Kontrolle des Promotors des Haushaltsgens EF1 α erfolgt, der in den untersuchten Zellen eine starke, Zelltyp-unabhängige Expression vermittelt. Beide Konstrukte wurden in humanen Zelltypen getestet: einer Fascin-negativen Keratinozytenlinie (*HaCaT*), einer Fascin-exprimierenden Neuroblastomlinie (*SHSY-5Y*) sowie in stark Fascin-positiven reifen Dendritischen Zellen (*DC*). Angegeben sind die Mittelwerte $\pm$ SEM von drei (*HaCaT*, *SHSY-5Y*) bzw. zwei (*DC*) Experimenten, die jeweils in Triplikaten durchgeführt wurden. Unterschiede in der Transfektionseffizienz wurden durch Cotransfektion mit pRL-CMV normalisiert.

Figur 9: Nucleotidsequenz des humanen Fascingens. Exonsequenzen sind fett wiedergegeben. Das Start- und Stopcodon sowie das putative Polyadenylierungssignal sind jeweils doppelt unterstrichen. Exon-Intron-Spleißstellen sind kursiv und unterstrichen.

Die nachfolgenden Beispiele erläutern die Erfindung.

Methodisches Setup

Zellkultur und Erzeugung von humanen dendritischen Zellen

Die humane Monozytenlinie THP-1 (Tsuchiya, Int. J. Cancer 26 (1980), 171-176) wurde in RPMI 1640 kultiviert, die humane Keratinozytenlinie HaCaT (Boukamp, J Cell Biol. 106 (1988), 761-771) in IMDM und die humane Neuroblastomlinie SHSY-5Y (Vinores, Cancer Res. 44 (1984), 2595-2599) in einer gleichvolumigen Mischung

aus DMEM und Nut Mix F12. Die Kulturmedien wurden jeweils mit 10 % fötalem Kälberserum (FCS), 2 mM L-Glutamin, 1 mM Natriumpyruvat, 100 IU/ml Penicillin sowie 100 µg/ml Streptomycin ergänzt. Alle Kulturmedien und -zusätze wurden von Life Technologies (Eggenstein) bezogen.

Menschliche dendritische Zellen (DC) wurden hergestellt aus peripheren mononucleären Blutzellen (peripheral blood mononuclear cells, PBMC) wie beschrieben in Jonuleit (Eur. J. Immunol. 27 (1997), 3135-3142). Sämtliche Zentrifugationsschritte wurden bei Raumtemperatur ausgeführt. Etwa 80 ml *buffy coat*, das aus gesunden Blutspendern gewonnen wurde, wurde 1:3 mit Phosphat-gepufferter Kochsalzlösung (PBS) verdünnt, die 2 mM EDTA und 0,4 IU/ml Natriumheparin enthielt. 15 ml Ficoll-Histopaque (Biochrom, Berlin) wurden mit dem doppelten Volumen des verdünnten *buffy coats* überschichtet. Nach Zentrifugation bei 1000 rpm über 20 min wurde ein 5 ml-Aliquot der oberen Phase (Serumfraktion) entfernt und Hitze-inaktiviert bei 56°C für 30 min. Die Serumfraktion wurde bei 3000 rpm für 5 min zentrifugiert und der Überstand wurde für nachfolgende Anwendungen als autologes Plasma bei 4°C aufbewahrt. Der Ficoll-Gradient wurde ein zweites Mal zentrifugiert (für 15 min bei 1500 rpm) und die PBMC-enthaltende Interphase wurde vorsichtig entfernt unter Verwendung einer Pasteur-Pipette. Die PMBC wurden mit 30 ml PBS, das 2 mM EDTA enthielt, verdünnt und dreimal gewaschen mit PBS/ 2 mM EDTA. Die Zellzahl wurde bestimmt und die Zellen wurden für eine Stunde bei 37°C in X-Vivo 15 (Bio-Whittaker, Walkersville, USA), das 3% des autologen Plasmas enthielt, in 6-Loch-Zellkultur Clusterplatten (Corning Costar, Bodenheim) in einem Volumen von 3 ml (5×10^6 Zellen/ml) inkubiert. Daraufhin wurde der Überstand entfernt, der die nicht adhärenen Leukocyten enthielt, und ersetzt durch neues Kulturmedium, das durch 800 U/ml rekombinantes menschliches (rh) GM-CSF und 1000 U/ml rhIL-4 ergänzt wurde. Die Cytokinbehandlung induzierte die Reifung der ursprünglich adhärenen monocytischen Zellen zu unreifen dendritischen Zellen. An jedem zweiten Tag wurde 1 ml des Mediums ersetzt durch Medium, das mit 1600 U/ml rhGM-CSF und 1000 U/ml rhIL-4 ergänzt wurde. Am Tag 7 wurden die nicht adhärenen Zellen, die die unreifen dendritischen Zellen repräsentieren, geerntet und zweimal mit Kulturmedium gewaschen. Unreife dendritische Zellen wurden bei einer Dichte von 10^6 Zellen/ml in 6-Loch-Platten in 3 ml Kulturmedium, das mit einer Mischung aus Cytokinen, die die weitere Reifung initiieren, ergänzt wurde,

angesetzt. Der Cytokincocktail enthielt rhGM-CSF (800 U/ml), rhIL-4 (500 U/ml), rhIL-1 β (10 ng/ml), rhIL-6 (1000 U/ml), rhTNF α (10 ng/ml) und Prostaglandin E2 (1 μ g/ml). Sämtliche Cytokine wurden erworben bei Strathmann Biotech (Hannover) außer rhGM-CSF (Sandoz, Nürnberg). Prostaglandin E2 wurde erworben bei Pharmacia & Upjohn (Erlangen). Nach 24 Stunden Inkubation wurden die reifen dendritischen Zellen zur Transfektion eingesetzt.

Alternativ wurden die dendritischen Zellen vor der Transfektion in eine CD83-positive und eine CD83-negative Fraktion aufgetrennt. Zur Erkennung des DC-Reifungsmarkers CD83 wurde der murine monoklonale Antikörper HB15e (BD Biosciences, Heidelberg) verwendet. Die CD83-positiven Zellen der DC-Kultur wurden mit antikörperbeschichteten paramagnetischen Kügelchen inkubiert, über magnetische Separation isoliert und nach DNase-vermittelterm Abbau der Kügelchen für die Transfektionen eingesetzt. Für die Zellseparation wurde das CELlection™ Pan Mouse IgG Kit (Dyna, Hamburg) nach den Vorgaben des Herstellers eingesetzt.

Reverse-transcription PCR (RT-PCR)

Es wurde mRNA isoliert unter Verwendung des QuickPrep™ Micro mRNA Purification Kit (Amersham Pharmacia Biotech, Uppsala, Schweden). Das RNA-Pellet wurde in 5 μ l Diethylpyrocarbonat-behandeltem Wasser gelöst. Die reverse Transkriptionsreaktion (RT) wurde mit 1 μ l mRNA in einem Gesamtvolumen von 20 μ l durchgeführt wie beschrieben in Ross (PCR Methods Applic. 4 (1995), 371-375). Die Amplifikations- und reversen Transkriptionsreaktionen wurden durchgeführt unter Verwendung eines Model 480 DNA Thermal Cycler (Perkin-Elmer, Foster City, USA). Sämtliche zur PCR eingesetzten genspezifischen Primer wurden von MWG-Biotech (Ebersberg) hergestellt. Die Menge und Qualität der erhaltenen cDNA wurde geprüft durch Amplifikation eines 366 bp-Fragments des housekeeping-Gens Hypoxanthin-Guanin-Phosphoribosyl-Transferase (HPRT) mit 1 μ l der RT-Reaktion als Matrize. Es wurden die Primer HPRT-3 (5'-GCTGACCTGCTGGATTACAT-3'; SEQ ID NO: 23) und HPRT-4 (5'-CATTATAGTCAAGGGCATATCC-3'; SEQ ID NO: 24) verwendet. PCR-Denaturierung (94°C für 1 min) und Extensionsschritte (72°C

für 1 min) waren konstant. Die Annealing-Temperatur wurde vermindert von 59°C auf 58°C, jeweils für zwei Zyklen, und auf 57°C für weitere 30 Zyklen (jeweils 1 min). Der finale Extensionsschritt (7 min bei 72°C) sicherte die komplette Doppelstrangpolymerisation und war Teil jeder PCR-Reaktion (siehe auch weiter unten). Zur relativen Quantifizierung des amplifizierten HPRT-cDNA-Fragments wurden Aliquots der PCR-Reaktionen einer Gelelektrophorese in 1,4%igem Agarosegel unterzogen und die Fluoreszenzintensitäten der Ethidiumbromid-gefärbten PCR-Produkte wurden mit einem Imaging-System (Herolab, Wiesloch) verglichen. HPRT-standardisierte Mengen der cDNA wurden als Matrize für eine RT-PCR eingesetzt, bei der unter Verwendung der Fascin-spezifischen Primer hFascin-3 (5'-GGCAAGCCTGGCTGTAGTAG-3'; SEQ ID NO: 25) und hFascin-4 (5'-CCAGAGTGAGATGCATGTTGG-3'; SEQ ID NO: 26), ein 277 bp-Fragment des 3'-UTR der humanen Fascin-cDNA (SEQ ID NO: 27) amplifiziert wurde. PCR-Denaturierung (94°C), Annealing und Extensionsschritte (72°C) wurden jeweils für 1 min durchgeführt. Die Annealingtemperatur wurde von 66°C auf 65°C vermindert, jeweils für zwei Zyklen, und auf 64°C für die weiteren 30 Zyklen.

Hybridisierungsproben

Eine genomische Bank wurde durchmustert mit der Sonde mFascin-ORF, die die ersten Zweidrittel (+73 bp bis +984 bp) des offenen Leserasters (ORF) der Maus-Fascin-cDNA (SEQ ID NO: 29) umfaßt. Innerhalb dieser Region zeigen Maus- und Mensch-Fascin-cDNA eine Homologie von 91% auf der Nucleotidebene (GenBank Hinterlegungsnummer U03057 und L33726). Die Sonde mFascin-ORF wurde hergestellt unter Verwendung der Primer allfas-1 (5'-GCCACCATGACCGCCAACGG-3'; SEQ ID NO: 31) und allfas-2 (5'-TGTGTCGCGGTCGATCTCCA-3'; SEQ ID NO: 32). Maus-cDNA aus dendritischen Zellen wurde zur Verwendung als Matrize Hitze-denaturiert (98°C für 5 min), und schnell auf Eis gekühlt vor dem Einsatz in der PCR. Während der PCR blieben die Denaturierungs- (95°C für 1 min) und Extensionsschritte (72°C für 2 min) konstant. Die Annealingschritte wurden für jeweils 1 min ausgeführt, wobei die Temperatur von 70°C auf 69°C und weiter auf 68°C vermindert wurde, jeweils für zwei Zyklen, und auf 67°C für weitere 30 Zyklen. Die Sonde wurde während der PCR markiert mit

Digoxigenin (DIG)-11-dUTP unter Verwendung des PCR DIG Probe Synthesis Kits (Roche Molecular Biochemicals, Mannheim).

Genomische Restriktionsfragmente, die die weiter distal gelegenen Bereiche des Gens enthalten, wurden identifiziert durch die Probe hFascin-UTR, die den distalen Bereich der 3'-UTR des menschlichen Fascin-Gens (+2388 bis +2664 in SEQ ID NO: 27) abdeckt. Die cDNA von humanen dendritischen Zellen wurde verwendet als Matrize zur Amplifikation unter Verwendung der Primer hFascin-3 und hFascin-4 (siehe oben). Das PCR-Produkt wurde in *Hinc II*-restringierten pUC18 cloniert und sequenziert. Zur DIG-Markierungsreaktion wurde subcloniertes hFas-UTR als Matrize verwendet. Die Matrize wurde Hitze-denaturiert bei 98°C für 5 min und schnell abgekühlt.

Hybridisierungsbedingungen

Alle Reagenzien für Filterhybridisierungen mit DIG-markierten Sonden und Signaldetektion wie auch Nylonmembranen wurden erworben bei Roche Molecular Biochemicals und die Experimente wurden durchgeführt nach Ross (BioTechniques 26 (1999), 150-155). Kurz zusammengefaßt wurde DIG Easy Hyb sowohl als Prähybridisierungs- als auch als Hybridisierungslösung verwendet und enthielt 50 mg/ml gescherte und hitzedenaturierte genomische DNA aus Lachsspermien. Sowohl Prähybridisierung als auch Hybridisierung wurde bei 38°C in einem Wasserbad bei leichtem Schütteln durchgeführt. Die DIG-markierte Sonde wurde detektiert durch mit alkalischer Phosphatase-gekoppeltem anti-DIG-Antikörper und CDP-StarTM als Chemilumineszenzsubstrat. Die Expositionszeit der Autoradiographie betrug typischerweise weniger als 30 min für Filter angeordneter Genbanken und weniger als 5 min für Southern blots und Koloniefilter. Zur Wiederverwertung wurden die Filter "gestrippt", wie es vom Hersteller empfohlen wird.

Isolierung und Charakterisierung der genomischen Clone

Die angeordnete menschliche genomische Genbank, die aus PBMC eines Kaukasiers gewonnen wurde, wurde hergestellt von P. Ioannou, C. Chen und B.

Zhao (Roswell Park Cancer Institute, Human Genetics Department, Buffalo, NY) (Ioannou, Nature Genetics 6 (1994), 84-89) und wurde erhalten vom Ressourcenzentrum des deutschen Humangenomprojekts am Max-Planck-Institut für Molekulare Genetik (Berlin). Die Genbank wurde gescreent durch Hybridisierung mit der Sonde mFascin-ORF. Nach DNA-Präparation (NucleoBond™ Plasmid Kit, Clontech, Palo Alto, USA) wurden positive PAC-Clone durch Southern blot verifiziert. Zur weiteren Charakterisierung wurde die PAC-Clon-DNA gespalten und die entstehenden Fragmente nach dem Zufallsprinzip in pZero2.1 (Invitrogen, Groningen, Niederlande) subcloniert.

DNA-Sequenzierung

Die Nucleotidsequenz wurde bestimmt durch Cycle-Sequencing und Analyse auf einem PE 373A-Sequencer (Perkin-Elmer). Wegen des hohen GC-Gehalts der Fascingensequenzen, vor allem im 5'-flankierenden Bereich, der vermutlich zur Ausbildung starker Sekundärstrukturen führte, ließen sich Teilbereiche nicht nach herkömmlichen Methoden sequenzieren, was allerdings durch folgende Maßnahmen überwunden werden konnte. GC-reiche DNA-Matrizen wurden vor der Sequenzierung linearisiert durch Restriktionsspaltung und Hitzedenaturierung (98°C für 5 min). Desweiteren wurde DMSO in einer Endkonzentration von 5% hinzugefügt. Die Temperatur des Denaturierungsschritts wurde reduziert auf 95°C und die Zyklenzahl auf 30 erhöht.

Konstruktion eines Fascin-Promotor-Photinus-Luciferase-Plasmids

Ein subcloniertes genomisches *Hind III*-Restriktionsfragment von 5,5 kb Länge enthielt 3 kb 5'-flankierende Sequenz des Fascingens, Exon 1 und ein Teil des Introns 1 (Figur 1). Das putative Promotorfragment, das die 5'-flankierende Genregion und einen Teil des 5'-UTR umfaßt, wurde herausgeschnitten (Restriktionsspaltung mit *Acc I*, Auffüllen der Überhänge, und weitere Restriktionsspaltung mit *Kpn I*) und ligiert in den promotorlosen Photinus-Luciferase-Expressionsvektor pGL3-Basic (Promega, Madison, USA). Die korrekte Länge und Orientierung des Promotorkonstrukts (pFascin-3,0) wurde überprüft durch die DNA-

Sequenzierung von beiden Enden (RVprimer3 und GLprimer2, Promega). Zur 5'-Deletionsclonierung wurde das Konstrukt mit *Kpn* I und *Stu* I gespalten und die Deletionsclonierung durchgeführt nach den Anweisungen des Herstellers (Amersham Pharmacia Biotech). Einige Deletionsclone wurden erzeugt durch Restriktionsspaltung mit *Kpn* I und einem zweiten Restriktionsenzym, das in pFascin-3,0 schneidet. Die Integrität und Länge der Plasmid-DNA jedes Deletionsclons wurde durch Gelelektrophorese überprüft. Um die absoluten Photinus-Reporterexpressionswerte zu normalisieren auf Grund von Differenzen in der Transfektionseffektivität, wurde ein CMV-Promotor-kontrollierter Renilla-Luciferase-Expressionsvektor (pRL-CMV, Promega) verwendet als Coreporter. Plasmid-DNA wurde unter Verwendung des QiagenTM Plasmidkits isoliert (Qiagen, Hilden). Die DNA-Konzentration wurde photometrisch bestimmt. Die Integrität der Plasmid-DNA wurde durch Gelelektrophorese überprüft.

Um die Expressionsstärke des humanen Fascin-Promotors in verschiedenen Zelltypen vergleichen zu können, wurde zur Normierung der Promotor des Haushaltsgens EF1 α (Wakabayashi-Ito, J. Biol. Chem. 269 (1994), 29831-29837) verwendet. Dies sollte ausschließen, daß etwaige Zelltyp-spezifische Unterschiede in der Expression des allgemein verwendeten Referenzkonstrukts unter Kontrolle des CMV-Promotors das Ergebnis verfälschen. Hierzu wurde der EF1 α -Promotor aus dem Expressionskonstrukt PEF-BOS-lacZ (Mizushima, Nucleic Acids Res. 18 (1990), 5322) mittels PCR amplifiziert und in pGL3-Basic kloniert. Parallel dazu wurde der CMV-Promotor des Coreporterkonstrukts pRL-CMV durch den EF1 α -Promotor ersetzt (pRL-EF1 α). Transfektionsexperimente in HaCaT, SHSY-5Y und DC, bei denen beide Coreporter in Parallelansätzen getestet wurden, zeigten, daß die Expressionsstärken des EF1 α - und CMV-Promotors in der gleichen Größenordnung liegen.

Transiente Zelltransfektion

Die Transfektionen wurden durchgeführt nach der biolistischen Gentransfertechnik unter Verwendung des Helium-angetriebenen PDS-1000/He-Systems (Bio-Rad,

Hercules, USA) nach Kalkbrenner (Meth. Mol. Biol. 83 (1996), 203-216). Dendritische Zellen (5×10^5) und THP-1 Zellen (10^6) wurden cotransfiziert mit 4,5 pMol des Testkonstrukts und 0,5 pMol des pRL-CMV. Der Abstand zwischen dem Makrocarrierhalter und dem *Transwell* (24 mm Durchmesser, 3 μ m Porengröße, Corning Costar, Bodenheim) betrug 6 cm. Es wurde ein Druck von 900 psi angewendet, basierend auf Optimierungsversuchen.

Die humanen Zelllinien HaCaT und SHSY-5Y wurden mittels Lipofektion transfiziert. Am Vortag wurden jeweils 5×10^5 Zellen pro Loch einer 24-Loch-Zellkultur-Platte ausgesät. Für jede Transfektion wurden 190 nMol Testkonstrukt und 20 nMol Coreporter (pRL-CMV) eingesetzt. Durch Zugabe denaturierter Lachssperma-DNA (Roche Diagnostics, Mannheim) wurde die Gesamt-DNA-Menge auf 1 μ g normiert. Als Transfektionsreagenz wurde GenePorter (PEQLAB, Erlangen) eingesetzt. Aufgrund entsprechender Optimierungsexperimente wurden für die Transfektion der HaCaT-Zellen 7 μ l GenePorter / μ g DNA und für SHSY-5Y 4 μ l / μ g DNA eingesetzt. Die Transfektion wurde entsprechend den Empfehlungen des Herstellers durchgeführt.

Luciferase-Assays

Es wurden Zellextrakte 24 Stunden nach der Transfektion präpariert. Die Zellen wurden pelletiert und gewaschen mit 2 ml PBS. Die Zellpellets wurden in Passive Lysis Buffer (Promega) resuspendiert in einer Konzentration von 10^6 Zellen/ 100 μ l und für 15 min bei Raumtemperatur inkubiert auf einem Taumelschüttler unter leichtem Schütteln. Die Zellysate wurden bei -20°C aufbewahrt. Die Proben wurden aufgetaut und auf Eis gestellt. 10 μ l des Zellysats wurden auf Photinus- und Renilla-Luciferaseaktivität hin analysiert mit dem Dual-LuciferaseTM Reporterassaysystem nach Angaben des Herstellers (Promega) in einem Turner-Luminometer TD-20/20 (Turner Design, Sunnyvale, USA). Die Luciferaseaktivität wurde für eine Dauer von 10 Sekunden gemessen. Die absoluten Werte wurden normalisiert durch Dividieren der Photinus-Luciferaseaktivität durch die Renilla-Luciferaseaktivität.

Beispiel 1: Isolierung und genomische Organisation des menschlichen Fascingens

Durch das Screening einer humanen, PBMC-abgeleiteten genomischen PAC-Bibliothek mit der Maus-Fascin-cDNA-spezifischen Sonde mFascin-ORF wurden 16 Clone identifiziert, wobei sich 8 Clone davon als positiv im Southern Blot herausstellten. Clon RPCIP704C24766Q3/4 wurde zur Charakterisierung des menschlichen Fascin-Genlokus verwendet. Die PAC-Clon-DNA wurde jeweils mit *Hind III*, *Pst I* und *Sac I* restringiert und die Restriktionsfragmente nach dem Zufallsprinzip in den Vektor pZero2.1 cloniert. Ein größeres Restriktionsfragment, das einen Teil des Exons 1, das komplette Intron 1 und Exons 2 bis 4 enthält, wurde durch Doppelverdau mit *Hinc II* und *Eco RI* cloniert. Clonierte Fragmente, die mit Fascin-Sonden hybridisierten, wurden weiteranalysiert und der Sequenzierung unterzogen. Figur 1 zeigt, daß das menschliche Fascingen eine genomische Region von annähernd 13 kb abdeckt und aus 5 Exons besteht. Exon 1 umfaßt die kurze 5'-nicht-translatierte Region (5'-UTR) des Gens und etwa die Hälfte der translatierten Sequenz. Intron 1 hat eine Länge von etwa 8 kb. Die kurzen Exons 2 bis 4 sind innerhalb einer Region von 800 bp gruppiert und codieren ein Drittel der translatierten mRNA. Exon 5 liegt etwa 1 kb weiter stromabwärts und enthält den verbleibenden codierenden Bereich sowie die komplette 3'-UTR des Gens. Die DNA-Sequenz des Gens ist in Figur 2 zu sehen. Die in der vorläufigen DNA-Sequenzierung ermittelten Gensequenzen wiesen zunächst noch einzelne Lücken in Intron 1 sowie eine Lücke in Exon 5 auf (SEQ ID NOs: 33, 10-15, 34 und 35), die mit der Bestimmung der Gesamtsequenz des Fascingens (SEQ ID NO:72) geschlossen wurden. Die letzten fünf Basenpaare der 3'-UTR der cDNA-Sequenz sind nicht in der genomischen Sequenz enthalten. Die Exon/Intron Grenzen entsprechen der GT/AG-Regel, abgesehen von einer Abweichung der 5'-Introngrenze des Introns 4 (GC anstatt GT).

Beispiel 2: Funktionelle Analyse der 5'-flankierenden Region

Wie es typisch ist für einen eukaryontischen Genpromotor, wurde ein Konsensus-TATA-Box-Motif (TATAAAA) in der Nähe des Transkriptionsstartpunkts identifiziert.

Um die Promotoraktivität der 5'-flankierenden Region des humanen Fascingens einzuschätzen, wurde ein Promotor-Reporter-Konstrukt (pFascin-3.0, Position 1 bis 3069 in SEQ ID NO:72 bzw. SEQ ID NO: 1) konstruiert, das die Gesamtlänge der isolierten 5'-flankierenden Region enthält sowie stromabwärts angrenzend die ersten 102 bp der 5'-nicht-translatierten Sequenz des Fascingens (Position 1 des publizierten humanen cDNA-Clons als Transkriptionsstartpunkt definiert). Ausgereifte dendritische Zellen, die Fascin stark exprimieren, wurden transient transfiziert durch einen biolistischen Gentransfer dieses Konstrukts. Während die negative Kontrolle (pGL3-Basic) nur Hintergrundluciferaseaktivität zeigt, resultierte die Transfektion mit pFascin-3.0 in einer starken Reporterexpression, die im Vergleich zu pGL3-Basic um das 101fache erhöht war (Fig. 3). Um cis-agierende Elemente in dem Promotor zu identifizieren, wurden 5'-Deletionsclone von pFascin-3.0 erzeugt (Figur 4, Position 1136 bis 3069, 1451 bis 3069, 1621 bis 3069, 1830 bis 3069, 2127 bis 3069, 2410 bis 3069, 2700 bis 3069, 2859 bis 3069 und 2915 bis 3069 in SEQ ID NO:72 bzw. SEQ ID NO: 2 bis 8, 21, 22) und in transienten Transfektionen eingesetzt. Die menschliche Monocytenzelllinie THP-1 wurde als Negativkontrolle verwendet, da eine RT-PCR keine endogene Fascinexpression anzeigte (Fig. 5).

Ein 5'-Deletionskonstrukt, das vom Promotor nur 53 Basenpaare einschließlich der Consensus-TATA-Box enthält (pFascin-0.05, Position 2915 bis 3069 in SEQ ID NO:72 bzw. SEQ ID NO: 22) ist nicht ausreichend für eine Luciferaseexpression (Fig. 6). Das erste Fragment, das eine basale Promotoraktivität zeigt, ist ein 211 bp langes Fragment, das 109 bp des Promotors enthält (pFascin-0.11, Position 2859 bis 3069 in SEQ ID NO:72 bzw. SEQ ID NO: 21). Die Transfektion mit Promotorkonstrukten, die in den weiter distal gelegenen Bereich der 5'-flankierenden Sequenz reichen, resultierte in einer weiteren schrittweisen Erhöhung der Luciferaseexpression. Die höchste Reporterexpression wurde für pFascin-1.6 (Position 1451 bis 3069 in SEQ ID NO:72 bzw. SEQ ID NO: 3) detektiert, das eine 3,4-fache Aktivität verglichen mit dem Minimalpromotor (pFascin-0.11) aufwies.

Terminal ausgereifte dendritische Zellen zeichnen sich gegenüber unreifen dendritischen Zellen durch die Oberflächenexpression des Markers CD83 aus. Die Zellen einer DC-Kultur wurden in eine CD83-positive und eine CD83-negative Fraktion aufgeteilt. Während in der CD83-positiven DC-Fraktion die

Reportergenexpression für die längeren Promotorkonstrukte pFascin-1.6 und pFascin-3.0 den fünffachen Wert des Basalpromotores pFascin-0.11 erreicht, wird in der CD83-negativen Zellfraktion durch Verwendung von pFascin-1.6 und pFascin-3.0 das Expressions-Niveau nicht über die Basalexpression hinaus gesteigert (Figur 7). Dieser Befund belegt, daß der Fascinpromotor vorwiegend in terminal ausgereiften, CD83-positiven dendritischen Zellen aktiv ist. Dies ist für den klinischen Einsatz besonders vorteilhaft.

Überraschenderweise weist der Promotor des menschlichen Fascingens auch eine basale Aktivität in THP-1 auf. Daraus geht hervor, daß zumindest ein unspezifisches aktivierendes Transkriptionselement in nächster Nachbarschaft 5' der TATA-Box lokalisiert ist. Jedoch nimmt im Gegensatz zu den dendritischen Zellen die Promotoraktivität in den THP-1-Zellen mit zunehmender Länge der transfizierten Konstrukte ab. Bei dem Konstrukt mit der vollen Promotorlänge ist die Reportergenexpression auf ein Drittel der Aktivität, die für das Konstrukt pFascin-0.11 detektiert wurde, reduziert. Zusammenfassend läßt sich sagen, daß innerhalb der untersuchten 5'-flankierenden Region des menschlichen Fascingens transkriptionsregulierende Elemente mit Spezifität für dendritische Zellen kooperieren, um die Expression des menschlichen Fascingens zu erhöhen. Zusätzlich enthält der Fascinpromotor Elemente, die die Transkription in Fascin-negativen Zelllinien wie THP-1 spezifisch reprimieren.

Auch ein zelltypübergreifender Vergleich der Fascinpromotorstärke - im Verhältnis zu einer Positivkontrolle, die zelltypunabhängig starke Reportergenexpression vermittelt - bestätigte die DC-Spezifität des Fascinpromoters. In der Fascin-negativen Keratinozytenlinie HaCaT bewegte sich die Fascinpromotor-abhängige Reportergenaktivität mit unter zehn Prozent der Positivkontrolle lediglich auf basalem Niveau (Figur 8). Selbst in der Fascin-positiven Neuroblastomlinie SHSY-5Y erreichte die vom Fascinpromotor getriebene Reportergenexpression nur ein Fünftel der Positivkontrolle. Demgegenüber wies der Facinpromotor in den stark Fascin-positiven DC die ca. anderthalbfache Stärke der Positivkontrolle auf.

PATENTANSPRÜCHE

1. Regulatorische Sequenz ausgewählt aus der Gruppe bestehend aus:
 - (a) regulatorischen Sequenzen, die die unter SEQ ID NO:72 von Position 1 bis 3069 oder die unter SEQ ID NO:1 angegebene Nucleotidsequenz umfassen;
 - (b) regulatorischen Sequenzen, die die Nucleotidsequenz umfassen, die in der Insertion des Klons DSM13274 enthalten ist und durch Amplifikation unter Verwendung zweier Oligonucleotide erhältlich ist, die die unter SEQ ID NOS: 36 und 37 angegebenen Sequenzen besitzen;
 - (c) regulatorischen Sequenzen, die eine Nucleotidsequenz umfassen aus SEQ ID NO:72 von Position 1136 bis 3069, 1451 bis 3069, 1621 bis 3069, 1830 bis 3069, 2127 bis 3069, 2410 bis 3069 oder 2700 bis 3069 oder ausgewählt aus der Gruppe bestehend aus: SEQ ID NOs:2 bis 8;
 - (d) regulatorischen Sequenzen, die eine Nucleotidsequenz umfassen, die in der Insertion des Klons DSM 13274 enthalten ist und durch Amplifikation unter Verwendung eines Paares von Oligonucleotiden erhältlich ist, wobei die Sequenzen der Oligonucleotide angegeben sind unter den SEQ ID-Nummern, ausgewählt aus der Gruppe von Paaren bestehend aus: 38 und 37; 39 und 37; 40 und 37; 41 und 37; 42 und 37; 43 und 37; und 44 und 37;
 - (e) regulatorischen Sequenzen, die zumindest einen funktionalen Teil einer unter (a) bis (d) genannten Sequenz umfassen und eine für dendritische Zellen spezifische Expression bewirken; und
 - (f) regulatorischen Sequenzen, die eine Nucleotidsequenz umfassen, die mit einer unter (a) bis (e) angegebenen regulatorischen Sequenz hybridisiert, und eine für dendritische Zellen spezifische Expression bewirken.
2. Regulatorische Sequenz nach Anspruch 1, die kombiniert ist mit mindestens einer der Nucleotidsequenzen,

- a) ausgewählt aus der Gruppe bestehend aus: den Abschnitten von SEQ ID NO:72 von Position 3911 bis 13398, 13556 bis 13637, 13760 bis 14004, 14173 bis 15414 und 16791 bis 16951 und SEQ ID NOs:9 bis 20, oder Teilen davon; oder
 - b) die in der Insertion des Klons DSM13274 enthalten ist und durch Amplifikation unter Verwendung eines Paares von Oligonucleotiden erhältlich ist, wobei die Sequenzen der Oligonucleotide angegeben sind unter den SEQ ID-Nummern, ausgewählt aus der Gruppe von Paaren bestehend aus: 45 und 46; 47 und 48; 49 und 50; 51 und 52; 53 und 54; 55 und 56; 57 und 58; 59 und 60; 61 und 62; 63 und 64; 65 und 66; 67 und 68; und 45 und 60.
3. Regulatorische Sequenz nach Anspruch 1 oder 2, die aus Mensch stammt.
4. Rekombinantes Nucleinsäuremolekül enthaltend die regulatorische Sequenz nach einem der Ansprüche 1 bis 3.
5. Vektor enthaltend die regulatorische Sequenz nach einem der Ansprüche 1 bis 3 oder das rekombinante Nucleinsäuremolekül nach Anspruch 4.
6. Rekombinantes Nucleinsäuremolekül nach Anspruch 4 oder Vektor nach Anspruch 5, der zusätzlich eine zu exprimierende Nucleotidsequenz enthält, wobei die Expression der Nucleotidsequenz von der regulatorischen Sequenz kontrolliert wird.
7. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, wobei die Nucleotidsequenz ein Antigen codiert.
8. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 7, wobei das Antigen tumor- oder pathogenspezifisch ist oder an der Bildung von Creutzfeldt-Jakob-Plaques oder Alzheimer-Plaques beteiligt ist.

9. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 7, wobei das Antigen ein Autoantigen oder ein Transplantationsantigen ist.
10. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 7, wobei das Antigen ein Allergen ist.
11. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, wobei die Nucleotidsequenz ein Protein codiert, das eine Immunantwort reguliert.
12. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 11, wobei das Protein ein Cytokin oder ein co-stimulatorisches Molekül ist.
13. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 11, wobei die Regulation eine Hemmung ist.
14. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 11, 12 oder 13, wobei das Protein IL-10 oder TGF- β ist.
15. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 11 oder 12, wobei die Regulation eine Steigerung der Immunantwort ist.
16. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 11, 12 oder 15, wobei das Protein IL-2, IL-4, IL-12, IL-15, IL-18, IFN-gamma, IFN-alpha, DC-CK1, MDC oder GM-CSF ist.
17. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 11, 12 oder 15, wobei das Protein ein Mitglied der B7-Familie, ICOS-Ligand oder CD40 ist.
18. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, wobei die Nucleotidsequenz für ein Apoptose-induzierendes Molekül codiert.

19. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 18, wobei das Apoptose-induzierende Molekül zur TNF-Superfamilie gehört.
20. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, wobei die Nucleotidsequenz eine Antisense-Sequenz ist oder ein Ribozym exprimiert.
21. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 20, wobei die Antisense-Sequenz oder das Ribozym spezifisch ist für eine mRNA, die ein Cytokin oder ein co-stimulatorisches Molekül codiert.
22. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, wobei die Nucleotidsequenz einen Transkriptionsfaktor codiert.
23. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, der zusätzlich eine zweite zu exprimierende Nucleotidsequenz enthält, wobei eine Nucleotidsequenz ein Antigen codiert und die zweite Nucleotidsequenz ein Protein codiert, das eine Immunantwort reguliert.
24. Vektor nach einem der Ansprüche 5 bis 23, der ein Virus ist.
25. Vektor nach einem der Ansprüche 6 bis 24, der zur Gentherapie oder DNA-Vakzinierung geeignet ist.
26. Rekombinantes Nucleinsäuremolekül oder Vektor nach Anspruch 6, wobei die Nucleotidsequenz ein Reportergen ist.
27. Verfahren zur Herstellung von genetisch modifizierten Wirtszellen, dadurch gekennzeichnet, daß man Wirtszellen mit einem Vektor nach einem der Ansprüche 5 bis 26 transfiziert und die transfizierten Wirtszellen in einem Kulturmedium kultiviert.
28. Wirtszelle, die genetisch modifiziert ist mit der regulatorischen Sequenz nach einem der Ansprüche 1 bis 3, dem rekombinanten Nucleinsäuremolekül nach

einem der Ansprüche 4, 6 bis 23 und 26 oder dem Vektor nach einem der Ansprüche 5 bis 26 oder erhältlich nach dem Verfahren nach Anspruch 27.

29. Wirtszelle nach Anspruch 28, die eine dendritische Zelle ist.
30. Wirtszelle nach Anspruch 28 oder 29, die aus Mensch stammt.
31. Nucleotidsequenz umfassend ein Fragment mit einer Länge von mindestens 15 Nucleotiden, das spezifisch mit einem Strang einer regulatorischen Sequenz nach einem der Ansprüche 1 bis 3 unter stringenten Bedingungen hybridisiert.
32. Verfahren zur Antigen-spezifischen Stimulierung von T-Zellen in vitro, enthaltend die Schritte:
 - (a) Transfektion von dendritischen Zellen mit einem Vektor nach einem der Ansprüche 7 bis 10, allein oder in Kombination mit einem Vektor nach einem der Ansprüche 11, 12, 15, 16, 17, 20 und 22, oder mit einem Vektor nach Anspruch 23;
 - (b) Cokultivierung der durch Schritt (a) gewonnenen transfizierten dendritischen Zellen mit T-Zellen; und
 - (c) Detektion der Aktivierung der T-Zellen aus Schritt (b).
33. Verfahren zur Herstellung eines Arzneimittels umfassend die Schritte (a) bis (c) nach Anspruch 32 und zusätzlich den Schritt
 - (d) Formulierung eines Arzneimittels durch Versetzen der durch den Schritt (c) gewonnenen stimulierten T-Zellen mit einem pharmazeutisch verträglichen Träger.
34. Verfahren zur Herstellung von T-Zell-stimulierenden dendritischen Zellen in vitro, enthaltend die Schritte:
 - (a) Transfektion von dendritischen Zellen mit einem Vektor nach einem der Ansprüche 7 bis 10, allein oder in Kombination mit einem Vektor nach

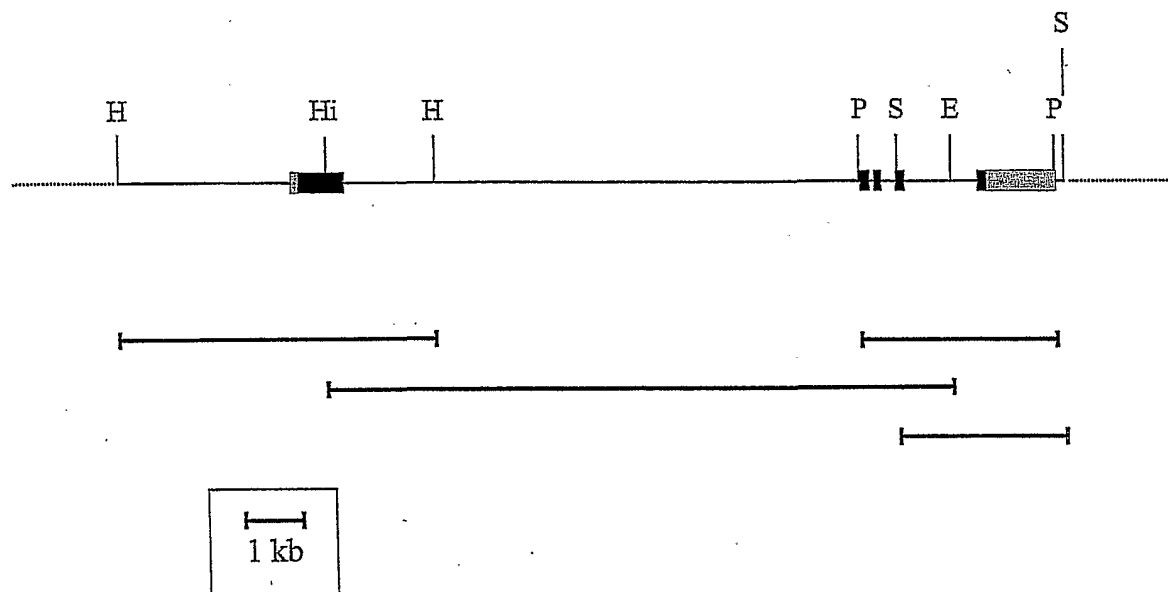
- einem der Ansprüche 11 bis 22 oder mit einem Vektor nach Anspruch 23; und
- (b) Kultivierung der transfizierten dendritischen Zellen in einem geeigneten Medium und/oder Detektion der T-Zell-stimulierenden Aktivität.
35. Verfahren zur Herstellung eines Arzneimittels umfassend die Schritte (a) und (b) nach Anspruch 34 und zusätzlich den Schritt
- (c) Formulierung eines Arzneimittels durch Versetzen der durch den Schritt (b) gewonnenen T-Zell-stimulierenden dendritischen Zellen mit einem pharmazeutisch verträglichen Träger.
36. Verfahren nach einem der Ansprüche 32 bis 35, wobei die dendritischen Zellen aus Mensch stammen.
37. Arzneimittel umfassend das rekombinante Nucleinsäuremolekül nach einem der Ansprüche 6 bis 23, den Vektor nach einem der Ansprüche 6 bis 25, die Wirtszelle nach einem der Ansprüche 28 bis 30, Antigen-spezifisch stimulierte T-Zellen, erhältlich nach dem Verfahren nach Anspruch 32, oder T-Zell-stimulierende dendritische Zellen, erhältlich nach dem Verfahren nach Anspruch 34, und optional einen pharmazeutisch verträglichen Träger.
38. Arzneimittel nach Anspruch 37, das ein Impfstoff ist.
39. Verwendung des rekombinanten Nucleinsäuremoleküls nach einem der Ansprüche 6 bis 8 und 10, des Vektors nach einem der Ansprüche 6 bis 8, 10, 24 und 25, allein oder in Kombination mit einem rekombinanten Nucleinsäuremolekül oder Vektor nach einem der Ansprüche 11, 12, 14 bis 17 und 22, der Wirtszelle nach einem der Ansprüche 28 bis 30, des rekombinanten Nucleinsäuremoleküls oder Vektors nach Anspruch 23, von Antigen-spezifisch stimulierten T-Zellen, erhältlich nach dem Verfahren nach Anspruch 32, oder T-Zell-stimulierenden dendritischen Zellen, erhältlich nach dem Verfahren nach Anspruch 34, zur Herstellung eines Arzneimittels zur Impfung gegen Viren, Bakterien, Pilze, Parasiten, Tumoren, Allergene,

Creutzfeldt-Jakob-Plaques, oder Alzheimer-Plaques oder zur gentherapeutischen Behandlung von Tumoren oder viralen, bakteriellen oder parasitischen Infektionen oder von Allergien.

40. Verwendung des rekombinanten Nucleinsäuremoleküls nach einem der Ansprüche 6, 7, 9 und 10 oder des Vektors nach einem der Ansprüche 6, 7, 9, 10, 24 und 25, allein oder in Kombination mit einem rekombinanten Nucleinsäuremolekül oder Vektor, nach einem der Ansprüche 11, 13, 14, 18 bis 22, des Nucleinsäuremoleküls oder Vektors nach Anspruch 23, oder der Wirtszelle nach einem der Ansprüche 28 bis 30 zur Herstellung eines Arzneimittels zur Behandlung von Autoimmunerkrankungen, Transplantatabstoßung oder Allergien.
41. Verwendung des rekombinanten Nucleinsäuremoleküls nach Anspruch 18 oder 19 oder des Vektors nach einem der Ansprüche 18, 19, 24 und 25 zur Herstellung eines Arzneimittels zur Vermeidung der Abstoßung von Transplantaten und von Autoimmunreaktionen.
42. Verwendung der regulatorischen Sequenz nach einem der Ansprüche 1 bis 3, des rekombinanten Nucleinsäuremoleküls nach einem der Ansprüche 4, 6 bis 23 oder des Vektors nach einem der Ansprüche 5 bis 25 zur spezifischen Expression von Antigenen oder immunregulatorischen Proteinen in dendritischen Zellen.
43. Verwendung der regulatorischen Sequenz nach einem der Ansprüche 1 bis 3, des rekombinanten Nucleinsäuremoleküls nach Anspruch 4, 6 oder 26 oder des Vektors nach Anspruch 5, 6 oder 26 zur Identifizierung und Isolierung von cis-Elementen aus der regulatorischen Sequenz, die eine für dendritische Zellen spezifische Expression vermitteln.
44. Verwendung der regulatorischen Sequenz nach einem der Ansprüche 1 bis 3, des rekombinanten Nucleinsäuremoleküls nach Anspruch 4, 6 oder 26 oder

des Vektors nach Anspruch 5, 6 oder 26 zur Bestimmung des Reifegrads von dendritischen Zellen.

45. Verwendung der regulatorischen Sequenz nach einem der Ansprüche 1 bis 3, des rekombinanten Nucleinsäuremoleküls nach Anspruch 4, 6 oder 26 oder des Vektors nach Anspruch 5, 6 oder 26 zur Identifizierung und Isolierung von Faktoren, die eine Expression spezifisch für dendritische Zellen vermitteln.
46. Verwendung der regulatorischen Sequenz nach Anspruch 1 oder Teilen davon zur Blockierung von Transkriptionsfaktoren durch Bereitstellung von Transkriptionsfaktor-Bindungsstellen in dendritischen Zellen.

**Figur 1**

1 AAGCTTCTAG AAGGGGTATA TTTTGCCCCC AGGCCCCAAA TCCTGGTCTT TGGACTACAG
 61 TCAAGACCTA GCACAGGGCT CAGGCCTAAA AAAAAGCTGT TGAGGAGGGT GTGAGGCTGC
 121 AACGTTGCGG TGAGAAGGGG GTCCCCAGGG AGGGCGCAGC AGGAAGCCCC AGGGAAGTGC
 181 CTGGGAGAGG GAGGTTGTGC AGCCAGAAGG AGCAGCCCGC AGCCTTTGGT TGTGACTCC
 241 TGCTTTGTAA GTGGCAGTCT GGTGGGTAG GACAGGGTCC GAGTCCTCAC TCAGGGAAGT
 301 GAGTCCAGGG TGAGCTGAGC AGCCCTTCCT GGTGGCCTCA CTTTCCCCTG CAGAGCCTGC
 361 TGCATGGTGT CCAGTGGCCA GCCTGGGAGG AGCTCAGGAC TGGCCTCACC TCGTGCCACG
 421 CCCTCGTACC AAGGTGGCTC AGATGGCACC TGGGTTCCTG AAGGGCCCAG GAACACAGCG
 481 TCCATGTCCC GATCCTTCCC CGGAGGGACT TGGGGTGGGG CCTGCAGGAA GGATCATGTG
 541 ACTTGTTCAA GGGCAGCTTG TCCTGCCCCG GACACAGAGG TGCCCATCGT AAGCGAGATG
 601 CAGGCAAGGT GAGGACAGGG CATGGTGCCG AGCAGGAATG ATTTTCGAAA ATGCTACTCT
 661 GTGGTCGGGC ACGGTGGCTC ACTCCTGTAA TCCCAGTACT TTGGGAGGCC CAGGCGGGCA
 721 GATCATGAGG TCAGGAGTTC AAGACCAGCT TGGCCAACAT GGTGAAACCC CTCTCTACT
 781 GAAAATACAA AAAACTACCC GGGCGTGCTG GTGGATGCCT GTAATCCCAG CTACTCAGGA
 841 GGCTGAGGCA GGAGAATCGC TTGAACCCCTG GAGGTGAAGG TTGCAATGAG CCGAGATTGC
 901 ACCACTGCAC TCCGACCTGG GCGACAGAGA AAGACTCCGT CTGAAAAACA AACAAACAAA
 961 ACCCCGAGAT TCTCTTTTCC CCCGTCCGGA GCTCTATGGC CATCTGGAGC TCGTCTGTCT
 1021 CACAAGGACA CATTTCTCTG CAGCAGCTGT GGACCAGGGC TCACTCACTC ACATCTGTAC
 1081 CCAATCTAGA GCAGGACAAA CACCTATCAC CTGCACTGGC AATGGACAGA GGACAGTGGC
 1141 AGCCCCATCA CCAGAGGCAT TCAAGCCAAG GCATTTTCTA TCGTCTATTT ATCTATCTAT
 1201 CTATCTATCT ATCTATCTAT CTATCTATCT ATCTATCTAG ATGAGATCTT GCTATGTTGC
 1261 CCAGGTTGAA CTCCTGGCCT CAAGCGATCC TCCTGCTTCA GCCTCCCAA GTGTTGGGAT
 1321 TTCAGGTGTG AACCCTGTG CCTGGCTGGA GTGCAGTGGC ACTGTCATAG CTCACTGCAG
 1381 CCTCCACCTC CTGGGCTCAA GGAATCCTCT TTCCTCAGCC TCCTGAGTAG CTGGGACCAC
 1441 AGGCATGCAC CATCACACCC AGCTAAACCT TGATTTTTTC CATAGTATAT GGCTCAGGGC
 1501 AGAGCAAAGA AGCACAAAAA CATGTTGGCT TGCGGTTGAG GGCTGATGGG AGGGGGTTCT
 1561 GGCCAAAGCC CAAAATTAAC AGCCACAACG TCAGTGTCTG TGGGAGGGGT TGCCAGGGGC
 1621 GTGCGGGTTC TGGGGCTCAA GGCCCTGCCG GTAAACCCAT TTGAAGCAGG ATGGCAAGAA
 1681 GGTGACACCA TCTTCCCCCG CGCACTGAAA GCCCCTGGCT GGGATGCTCT GGGCAGACA
 1741 CAGGCTCGGA CCAGCCCCAG CAATCCCAGT TTATCAGCGA GCCGGCTGAG GGCCCGGAGT
 1801 TATCTCAGTG CCCGGCCTGA GACCTTGTGG GCAGCCTGTG GTCATGCTGG CATTCCAGGG
 1861 GCGTTTGTGC ATGTGGGGAA TGTCCAGGAA AAGCCTCAGC CTTCCGTGAG GCGCAGAAAA
 1921 GGGAAGTGTC CCTAGAGGGG GTGGGTGAGG GCGTGGGAGG TGGTGTCTGC AGGGAATGTC
 1981 CCCTTTGGGG GAGGAGGATG GAGGGTTGGG ATTCTGAGGA TGGGGGGGGG GGCTGTAGCC
 2041 AGCACCATGT CCCTCCTGTG TGACCAGCTC AGAGTCCCAT GAAATTGGGG CTTGGGAGGG
 2101 GAAGGGACAC TGGCCTGGGA ACCAGAGACC TGGGCTGGTC TGCTCACAG TTGCTGGACC
 2161 TCTGTGATCC GTGTCAAAAA ACGAGAACAC CAATTCTGT CCTGCCACT CACCACCAGG
 2221 TGGGACCTGA ACCCTGGCAT CGCCAGCATT GGGAAATGTCG GCCACTGACT CAACCACTCT
 2281 CCCGGAGACC TATTTGGGCC ACCCGAGGCG GGTGCCTGGG CCACACGGAG GGGTCCTGGT
 2341 GGTCTTCAGG GCAGCGGCTG TGGGGCTGAA GCCTCAAGGA ACCACATCTC TGCATAGGAG
 2401 GGCCAGGCTG CAGGGCCTCG GGAGACAACC TAGTTGGGCG TGTGGGGTCT TGTGGGTCCC
 2461 AGGTCCTGGC CTCACCGGGT CCCCACCGCG CTGTCAGCTC CCAGCCTCTT TCCCTCGTCT
 2521 GCCTCTGGGC TTCTGTAAAG CTATGTGCTC CAAGGCCACC TCCTCCAGG ACCCTCAGAG
 2581 CCCCCACCTT CGCCAGTAC CGATCTGCAC CGTGGTCTCT GAAGTCTCCA TCGTGACCTC
 2641 AAACCTCGCT CGTCCTTGCT CCTAGCGAGG CTTGGGGTCTG GGGTGTCCGA GGTGGGGGAC
 2701 ATCCGGGGGG GTTAGGTGGC TGGCGCGGGG AGCCGGGGTT GTGAGGGGTG ATGTCCTCAG
 2761 GCGGCGGGCG TGGGGGTGCG GGCGAGGACA CCGGTGGGGT GAGAGCACCG GCGGGGACG
 2821 AGCGGGGGCC GCAGCGCCGG GTCCCTCGGC CCGGGGCCCC TCCCGCGCGG AGCCAGGGGC
 2881 GGGACAGGGG GCGGTGGCCT GGTGGCGCTG ACGTACCTC GCCTATAAAA TGTCCGGGGC
 2941 GCGCTAGCT GGGCTTTGTG GAGCGCTGCG GAGGGTGCCT GCGGGCCGCG GCAGCCGAAC
 3001 AAAGGAGCAG GGGCGCCGCC GCAGGGACCC GCCACCCACC TCCCGGGGCC GCGCAGCGGC
 3061 CTCTCGTCTA CTGCCACCAT GACCGCCAAC GGCACAGCCG AGGCGGTGCA GATCCAGTTC
 3121 GGCCTCATCA ACTGCGGCAA CAAGTACCTG ACGGCCGAGG CGTTCCGGGT CAAGGTGAAC
 3181 GCGTCCGCCA GCAGCCTGAA GAAGAAGCAG ATCTGGACGC TGGAGCAGCC CCCTGACGAG
 3241 GCGGGCAGCG CCGCCGTGTG CCGTGGCAGC CACCTGGGCC GCTACCTGGC GCGGACAAG
 3301 GACGGCAACG TGACCTGCGA GCGCGAGGTG CCCGGTCCCG ACTGCCGTTT CCTCATCGTG
 3361 GCGCACGACG ACGGTGCTGT GTGCTGCGAG TCCGAGGCGC ACCGGCGCTA CTTGGCGGCG
 3421 ACCGAGGACC GCCTGTCTCT CTTGCGCGAG ACGGGTCCCC CGCCGAGAAG TGGAGCGTGC
 3481 ACATCGCCAT GCACCTCAG GTCAACATCT ACAGCGTCAC CCGTAAGCGC TACGCGCACC
 3541 TGAGCGCGCG GCGGGCCGAC GAGATCGCGG TGGACCGCGA CGTGCCCTGG GCGGTGCACT
 3601 CGCTCATCAC CCTCGCCTTC CAGGACCAGC GCTACAGCGT GCAGACCGCC GACCACCGCT

3661 TCCTGCGCCA CGACGGGCGC CTGGTGGCGC GCCCCGAGCC GGCCACTGGC TACACGCTGG
3721 AGTTCCGCTC CGGCAAGGTG GCCTTCCGCG ACTGCGAGGG CCGTTACCTG GCGCCGTCGG
3781 GGCCCAAGCG CACGCTCAAG GCGGGCAAGG CCACCAAGGT GGGCAAGGAC GAGCTCTTTG
3841 CTCTGGAGCA GAGCTGCGCC CAGGTGCTGC TGCAGGCGGC CAACGAGAGG AACGTGTCCA
3901 CGCGCCAGGG TGAGTGGGGA CGCTGCCCCC GCCTCTCCTG GTCCGTGCAC AAAGCGCACC
3961 CCACCCGCGC CCCTCCAGCC TCCCGCCCTT TCTCGCTCGC GCGCCGCTG CCGTCCGGAG
4021 CACTGCCCAT TCGCGCCCCG CTAGGCACGC GGGCTACCCC GCCTGGAGGG GCGGAGGAGT
4081 GGGGCTTTGC CCATCCGCGG [50] GTGTCA GCCCTTCCCC CCAGCCCCCTC CTCCGGCGTC
4191 TGCCCCGCGC TCGAGGCGCG GGCCTTTGTG AGCAGGGGGG CGGGTGCGCT CGACTGGGTC
4251 CTCCCTCGCC CCCTTTGTCC TGCTCCATCT CTGCAGATGG GAAAACCAGA TCGGCGGGG
4311 CGGGGGAGGG GATCGGCTTT TGCGGTTTCA CCCTGCAGAG GAGCCCCCGG CCGCGCCCCC
4371 GGGCCAGGG CTCGGGTCCC GAGGTTCCCC AGGAGGCGGT CTCCCTCCTC GTCGCGCGGG
4431 CCCGGGAACG GCGTGGCGCG CTGAGCTCCG CTCCGCGGGC GACCGAGGGC GGCTTCAGCG
4491 CGGTCTCCA GGCCGGTGGC CTGAGCTCCG CTCCGCGGGC GACCGAGGGC GGCTTCAGCG
4551 CGAGCCGGGA GACCCAGGC CAGCTCCCTC GGGGAAGCCCC TACCTCTGG TGAACCCATC
4611 CCCTAGGGCG CTCGCCGGGA ACAATTGACT GCAACTCGAT CCGTCCCCCTC CGGGGCCTCC
4671 TGCAGGCACG GCTATTTGCG ACCGGCCTGT GCGCCACTCT TTCCCTTCTT TCTCAGATTC
4731 TCCCCGAGGG GCTTTCTCTT CCTTTTGGCG TCCTCTTCCA CTCCTGGCGG AGAAGCTGTT
4791 CCTTGCTTGC AGACAGAGGC CCCGCTGGAG GGAGGGGCGT GGGGGGATCT TTTTCCCTGG
4851 ATAGGAAGTC GGACCCAGG CCTTGGAGAT GGCTGGACGG GAGCCCACTT TGTGTCTCA
4911 TCCCTCCTCT TGGAAGGAGC CCTCCTGACG GCCCCCTCT TGGCTGTGGG CTCTGCAGGA
4971 GGGGGAAAGA CCCCCATGG CAGTGGGGGT GAGGTGGAG GCCTGGGTGG GGTAAATGGCC
5031 ATTTTGTGCC TGAAGTTTGC TACCTTAAGT CCCAGAGAGG TGAAGCTGCT TGTATGGGT
5091 CACACAGCAA GTGACCACAA GGAAGCAGCA TGCAGAAGTG GGTGCATTTG GCTCCAGAAA
5151 CCCCCTTTCC TGTTTCCAC AGCGCATGGC GGGCAGAGCT CCTCACCTGA TCAGGAGCA
5211 TTGAGCCCTA CTAAGGGGAC CTGCTGAGCC ATGAGTGAAC CAGTGGGGTT CACGTTCTCT
5271 TGGGGACATT ACGAAGTGTC AAGGAAAAGG CAGCGGGTAC AGGTTATTAG ATGGCTGAGT
5331 CCTTTCTGAG CAGGACATTT GAGCTGGCCT GAAGGATGAA TAGAAGGTGG CAGGTATGGG
5391 GGGGGTGCTG TGTGGACCAG CGTTGTAGGC TGTGGGTGTC TCCTGTGCAA AGGTCTTATG
5451 GCAGGAAGGT ACAGAGAAAAG CCCAGTGTGG CTGGGAGGAA GCTTGTGAGG TCTGGGCCAC
5511 GGGACCTTGG GTGAGAGCTT AGTGCAGGGC TATGGTCCTT GCCTTGGGGT TGCAGCCACA
5571 GCCTCCCGTG GGAAGAAGT CGCAGAAGTA GGTGGCCTTG GCCCTGGGGG CCAGGAAGTT
5631 GCAGAGAGCC TGAGGGTGT TTTGGCGAGC CTGGGCAGAT AACTCCCCCTC TCCTGAGAAG
5691 CTTGCTGGGG GCGTCTGGTG TGTGATTCAG GCTGATAGGG TGGGAGGAAC CAGAAATTGC
5751 AGCCAGGAGA GGTGGTGGTG ATGGCTCAGC TTGGAGTTGA AAGGGATGTG GGACGGCGCG
5811 GGGGTGGGGG TACTTGGGGG CCGAGTTGG GGCTCCTCTG GCCTCAGATC TTGGGGTCTC
5871 TATCTCGTTT CTATGTCAGC CAACTCCCCG GGGGCTTTTG GGAAGTGGAG TCTGCAGGGG
5931 TGGGCGTGGG GGGCGTGTCT GCAGAGGCCA TTTTGGACGG GCCTGCGTGG CAAGGGGGAG
5991 GCGTCTGCTT CCGACACACA TGCTGGGGAA CGCCAGCCAG GAAGGGGAGG ATCCGAGCTC
6251 [200] GGGCC CCGTGGGGG AGGGTGTCTC CCTTTGAAAA ACGCTGCGCC CTGCTCTTCA
6311 GCCTCAGTTT CCCCATCTGT AAAGTGTGCA TTCCAGGCCT CCTTGGGCTG GCCCAGAGCT
6371 GCCCTGGGCA GGGGCTTTCA GCCCTTCA CAGAGGGCTG GGTCAGCTGA GTGAGTGCTC
6431 CTTTGTCTCT CCGCCATGCC CCAGGTCCT CCCTGCTCCC CAGGATCCAC ACCTACCTTG
6491 GCCCCGACCC TGGGCCATGC CTCACCACAC TTCCTGTCTT CTCTCACATC TCTCACATCT
6551 GGGAGTCCTC TCCCGCCAGC CTGTGCTTGC ATCTGGGGAT CCATGCCAGG GATAGGACCT
6611 G[400] CTGC AGTGAGCCGA GATCGTGCCA CTGCACTCCA GCCTGGGCAA CAGAACGAGA
7071 CCCTGTCTTA AAAAAATTTT TTTTGGCCTG GTGCGGTGGC TCACGCCTGT AATGCCAGCA
7131 CTTTGGGAGT CCAAGGTGGG TGGATCACCT GGAGGTCGGG AGTTTGTAGC CAGTCTGACC
7191 AACATGGAGA AACCCCATCT CTACTAAAAA CACAAAAATT AGCCGGGCGT GGTGGCGCAT
7251 GCCTGTAATC CCAGCTACTC TGGAGGCTGA GGCAGGAGAA TCCCTTGAAC CCGGGAGGCG
7311 GAAGTTGCAG TGAGCCGAGA TCGCACCATT GCACTCCAGC CTGGGCAACA AGAGTGAAC
7371 TCCATCTCAA AAAACAAAAC AAAACAAAAA AAACAACCAA AACAAACAAA CAAAAACTG
7431 TAATTAATAA AATTAAGA AAAAAGTTAA AAAAAAAAAA AAGAGAGGAT GGCTCAAAGG
7491 CTGGAAACAG GGAAGGCCAC TGTGAGGGGA GGACAGGCAG GGCTAGACCT CTCTGGAAGG
7551 AGGAGGGAGA GGTACC[300] GGGACTGAG TCGTGGTACA GATGGCGGCA GTTGTGTGGC
7911 CCTGGGGCCC AGGAATGGCT GGAAGCTCC CTTTCTTCTC TGGCTTGGGG GATGAGTTA
7971 GTGTAATTT CATGGAGTTG CCAGGAGGAT TAAGGCCGAC GTGCATTAC ACCTGGCCCC
8031 GTGGCTGCCG CCAGTTAAGC CCGCGCAGC CATTGGAAGT CTCATTGAGA AGGAGCCTCC
8091 CGGCTTCTTG CATTTGTTCC AGGCGGGCTG GTAAGCGCCC TGCTTATTTG CAAGGCTGTG
8151 TGAGGACGAG CCCTGTAAAC CTAGGACGA GGCTGTGCTA ATGTGGTTTC TGTTCCACTG
8211 TCTCCCTTTC TCCTTTTGCC CCTCTTTCTG GGAACCTGG AGCCCTGAT GCCCTCTGTT

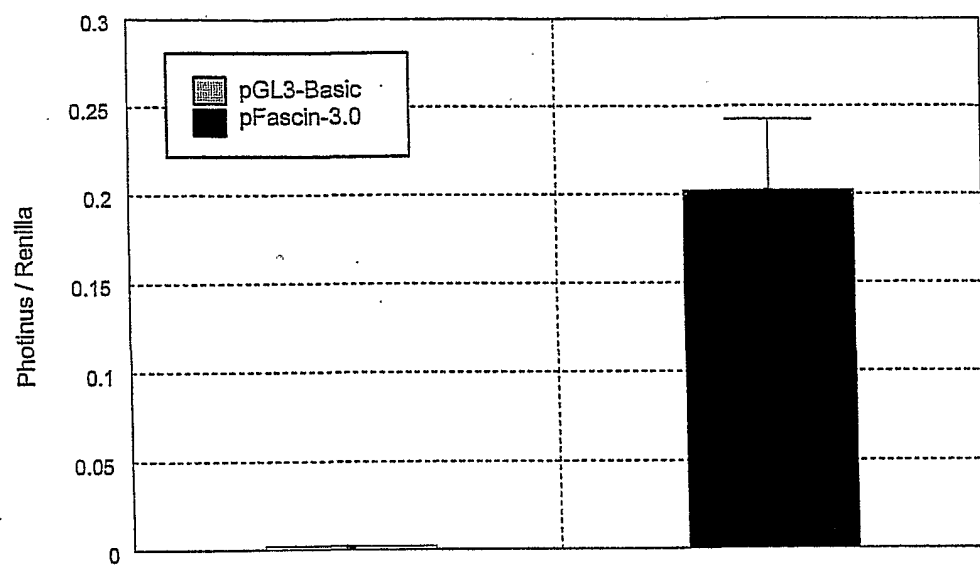
8271 TTTTCCTGCT GCTGGGGTGG ACCTAGAACC TCTCCCTCCA TCCTTTCTGC TGCTGGGGTG
8331 GGCTTGCAGA AGACTGTCTT CTTACCTGAG TCCCTAACTT CCCTCCGATC AGCGGGGGCT
8391 TCAGCGGAGC CCCAGCGCAT TGGACAGGTG GTGCTCTGTG GCCGGCACAG GCCAGGGCCC
8451 AGGGTGCAGC CCCTGCCTGT AGGTAGACAT GCAGGCCCTT CTGAGCAGAA TGAGGGGGCG
8511 GTGAGGGCAG CTGTGCGGGT GTGGCTGGCC TGACGGCTCC CTGAGGCCCT GCGGCTTCTG
8571 TGCAGTGGGG GTGGGGCGGG CCAGACCTTT GCAGGGCCTC TTGGCTGGTG GAGGGCTGAG
8631 TGAGCAGAGG CCCAGGCCCT CGTGTTCCTT GGGGTGTGGC CTTAGGATGG TCCCGGCACC
8691 CTGGGGACCC AGCCCCTGCT CACCTTATCC TCCTCCCGTG CTTGGCTGGG GTGTGGTGGC
8751 TGAGCCAGCC CACCCAGTGC GGTGGATGCT CGTCGGAGGG CAGAGGGTGG GCTACCGGCT
8811 TAAGGGGCCC CAGGGAACCT GGGGGGTGGA GCCCAAGGGG TTCCAAGAA GGGCGGGGCC
8871 AGGAACCCAT AGACAAGAGG GTGGGCGAGG GAAAGGGCGT GGCAAGAGGG CCACCCACCT
8931 CCGGTCCAGA GAGCCAGCCA GACCCCTTACT TTCTCTTGCT GTGTGACCTT AGGCAAGGAC
8991 ATGCCACCAC CGGCCCTTAGT CTCCCTCTTT GTGAAGTAGG ATGAAGATCC CCACCTGTGG
9051 AGCAGATGTG GGGCTGCATG GTTGGGTGAG GATGGAACAG GATGGGGAGG CCGGGCGTGG
9111 TGGCTTACCC CTGTAATCCC AGCACTTTGG GAGGCAAAGG TGGGAGGACT GCTTGAGTCC
9171 TGGAGTTTGG GCAACATAGC GAGACCACCC CCATCTCTAC AAAATAATGT TAAAGTTAGC
9231 CAGGTGTAGT GGTGAGTGCC TGTGGTCCCA CCTACTGGGG AGACTGAGGC AAGAGGATCC
9291 TTTGAGC[1. 500]CTGCAG CCTCCGCTC CTGGGTTCAA GCAATTCTCC TGCGTCAGCC
10851 TCCCGAGCAG CTGGGATTAC AGGCGCCCGG CACCACCCCC AGCTAATTTT TTTTTTTTTG
10911 AGACAGAGTC TCGCATTGTC GCCCAGGCTG GAGTGCAGTG GCGCGATCTT GGCTCACTGC
10971 AAGCTCCACC TCCCGGGTTC ACGCCATTCT CCTGCCTCAG CCTCCCGAGT AGCAGGGACT
11031 ACAGGCGCCC GCCACCGTGC CCGGCTAATT TTTTTTTTTT GTATTATTAA TAGAGACGGG
11091 TTTACCCGTG TTAGGATGAT CTCGATCTCC TGACCTCGTG ATCCGCTGCT CTCGGCCTCC
11151 CAAAGTGCTG GGATTACAGG CATGAGCCAG CTTGCCTGGT CCTTTTTTTT GTATTTTTAG
11211 TAGAGATGAG GTTTCACCAT GTTGGCCAGG CTGGTCTCAA ACTCCTGACC TATGATCCCG
11271 CCCGCTCGG C[560]CTGC AGAACACCCC TGTGTTGGCA AATACATCTC TTTCTTTTTG
11891 ATTCAGGGTC TTGCTCTGTT GCCCAGGTTG GAGTGCAGTG GCACAATCAT AGCTCACTGT
11951 AGCCTCGACT TCCTGGGCTC AAGTGATCCT CCCACCTCAG CCTCCCAAAT AGCTGAGACT
12011 ACAGGCACAC ACCACCACAT CAAGTTAATT TTTTTTTCTA TTTGGTACAG ACAGGGTCTG
12071 GCTATGTTGC CTAGGATGGT CTTGAACTCC CGGCCTCACA TGATCCTCCT AACTCAGCCT
12131 CCCAAAGTGC TGAGGTTACA GACACAAGCC ACCATGTGCC CAGCCAGTGA ACCACAACCT
12191 TTCTAGCTGT TTTTTCTTA TACATAAGAA GCATGCGGGT TGCCACGGCG GCTCACGCTT
12251 GTAATCCCAG CACTTTGGGA AGCTGAGGCG GGTGGATCAC CTGAGGTCAT GAGTTCAAGA
12311 CCAGCCTTG[1. 500]CAAT CCTCCTGCCT TGGGCCTCCC AAAGTGTGTC GATCACAGGC
13871 ATGAGCCCTT GTCCCTGGCC CCTATTATTT CTCTAACTGG GGAAAGTCAT GTGGGAACAG
13931 ATGTAGCTTG CCTTGGCCTC TGACCGGCCC TGCTTGCCTT CCTGGGTGCT CTCTGCTGCT
13991 TCTCATGTGT GCCACTGTGG GGACTCGGCC GCCCACCCCA CCCCGTGGTG TTACCTTGCG
14051 TGTGTAGTTC TGTGAGCTCA GGGCTATGGT CTGCCAGAAC TAGGGGGCGT GGGGCCCCAG
14111 TACCAGCCCA AGGCTCCTC TCTGCAAGTA TGGACCTGTC TGCCAATCAG GACGAGGAGA
14171 CCGACCAGGA GACCTTCCAG CTGGAGATCG ACCGCGACAC CAAAAAGTGT GCCTTCCGTA
14231 CCCACACGGG CAAGTACTGG ACGCTGACGG CCACCGGGGG CGTGCAGTCC ACCGCCTCCA
14291 CCAAGTGAAGT GCCTCGCTCC CACCTGTCAC CGCCCCACC ACCTTGCTG GGTACCCCG
14351 CCTGACCTG TCCCGCCATC CCCCAGGAAT GCCAGCTGCT ACTTTGACAT CGAGTGGCGT
14411 GACCGGCGCA TCACACTGAG GCGCTCCAAT GGCAAGTTTG TGACCTCCAA GAAGAATGGG
14471 CAGCTGGCCG CCTCGGTGGA GACAGCAGGT AACACTAAAG CCCAGTTCC CTGGAGCCGT
14531 CCTGGAGTCC TGGAGGGTCT GGCCATGCCG TGGTCACTTG GTAGCCCCAG CCAAGGCCTG
14591 CTCTGTGCTG GGCATCCCC CGGACTGGCC CCGCACTGTC CTACCCTGGG GACTGCTGTG
14651 TGACCCCAGC TCCTGGCCCT CCCTCTCTGG TCACCCCAGC CTCCACCCCA CTCCTGCCA
14711 GGAGGCTCAC TGAATCCCCT CTTTCTGGGA CAGGGGACTC AGAGCTCTTC CTCAATGAAGC
14771 TCATCAACCG CCCCATCATC GTGTTCCGCG GGGAGCATGG CTTCAATCGG TGCCGCAAGG
14831 TCACGGGCAC CCTGGACGCC AACCGCTCCA GCTATGACGT CTTCCAGCTG GAGTTCAACG
14891 ATGGCGCCTA CAACATCAA GGCAGGTTCT CCTGTGGGCA GCTGCTGGGC AGGGAACCCC
14951 TCGGTGCGGG CTGGGGTCCAG TGCTGCGGGG AGCGCCCTCT GCATCCACAC TGGACCTGG
15011 CTTGGCTCAG GGCCATTCCA GGCCCTAAAG GGACAGGTGT CTGATGGCCA CCAGGGGGCT
15071 CTGGGATGCA AGCAGCCCCT TTCCCTCTTG TCTGTGTGGT TGGGGGGACT TACCTTGCCC
15131 ACCTGACAGA GAGGTGTGTG GAGGGGAGAG CAGGGAGGGA AGGAGACCAG GAAAGGGAGG
15191 GAGGAGGACA GGGGAGGGA GAGCGGGGAA GAGAGGAGAG CAGGGGTGGG GAGGTTTCTG
15251 GAAAGGGTGT GCAAGGGGAG GACGCGCCTC GGTTATGGGA CTGGAGCCCC TTCCAGGAG
15311 GACCCCCAAC AATCCAGAGG TGCCTGTTAG GATTAGAAG ATGGTTTTTT TGTGTTTTTT
15371 TTTGAGACTC ACTCCCTCAC CCTGGCTGGA CTTGCAGTGG CGTTATCTCG GCTCACTGCA
15431 ACCTCTGCCT CCTGGGTTCA AGTGATTCTC CTGCCTCAGC CTCCAAGTA GCTGGGATTA

```

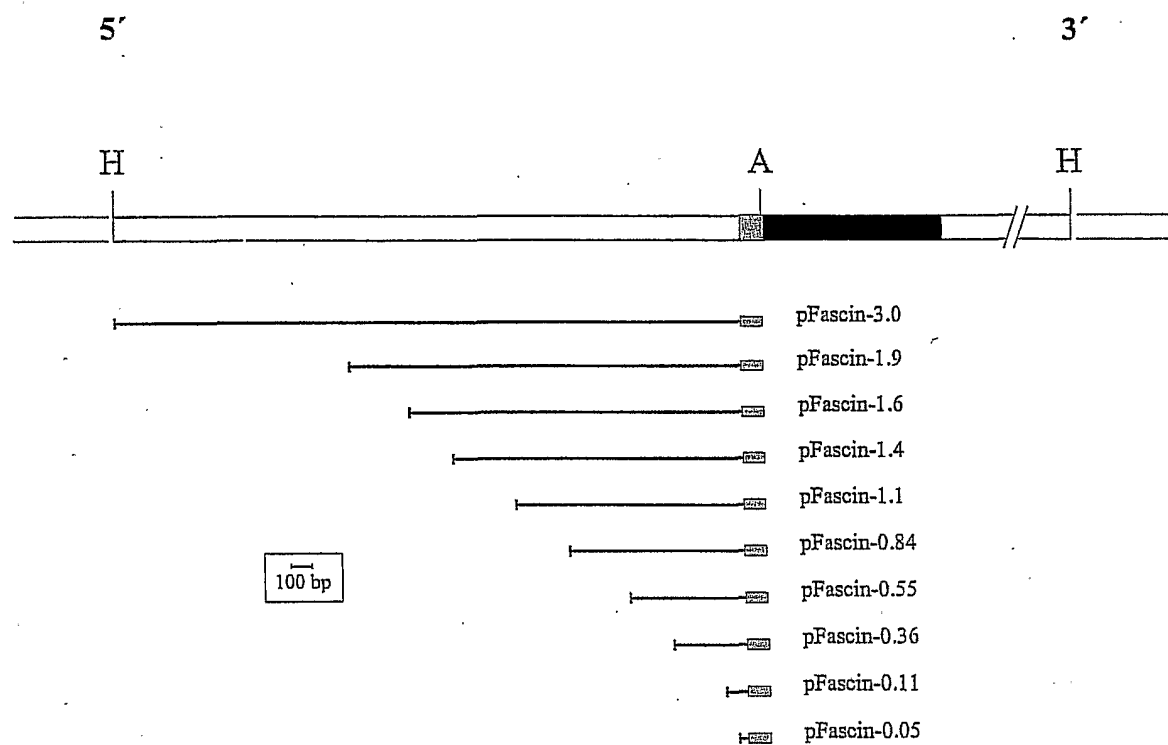
15491 CAGGTGTGCA CCACCACGCC TGGCTAATTT TTATATTTTT AAAATTTATT ATTTATTTAT
15551 TTTGAGACCC AGTCTGGAGT GCAGTGGCGT TATCTCGGCT CTCTGCAACC TCTGCCTCCT
15611 GGGTTCAAGC GATTCTCCTG CCTCAGCCTC CCGAGTAGCT GGGACTATGT GTGGGAGCCA
15671 CCATGCCTGG CTAATTTTTT TGTATTTTTT ATAGAGACGG GTTTCACCAT GTTGTCCAGG
15731 CTGGTCTTGA ATTCTGTGGC TCAAGTGATC CGCCACCTC AGCCTCCCAC AGTGCTGGGT
15791 TTATAGGTGT GAGCCACCAC ACCCGGCTAA TTGTTTTGTA TTTTATAGTAG AGACGGAGCT
15851 TCACTATGTT GGCAAGGCTG GCTCGAACTC CTGACCTCAA GTGATCCGCC CACCTCAGCC
15911 TCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGCGGCC GAGCAGAACA CGTTC TAGGA
15971 CCCTTGTTCA TGTGTCCATC ATGGACAGGA GGACGTGCGG GCCATAGGGA CCTGGCTCA
16031 TTCCGGAGCC GGGACTGGAG GGTGGGGCGT CACCCTTGGG AACACCCGTG CCCACCCTCC
16091 GCTGCCCAGG GTAGGGGTGG GGAGCCAGGC TTTGGGCCCC ACTTGATAAA GTCCCTCCC
16151 CAGACTCCAC AGGCAAATAC TGGACGGTGG GCAGTGAATC CGCGGTACAC AGCAGCGGCG
16211 ACACTCCTGT GGACTTCTTC TTCTGAGTTCT GCGACTATAA CAAGGTGGCC ATCAAGGTGG
16271 GCGGGCGCTA CCTGAAGGGC GACCACGCAG GCGTCTGAA GGCTCGGCG GAAACCGTGG
16331 ACCCCGCCTC GCTCTGGGAG TACTAGGGCC GGCCCGTCCT TCCCCGCCCC TGCCACATG
16391 GCGGCTCCTG CCAACCCTCC CTGCTAACCC CTTCTCCGCC AGGTGGGCTC CAGGGCGGGA
16451 GGCAAGCCCC CTTGCCTTTC AAAGTGGAAA CCCCAGAGAA AACGGTGCCC CCACCTGTGG
16511 CCCCTATGGA CTCCCCACTC TCCCCCTCCG CCGGGTTCCC TACTCCCCCTC GGGTCAGCGG
16571 CTGCGGCCCTG GCCCTGGGAG GGATTTTCTG TGCCCTTGCC CTCTTGTCTG CCACGGGGCG
16631 AGTCTGGCAC CTCTTCTTTC TGACCTCAGA CGGCTCTGA [ 90] CCACCTC CTCCCAGCCC
16781 CCCAGGAGAG CTGGGCACAT GTCCCAAGCC TGTGAGTGGC CCTCCCTGGT GCACTGTCCC
16841 CGAAACCCCT GCTTGGGAAG GGAAGCTGTC GGGTGGGCTA GGACTGACCC TTGTGGTGT
16901 TTTTGGGGTG GTGGCTGGAA ACAGCCCCTC TCCCACGTGG CAGAGGCTCA GCCTGGCTCC
16961 CTTCCCTGGA GCGGCAGGGC GTGACGGCCA CAGGGTCTGC CCGCTGCACG TTCTGCCAAG
17021 GTGGTGGTGG CGGGCGGGTA GGGGTGTGGG GGCCGTCTTC CTCTGTCTC TTTCTTTTCA
17081 CCCTAGCCTG ACTGGAAGCA GAAAATGACC AAATCAGTAT TTTTTTTAAT GAAATATTAT
17141 TGCTGGAGGC GTCCCAGGCA AGCCTGGCTG TAGTAGCGAG TGATCTGGCG GGGGGCGTCT
17201 CAGCACCTC CCCAGGGGGT GCATCTCAGC CCCCTCTTTC CGTCTTCCC GTCCAGCCCC
17261 AGCCCTGGGC CTGGGCTGCC GACACCTGGG CCAGAGCCCC TGCTGTGATT GGTGCTCCCT
17321 GGGCCTCCCG GGTGGATGAA GCCAGGCGTC GGGCCTCCG GGAGCCCTGG GGTGAGCCGC
17381 CGGGGCCCCC CTGCTGCCAG CCTCCCCCGT CCCCACATG CATCTCACTC TGGGTGTCTT
17441 GGTCTTTTAT TTTTGTAAAG TGTCAATTTGT ATAAGTCTAA ACGCCCATGA TAGTAGCTTC
17501 AAAGTGGAAA TAGCGAATA AAATAACTCA GTCTGCAGCC CCAGGCCGGC CTGTGTGTGT
17561 CTTGGGGCTG AGGTGGGTGG GGGGGCTGAG GTGGGTGGGA GGGCTGGCGG GACAGGTAGG
17621 CGCCCTGGCT CCCCAGCTCA GTGCTGGGAG TGTGAGTGG GAGGGAGGCC GTGGCTCCAG
17681 TGGGTGCTCC GGAGCTC

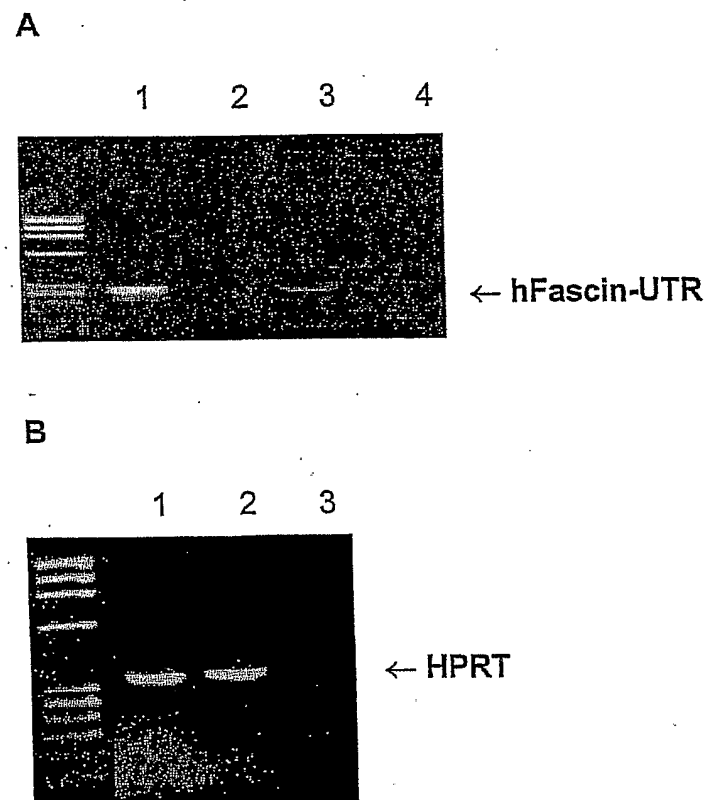
```

Figur 2

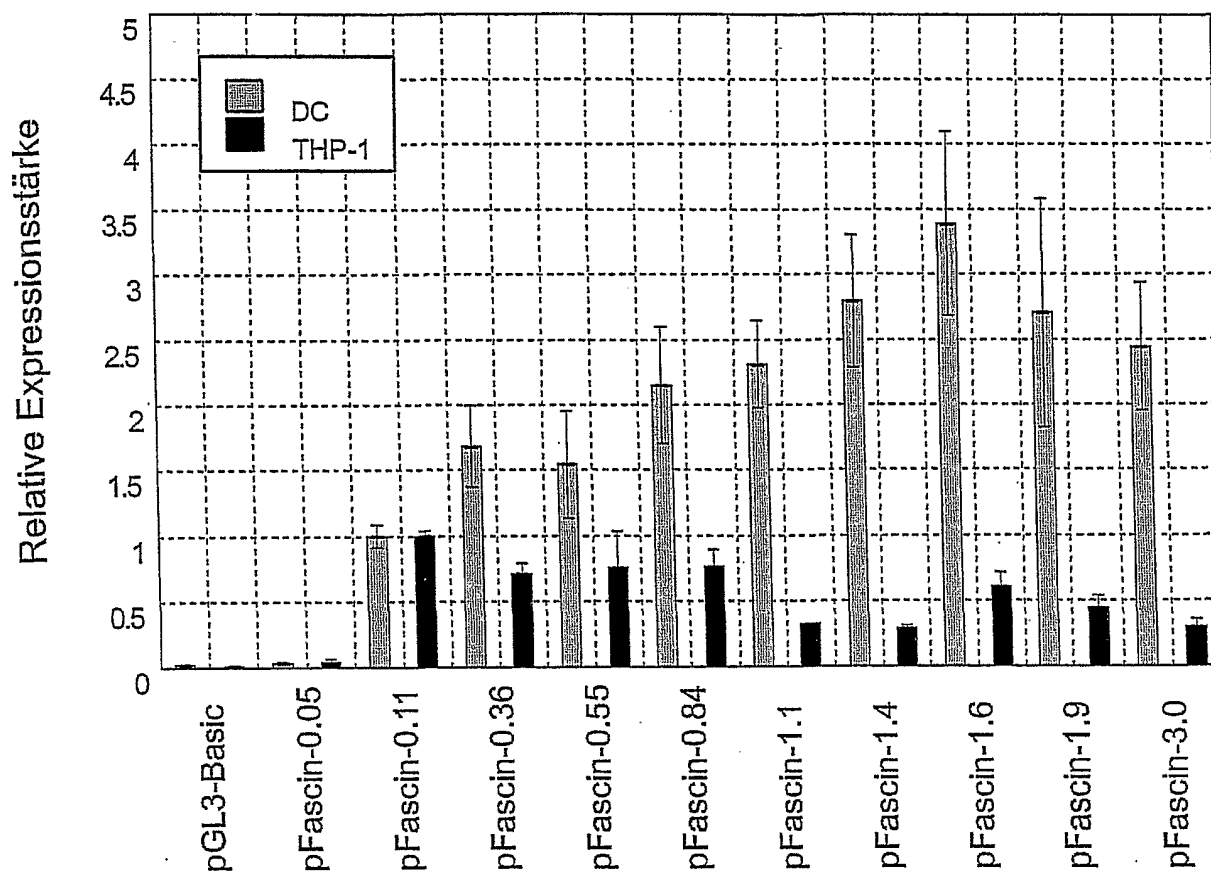


Figur 3

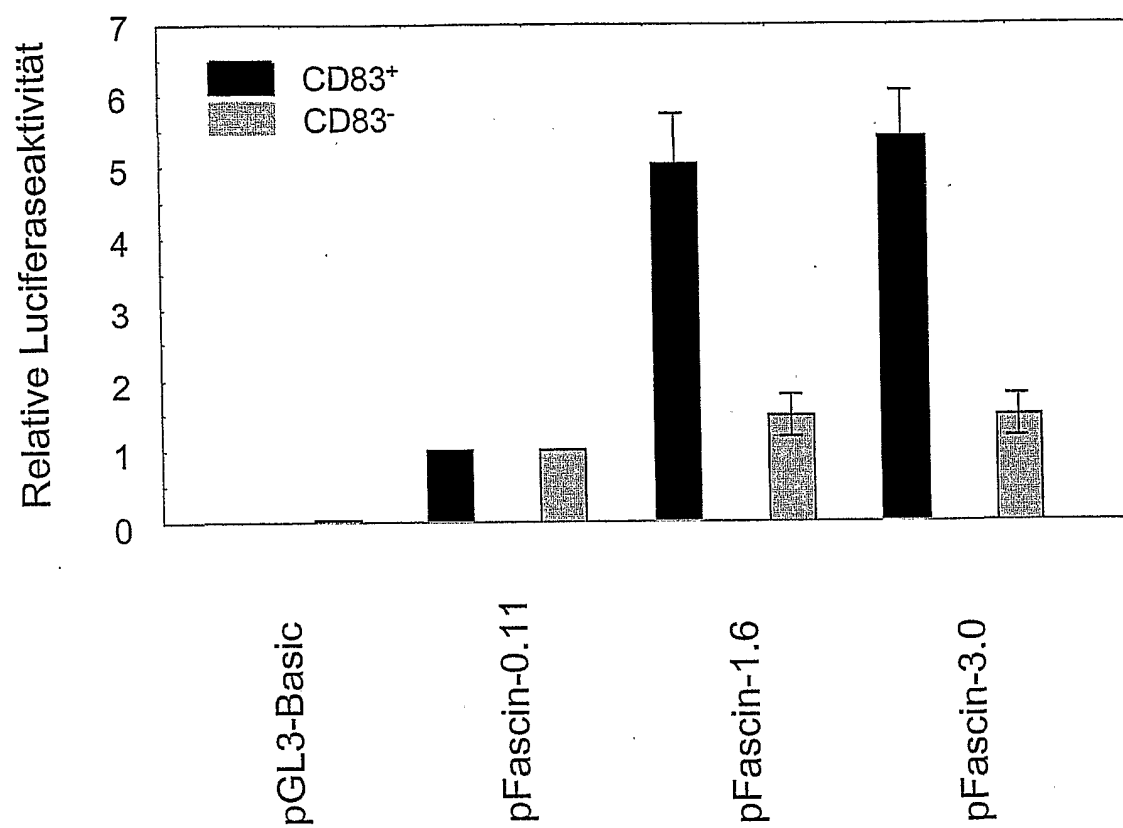
**Figur 4**



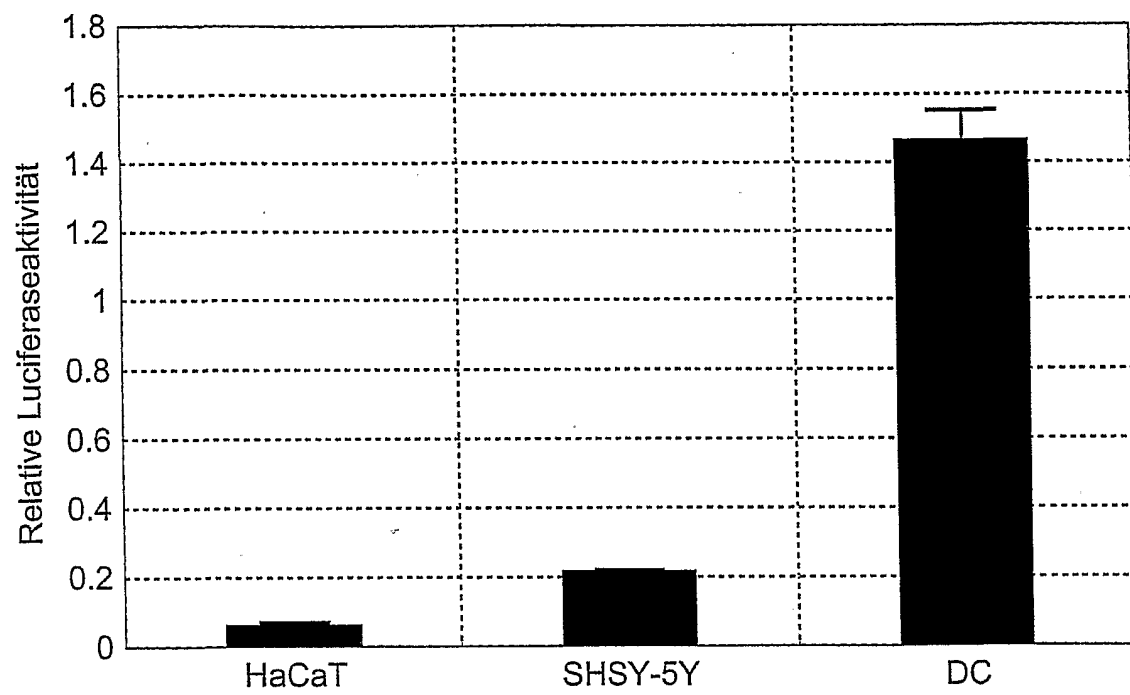
Figur 5



Figur 6



Figur 7

**Figur 8**

```

1  AAGCTTCTAG AAGGGGTATA TTTTGCCCCC AGGCCCCAAA TCCTGGTCTT TGGACTACAG
61  TCAAGACCTA GCACAGGGCT CAGGCCTAAA AAAAAGCTGT TGAGGAGGGT GTGAGGCTGC
121 AACGTTGCGG TGAGAAGGGG GTCCCCAGGG AGGGCGCAGC AGGAAGCCCC AGGGAAGTGC
181 CTGGGAGAGG GAGGTTGTGC AGCCAGAAGG AGCAGCCCGC AGCCTTTGGT TGTGACTCC
241 TGCTTTGTAA GTGGCAGTCT GGTGGGTAG GACAGGGTCC GAGTCCTCAC TCAGGGAAGT
301 GAGTCCAGGG TGAGCTGAGC AGCCCTTCCT GGTGGCCTCA CTTTCCCCTG CAGAGCCTGC
361 TGCATGGTGT CCAGTGGCCA GCCTGGGAGG AGCTCAGGAC TGGCCTCACC TCGTGCCACG
421 CCCTCGTACC AAGGTGGCTC AGATGGCACC TGGGTTCCCG AAGGGCCCAG GAACACAGCG
481 TCCATGTCCC CATCCTTCCC CGGAGGGACT TGGGGTGGGG CCTGCAGGAA GGATCATGTG
541 ACTTGTTCAA GGGCAGCTTG TCCTGCCCCG GACACAGAGG TGCCCATCGT AAGCGAGATG
601 CAGGCAAGGT GAGGACAGGG CATGGTGCCG AGCAGGAATG ATTTTCGAAA ATGCTACTCT
661 GTGGTCGGGC ACGGTGGCTC ACTCCTGTAA TCCCAGTACT TTGGGAGGCC CAGGCGGGCA
721 GATCATGAGG TCAGGAGTTC AAGACCAGCC TGGCCAACAT GGTGAAACCC CTTCTCTACT
781 GAAAATACAA AAAACTACCC GGGCGTGGTG GTGGATGCCT GTAATCCCAG CTACTCAGGA
841 GGCTGAGGCA GGAGAATCGC TTGAACCTG GAGGTGAAGG TTGCAATGAG CCGAGATTGC
901 ACCACTGCAC TCCGACCTGG GCGACAGAGA AAGACTCCGT CTGAAAAACA AACAAACAAA
961 ACCCCGAGAT TCTCTTTTCC CCCGTCCGGA GCTCTATGGC CATCTGGAGC TCGTTCTGTC
1021 CACAAGGACA CATTTCCCTG CAGCAGCTGT GGACCAGGGC TCACTCACTC ACATCTGTAC
1081 CCAATCTAGA GCAGGACAAA CACCTATCAC CTGCACTGGC AATGGACAGA GGACAGTGGC
1141 AGCCCCATCA CCAGAGGCAT TCAAGCCAAG GCATTTTCTA TCGTCTATTT ATCTATCTAT
1201 CTATCTATCT ATCTATCTAT CTATCTATCT ATCTATCTAG ATGAGATCTT GCTATGTTGC
1261 CCAGGTTGAA CTCCTGGCCT CAAGCGATCC TCCTGCTTCA GCCTCCCAA GTGTTGGGAT
1321 TTCAGGTGTG AACCACGTGA CCTGGCTGGA GTGCAGTGGC ACTGTCATAG CTCACTGCAG
1381 CCTCCACCTC CTGGGCTCAA GGAATCCTCT TTCCTCAGCC TCCTGAGTAG CTGGGACCAC
1441 AGGCATGCAC CATCACACCC AGCTAAACCT TGATTTTTTC CATAGTATAT GGCTCAGGGC
1501 AGAGCAAAGA AGCACAAAAA CATGTTGGCT TGCGGTTGAG GGCTGATGGG AGGGGGTTCT
1561 GGCCAAAGCC CAAAATTAAC AGCCACAACG TCAGTGTCTG TGGGAGGGGT TGCCAGGGGC
1621 GTGCGGGTTC TGGGGCTCAA GGCCCTGCCG GTAAACCCAT TTGAAGCAGG ATGGCAAGAA
1681 GGTGACACCA TCTTCCCCCG CGCACTGAAA GCCCCTGGCT GGGATTGCC TGGGCGAGCA
1741 CAGGCTCGGA CCAGCCCCAG CAATCCAGT TTATCAGCGA GCCGGCTGAG GGCCCGAGT
1801 TATCTCAGTG CCCGGCCTGA GACCTTGTGG GCAGCCTGTG GTCATGCTGG CATTCAGGG
1861 GCCTTTTGGC ATGTGGGGAA TGTCCAGGAA AAGCCTCAGC CTTGCTGAG CCGCAGAAAA
1921 GGGAAAGTGT CCTAGAGGGG GTGGGTGAGG GCCTGGGAGG TGGTGTCTGC AGGGAATGTC
1981 CCCTTTGGGG GAGGAGGATG GAGGGTTGGG ATTCTGAGGA TGGGGGGGGG GGCTGTAGCC
2041 AGCACCATGT CCCTCCTGTG TGACCAGCTC AGAGTCCCAT GAAATTGGGG CTTGGGAGGG
2101 GAAGGGACAC TGGCCTGGGA ACCAGAGACC TGGGCTGGTC TGGCTCACAG TTGCTGGACC
2161 TCTGTGATCC GTGTCAAAAA ACGAGAACAC CAATTCCTGT CTTGCCACT CACCACCAGG
2221 TGGGACCTGA ACCCTGGCAT CGCCAGCATT GGGAAATGTC GCCACTGACT CAACCACTCT
2281 CCCGGAGACC TATTTGGGCC ACCCGAGGCG GGTGCCTGGG CCACACGGAG GGGTCCTGGT
2341 GGTCTTCAGG GCAGCGGCTG TGGGGCTGAA GCCTCAAGGA ACCACATCTC TGCATAGGAG
2401 GGCCAGGCTG CAGGGCCTCG GGAGACAACC TAGTTGGGCG TGTGTTGGGTC TGTGGGTCCC
2461 AGGTCTTGGC CTCACCGGGT CCCCACCGCG CTGTCAGCTC CCAGCCTCTT TCCCTCGTCT
2521 GCCTCTGGGC TTCTGTAAAG CTATGTGCTC CAAGGCCACC TCCTCCAGGC AGCCCTCAGA
2581 CCCCCACCTT CGCCAGTAC CGATCTGCAC CGTGGTCTCT GAAGTCTCCA TCGTGACCTC
2641 AAACCTCGCT CGTCCTTGCT CTAGCGAGG CTTGGGGTCT GGGTGTCCGA GGTGGGGGAC
2701 ATCCGGGGGG GTTAGGTGGC TGGCGGGGG AGCCGGGGTT GTGAGGGGTG ATGTCCTCAG
2761 GCGGCGGCGC TGCGGGGTGC GCGAGGACA CCGGTGGGGT GAGAGCACCG CGGGGCGAGC
2821 AGCGGGGGGG GCAGCGCCGG GTCCCTCGGC CCGGGGCCCC TCCGCGCGGG AGCCAGGGGC
2881 GGGACAGGGG GCGGTGGCCT GGTGGCGCTG ACGTCACTC GCCTATAAAA TGTCCGGGGC
2941 GCCGCTAGCT GGGCTTTGTG GAGCGCTGCG GAGGGTGCGT GCGGGCCGCG GCAGCCGAAC
3001 AAAGGAGCAG GGGCGCCGCC GCAGGGAACC GCCACCCACC TCCCGGGGCC GCGCAGCGGC
3061 CTCTCGTCTA CTGCCACCAT GACCGCCAAC GGCACAGCCG AGGCGGTGCA GATCCAGTTC
3121 GGCCTCATCA ACTGCGGCAA CAAGTACCTG ACGGCCGAGG CGTTGCGGTT CAAGGTGAAC
3181 GCGTCCGCCA GCAGCCTGAA GAAGAAGCAG ATCTGGACGC TGGAGCAGCC CCCTGACGAG
3241 GCGGGCAGCG CGGCCGTGTG CCTGCGCAGC CACCTGGGCC GCTACCTGGC GCGGGACAAG
3301 GACGGCAACG TGACCTGCGA GCGCGAGGTG CCCGCTCCCG ACTGCCGTTT CCTCATCGTG
3361 GCGCACGACG ACGGTGCGTG GTCGCTGCAG TCCGAGGCGC ACCGGCGCTA CTTGCGCGGC

```

Fig. 9

3421 ACCGAGGACC GCCTGTCCTG CTTTCGCGCAG ACGGTGTCCC CCGCCGAGAA GTGGAGCGTG
 3481 CACATCGCCA TGCACCCTCA GGTCAACATC TACAGCGTCA CCCGTAAGCG CTACGCGCAC
 3541 CTGAGCGCGC GGCCGGCCGA CGAGATCGCC GTGGACCGCG ACGTGCCCTG GGGCGTCGAC
 3601 TCGCTCATCA CCCTCGCCTT CCAGGACCAG CGCTACAGCG TGCAGACCGC CGACCACCGC
 3661 TTCCTGCGCC ACGACGGGCG CCTGGTGGCG CGCCCCGAGC CGGCCACTGG CTACACGCTG
 3721 GAGTTCCGCT CCGCAAGGT GGCCTTCCGC GACTGCGAGG GCCGTTACCT GGCGCCGTCG
 3781 GGGCCCAGCG GCACGCTCAA GGCGGGCAAG GCCACCAAGG TGGGCAAGGA CGAGCTCTTT
 3841 GCTCTGGAGC AGAGCTGCGC CCAGGTCTGT CTGCAGGCGG CCAACGAGAG GAACGTGTCC
 3901 ACGCGCCAGG GTGAGTGGGG ACCTGCCCC CGCCTCTCCT GGTCCGTGCA CAAAGCGCAC
 3961 CCCACCCGCG CCCCTCCAGC CTCCCGCCCT TTCTCGCTCG CGGCGCCGCT GCGGTCCGGA
 4021 GCACTGCCCA TTGCGCCCC GCTAGGCACG CGGGCTACCC CGCCTGGAGG GGGCGAGGAG
 4081 TGGGGCTTTG CCCATCCTCG GGTGCGCTG CCCACCTCCC ACCCCGGGCT GGGATCATGG
 4141 GCTCCCCTAG GCCCCGCGGA GTCGCAACCG TACGTGCACC CTCCTAACCC CCCCCCCGCG
 4201 CCAATCTTGG CTCTCCCCAC GCGCGCTGAT CCCTCGGATC AGCTCTTGCGG
 4261 AGTGGGAGCC CTTACCCAT TTCTCTGTGC CCTCCCCC GCGGCTGTGCT CTGGAGCTCG
 4321 GGTGTCGAG GCAAGGGTCG CCACCCGCAA GGGCGCGCCT CCACCCCCAC CGGCAGCCTT
 4381 TCGCGGGCGA GATGGGGGAG GTTAGCCAGG CCTTTGATCC CGGCGGGGCG CGCCTCCACC
 4441 TCCCCGTCTG CCCGCGCTTC CTCCACCTCC TCCCGCTGCC GGGCGGGGTC GGCCTCCGCT
 4501 GGTGGGGGGG GCGCGGGGGT GTCAGCCCTT CCCCCAGCC CCTCCTCCCG CGTCTGCCCC
 4561 GCGCTCGAGG CCGCGGCCCT TGTGAGCAGG GGGCGGGGTC GCCTCGACTG GGTCTCCCT
 4621 CGCCCCCTTT GTCTTGCTCC ATCTCTGCAG ATGGGAAAAC CAGATGCGGC GGGCGGGGGG
 4681 AGGGGATCGG CTTTTGCGGT TCACCCCTGC AGAGGAGCCC CCCGCGCCGC CCCCCGGGCC
 4741 AGGGCTCGGG TCCCAGGTT CCCCAGGAGG CGGTCTCCCT CCTCGTCGCC GCGGCCCGGG
 4801 AACGGCGTGG CGCGGATGGC GGCCCTCCAG GCACCCCGCC CTTGCGCCCG CGCGCGCGTC
 4861 TCCAGGCCGG TCGCTGAGC TCCGCTCCGC GGGCGACCGA GGGCGGCTTC AGCGCGAGCC
 4921 GGGAGACCC AGGCCAGCTC CTTGCGGAAG CCCCTACCT CTGGTGAACC CATCCCCTAG
 4981 GCGCTCGCC GGGAAACAATT GACTGCAACT CGATCCGTCC CCTCCGGGCT CTCCTGCAGG
 5041 CACGGCTATT TGCGACCGGC CTGTGCGCCA CTCTTTCCCT TCTTTCTCAG ATTCTCCCCG
 5101 AGGGGCTTTT TCTTCTTTT GCGCTCTCT TCCACTCCTG GCGGAGAAGC TGTTCCTTGC
 5161 TTGCAGACAG AGGCCCCGCT GGAGGGAGG GCGTGGGGGG ATCTTTTTC CTGGATAGGA
 5221 AGTCGGACCC CAGGCCCTGG AGATGGCTGG ACGGGAGCCC AGTTTGTGTC CTCATCCCTC
 5281 CTCTTGAAG GAGCCCTCCT GACGGCCCC CTCTTGCTG TGGGCTCTGC AGGAGGGGGA
 5341 AAGACCCCC ATGGCAGTGG GGGTGAGGGT GGAGGCCTGG GTGGGGTAAT GGCCATTTTG
 5401 TGCTGAAGT TTGCTACCTT AAGTCCCAGA GAGGTGAAGC TGCTTGTGAT GGGTCACACA
 5461 GCAAGTGACC ACAAGGAAGC AGCATGCAGA AGTGGGTGCA TTTGGCTCCA GAAACCCCGT
 5521 TTCCTGTTT CCACAGCGCA TGGCGGGCAG AGCTCCTCAC CTGATCAGCA GGCATTGAGC
 5581 CCTACTAAGG GGACCTGCTG AGCCATGAGT GAACCAAGTG GGTTCACGTT CTCTTGGGGA
 5641 CATTACGAAG TGTCAGGAA AAGGCAGCG GTACAGGTTA TTAGATGGCT GAGTCCTTTT
 5701 TGAGCAGGAC ATTTGAGCTG GCCTGAAGGA TGAATAGAAG GTGGCAGGTA TGGGGGGGGT
 5761 GCTGTGTGGA CCAGCGTTGT AGGCTGTGGG TGTCTCCTGT GCAAAGGTCT TATGGCAGGA
 5821 AGGTACAGAG AAAGCCAGT GTGGCTGGGA GGAAGCTTGT GAGGTCTGGG CCACGGGACC
 5881 TTGGGTGAGA GCTTAGTGCA GGGCTATGGT CCTTGCCTTG GGGTTGCAGC CACAGCCTCC
 5941 CGTGGGGAAG AAGTCGCGA AGTAGGTGGC CTGCGCCCTG GGGGCCAGGA AGTTGCAGAG
 6001 AGCCTGAGGG TGTTTTGGC GAGCCTGGGC AGATAACTCC CCTCTCCTGA GAAGCTTGCT
 6061 GGGGCGCTCT GGTGTGTGAT TCAGGCTGAT AGGGTGGGAG GAACCAGAAA TTGCAGCCAG
 6121 GAGAGGTGGT GGTGATGGCT CAGCTTGGAG TTGAAAGGGA TGTGGGACCG CGCGGGGGTG
 6181 GGGGTACTTG GGGGCCGAGG TTGGGGCTCC TCTGGCCTCA GATCTTGGGG TCCTTATCTC
 6241 GTTTCTATGT CAGCCAACTC CCGGGGGGCT TTTGGGAACT GGAGTCTGCA GGGGTGGGCG
 6301 TGGGGGGCGT GTCTGCAGAG GCCATTTTGG ACGGGCCTGC GTGGCAAGGG GGAGGCGTCT
 6361 GCTTCCGACA CACATGCTGG GGAACGCCAG CCAGGAAGGG GAGGATCCGG ACTTAGCTGG
 6421 CAGGGGGGAT GTGAATCATC TCTGCAGGCC TGCACGGGCG GGCCCGGTGG GGGCAGGGTG
 6481 CTCCCCTTTG AAAAACGCTG CGCCCTGCTC TTCAGCCTCA GTTTCCCCAT CTGTAAAGTG
 6541 TGCATTCCAG GCCTCCTTGG GCTGGCCCAG AGCTGCCCTG GGCAGGGGCT TTCAGCCCTT
 6601 CACACCAGAG GCTGGGTCAG CTGAGTGAGT GCTCCTTTGT CCTCCCGCCA TGCCCCAGGC
 6661 TCCTCCCTGC TCCCCAGGAT CCACACCTAC CCTGGCCCCG ACCCTGGGCC ATGCCTCACC
 6721 AACTTCTCTG TCTTCTCTCA CATCTCTCAC ATCTGGGAGT CCTCTCCGCG CAGCCTGTGC

Fig. 9 cont.

6781 TTGCATCTGG GGATCCATGC CAGGGATAGG ACCTGTCTCTC CCTGCTGGCT TGGGACATGC
 6841 CCTCTCCCAG CCACTCTGAG GAGGGCTGAC TGAGGAAGGG CTCGTCAAGC TGGGCTTTGC
 6901 AGGATGTGTA AGAGTTCTCC CCACGAGGCG CAGGGCATTG TGAGCCCAGG GAATGGCTTG
 6961 TATAAGGATG CAGAGGCATT TGAAATGGCC AAGTAGTTGC AGGAATCGAC TGGAAACCGG
 7021 GGTGGTAAGG TGAAGCCATA GAACATTCCA GGGCCCCCTCC CCTAAATGAG ATGGAAGGAG
 7081 TGCCTGTTTT GAACAAGCCA GGGCATCTGG GGACCGTTAA GGCCTGGGGG TGGTGATGGG
 7141 GACTGGAGGG TGTGAGGCAA CCAGGGGGTG CCCCTCTGAG CCACCACAGA CAGAATGGTT
 7201 CAGAAGGCCA GGCACACTGG CTCTCGCCTG TAATCCAAGC ACTTTGGGAG GCCAAGGAGG
 7261 GAAGATGGCT TCGGCCAAGG AGGGAAGATG GCTTCAGCCC AGGAGTTCAA GGCCAGCCTG
 7321 GAAAACATGG CAAAACCTCT GTCTACAAAA AATACAAAAA ATTAGCCAGG CATGGTGATG
 7381 CACGCTGTGC GTCCCAGCTA CTCAGGAGGC TTAGATGGGA GGATCACTTG AGCCGGGGAG
 7441 TTTGAGGCTG CAGTGAGCCG AGATCGTGCC ACTGCCTCC AGCCTGGGCA ACAGAACGAG
 7501 ACCCTGTCTT AAAAAAATTT TTTTGGCCCT GGTGCGGTGG CTCACGCTG TAATGCCAGC
 7561 ACTTTGGGAG TCCAAGGTGG GTGGATCACC TGGAGGTCGG GAGTTTGAGA CCAGTCTGAC
 7621 CAACATGGAG AAACCCATC TCTACTAAAA ACACAAAAAT TAGCCGGGCG TGGTGCGCA
 7681 TGCCTGTAAT CCCAGCTACT CTGGAGGCTG AGGCAGGAGA ATCCCTTGAA CCCGGGAGGC
 7741 GGAAGTTGCA GTGAGCCGAG ATCGCACCAT TGCCTCCAG CCTGGGCAAC AAGAGTGAAA
 7801 CTCCATCTCA AAAAAACAAA CAAAAACAAA AAAACAACCA AAACAACAA AAAAAAACT
 7861 GTAATTAAAA TAATTAAAG AAAAAAGTTA AAAAAAATAA AAAGAGAGGA TGGCTCAAAG
 7921 GCTGGAAACA GGAAGGCCA CTGTGAGGGG AGGACAGGCA GGGCTAGACC TCTCTGGAAG
 7981 GAGGAGGGAG AGGTACCCGT GGGCCCCGCA CAGGAGACTC TTAATCTCCT GGCTCCCGGG
 8041 GGCTTCTGTG GGCCTGTGAC TCAGGATTTT TGTGCCTCTG TTGATGAAAA GAAAAATCCT
 8101 GAAGAAAACA CTGTTCCCAT AGTAACACAG CTCTAATTAG AGACTCCAAG GCCAAGCGAG
 8161 GGCCTTGGAG CCAGGAGGGG CCCTGCTCTA CTCGGGGGAC GACACCCCTT TCCCCTATCC
 8221 CTCCCTGCTG GGGTGTGGCC TGATGTCTAC GTGGCACCAG GCCCCTAAAT CTTTCAACA
 8281 TCCTGGCGGA GCGGAGAGGC TGGGCCCTTT TGACAGGTGG GGAACCTGAG GAAGGTTGCA
 8341 GGAAGGCAC TCATCCAAGG GTGCCTGGCC CCCC CGGTG GGTGCTCAC TGGCTGGCCT
 8401 AAACCTTTGC ATCAAACCAG TTCCAGGATT GAACGCAACA GGCTGATGGG GACTGAGTCG
 8461 TGGTACAGAT GCGGCGAGTT GTGTGCGCCT GGGCCAGGA ATGGCTGGGA AGCTCCCTTT
 8521 CTTCTCTGGC TTGGGGGATG AGGTTAGTGT AAATTTTCATG GAGTTGCCAG GAGGATTAAG
 8581 GCGGACGTGC ATTCACACGT GGCCTCCGTC CTGCCGCCAG TTAAGCCCTG CGCAGCCATT
 8641 GGAAGTCTCA TTGAGAAGGA GCCTCCCGGC TTCTGCATT TGTTCAGGC GGGCTGGTAA
 8701 GCGCCCTGCT TATTTGCAAG GCTGTGTGAG GACGAGCCCT GTAAACCTTA GGACGAGGCT
 8761 GTGCTAATGT GGTCTCTGTT CCACTGTCTC CCCTTCTCCT TTTGCCCTC TTTCTGGGGA
 8821 ACCTGGAGCC CCTGATGCCC TCTGTTTTTT CTTGCTGCTG GGGTGGACCT AGAACCTCTC
 8881 CCTCCATCCT TTCTGCTGCT GGGGTGGGCC TGCAGAAGAC TGTCCTCTTA CCTGAGTCCC
 8941 TAACCTCCCT CCGATCAGCG GGGGCTTCAG CGGAGCCCCA GCGCATTTGA CAGGTGGTGC
 9001 TCTGTGGCCG GCACAGGCCA GGGCCCAGGG TGCGGCCCTT GCCTGTAGGT AGACATGCAG
 9061 GCCCTTCTGA GCAGAATGAG GGGGCGGTGA GGGCAGCTGT GCGGGTGTGG CTGGCCTGAC
 9121 GGCTCCCTGC AGCCCTGCGG CTTCTGTGCA GTGGGGGTGG GCGGGGCCAG ACCTTTGCAG
 9181 GGCCTCTTGG CTGGTGGAGG GCTGAGTGAG CAGAGGCCCC AGCCCTCGTG TTCCCTGGGG
 9241 TGTGGCCTTA GGATGGTCCC GGCACCCTGG GGACCCAGCC CTGCTCACC TTATCCTCCT
 9301 CCCGTGCTTG GCTGGGGTGT GGTGGCTGAG CCAGCCCACC CAGTGCAGGT GATGCTCGTC
 9361 GGAGGGCAGA GGGTGGGCTA CCGGCTTAAG GGGCCCCAGG GAACCTGGGG GGTGGAGCCC
 9421 AAGGGGTTCC AAGAAGGGG GGGGCCAGGA ACCCATAGAC AAGAGGGTGG GGCAGGGAAA
 9481 GGGCGTGGCA AGAGGGCCAC CCACCTCCGG TCCAGAGAGC CAGCCAGACC CTACTTTCT
 9541 CTTGCTGTGT GACCTTAGGC AAGGACATGC CACCACCGGC CTTAGTCTCC CTCTTTGTGA
 9601 AGTAGGATGA AGATCCCCAC CTGTGGAGCA GATGTGGGGC TGCATGGTTG GGTGAGGATG
 9661 GAACAGGATG GGGAGGCCGG GCGTGGTGGC TTACCCCTGT AATCCAGCA CTTTGGGAGG
 9721 CAAAGGTGGG AGGACTGCTT GAGTCCTGGA GTTTGGGCAA CATAGCGAGA CCACCCCAT
 9781 CTCTACAAAA TAATGTTAAA GTTAGCCAGG TGTAGTGGTG AGTGCCTGTG GTCCACCTA
 9841 CTGGGGAGAC TGAGGCAAGA GGATCCTTTG AGCCAGGAG GTGGAGGCTG TAGAGAGCCA
 9901 TGCTTGGGCC ACTTGCACTC CAGCCTGGGC AACAGAGCAA GACCCTATCT CTAACAAAAA
 9961 AATGGAATTG AGGATGTGTC CTCTGGGTGT GGGCTCTACC ACCCACCAGC CTCCCTCTCC
 10021 TGTGGGTGCT GCCTGCCACC CACCGCTGAG GTCCCTGCAG CATCAAGGTG CCCAACGAA
 10081 GACACTCTCC AGCTCTGAGC CAGGCACCCA TACAGAGCCC GGCAGACGCC TCTGCTGCTG
 10141 CAGTCTTAG AATCCAGAGC GCGTGGGGAA TGTGAATTTG TGCTGCTGGA GCCAACATC

Fig. 9 cont.

10201 CCACCTCCAG GTTTTAGCTG GAGGAATCCA TGTGGATATG CAAGCAATGT AGTTATAATG
 10261 AATAGCAATT GTGGGCTGGG CGCAGTGGCT CATACTGTA ATCCCAGCAC TTTGGGAGGC
 10321 TGAGGTGGGT GGATCACCTG AGGTCAGGAG TTTGAGACCG GCCTGGCCAC ATAATGAAAC
 10381 TCCGTCTCTA CTAAAAATAC AAAAAAATTA GCCGGGCATA GTGGCGGGTG CCTGTAATCC
 10441 CAGCTACTCA GGAGGCTGAG GCAGGAGAAT CACTTGAAAC CCAGGAGGCA GAGGTTTCAG
 10501 TGAGGCGAGA TCGTGCCATT GCACTCCAGC CCGGGAGACA CAGCGAGACT CTCTTTCTCA
 10561 AAAAAAGAAT AGCAGCTTTG CTATAGCTTA GGGGAGCAAA ACCCGAAAGC CACCTTAAGT
 10621 TTCCCGAAGA TAGAGATCCC TGGGGTCCTA TTCTGAGACA GAGTCTTGCT CTGTCGCCCA
 10681 GGCTGGAGTG CAGTGGCGGA TCTCAGCTCA CTGCAGCCTC CGCCTCCTGG GTTCAAGCAA
 10741 TTCTCCTGCG TCAGCCTCCC GAGCAGCTGG GATTACAGGC GCCCGGCACC ACCCCAGCT
 10801 AATTTTTTTT TTTTGTAGAC AGAGTCTCGC ATTGTCGCCC AGGCTGGAGT GCAGTGGCGC
 10861 GATCTTGGCT CACTGCAAGC TCCACCTCCC GGGTTTCAGC CATTCTCCTG CCTCAGCCTC
 10921 CCGAGTAGCA GGGACTACAG GCGCCCGCCA CCGTGCCCGG CTAATTTTTT TTTTTGTAT
 10981 TATTAATAGA GACGGGTTTC ACCGTGTTAG GATGATCTCG ATCTCCTGAC CTCGTGATCC
 11041 GCCTGCCTCG GCCTCCCAA GTGCTGGGAT TACAGGCATG AGCCACCTTG CCTGGTCTTT
 11101 TTTTGTGTAT TTTTAGTAGA GATGAGGTTT CACCATGTTG GCCAGGCTGG TCTAAACTCC
 11161 TGACCTCATG ATCCGCCCCG CTCGGCCTCC CAAAGTGCTG GGATTACAGG TGTGGGCCAC
 11221 TGTGCCCCGG CAGTCCCATT CTACTTTTTT GGGACATGGA ATCATGTAAA GATGCCATAC
 11281 AGAAAACAGT ATGCAAATCC AGAGACAGAG GACATCTGTG CATGTTTCTT TCTTCTTCC
 11341 TTTTTTTTTT TTTTTTTTTT TTTGAGATGG ATTCTTGCTC TGTGCGCCAG GCTGGAGTGC
 11401 AGTGGTGCAA TCTCAGCTCA CTGCAACCTC TGCCTCCCAG GTTCAAGTGA TTCTCCTGCC
 11461 TCAGACTTCC AAGTAGCTGG GATTACAGGC ATGCCCCACC ATGCCAGCT AATTTTTGTA
 11521 TTTTGTAGTAG AGGCGGGGTT TCACCATGTT GGCAAGGCTG GTCTTGAAGT CATGACCTCA
 11581 GGTGATCCAC CCGCCTCAGC TTCCCAAAGT GCTGGGATTA CAGGCGTGAG CCGCCGTGGC
 11641 AACC CGCATG CTCTTATGT ATAAGAAAAA AACAGCTAGA AAGTTGTGGG TTCCTGGCT
 11701 GGGCACATGG TGGCTTGTGT CTGTAACCTC AGCACTTTGG GAGGCTGAGT TAGGAGGATC
 11761 ATGTGAGGCC GGGAGTTCAA GACCATCCTA GGCAACATAG CCAGACCCTG TCTGTACCAA
 11821 ATAGAAAAAA AAATTAACCT GATGTGGTGG TGTGTGCCTG TAGTCTCAGC TATTTGGGAG
 11881 GCTGAGGTGG GAGGATCACT TGAGCCCAGG AAGTCGAGGC TACAGTGAGC TATGATTGTG
 11941 CCAGTGCACCT CCAACCTGGG CAACAGAGCA AGACCCTGAA TCAAAAAGAA AGAGATGTAT
 12001 TTGCCAACAC AGGGGTGTTT TGCAGTGGGG TGGGAAGGGG ACGTGTAGGA CTTCATAGTT
 12061 TCCTTTTTTT TGTTTTAGAG ACAGGGTTTA AAAAAAATT TTTTTTTTTT TTTTGAGGCG
 12121 GAATTTTATT CTGTGTGCCC AGGCTGGAGT GATCTCGGCT CACTGCCATC TCCGCTCCC
 12181 AGGTTGAAGT GATTCTCCTG CCTCAGCCTC AGGCAATTTT GTATTTTTTAG TAGAGACGGG
 12241 GTGACTCCAT GTTGGTCAGG CTGGTCTCAA ACTCCCGACC TCGGGTGATC TGCCTGCCTT
 12301 GGCCTCCCAA ACTGCTGGGA TTACAGGTGT GAGCCACTGT GCCTGGCCCC TTTTTTTTTT
 12361 TTTTAAGTAG GGTCTTGCTA AGTTGTCCAG GCTGGTCTTG AACTCCTGGT CTCAAATGAT
 12421 CCTCCTGACT TGGCCTCCCA AAAGGCTGGG ATTGCAGGAG ATATGAGCCA CCGCACCCCG
 12481 CCCTTTTATA AGTTTTCTAA AGGAAAGGAA CATGTACCAT CGTGTTTTAT TTTAATTTTA
 12541 CATAGAAGCA GGAACAATAG CATGAACCCC CAGACACGGC CCCCCAAGAT GCGCTGTTCC
 12601 GCTGTGCTGC AGCCCCATTT GTCTTTTCT ATACCATTTT TGGTTGTATT TGAAGCAAAT
 12661 CCCTGACATC AGGGCATTTT ATCCCAAAAT ACTTCCATCT GTGTTTCTAA AAAAACGAGG
 12721 CCATTTTCTT ATGTAACCAC AACGTGATCA ACAGTAATCA ATGCTTAACA TCTAATTCTG
 12781 GGCCACGTTA AGACCCTCCT GATTGGCTCA AGAATGTCTT TTGTGCAGTT GGTTCCTTGA
 12841 TTCTTTTGT TTTTGTGAG ACAGGGTCTC ACTGTGTCAC CCAGGCTGGA GTGCAGTGGT
 12901 GCAATCACGG CTCACTGCAG CCTCAGCCTC CTGCCTCAGC GATCCTCCTG CCTCAGCCTC
 12961 CCAAGTAGCT GGGACCACAG GTTTGCACCA CCACTCCCAG CTAATTAAAA AAAGTAATTT
 13021 GTAGAGACGG GGGTGGGGTG GGGGTCTTGC TATGTTGCCC AGGCTGGTCT CAAACTCCTA
 13081 GACTCAAACA ATCCTCCTGC CTGGGCTTCC CAAAGTGTTG CGATCACAGG CATGAGCCCC
 13141 TGTCCCTGGC CCCTATTATT TCTCTAAGTG GGGAAAGTCA TGTGGGAACA GATGTAGCTT
 13201 GCCTTGGCCT CTGACCGGCC CTGCTGCGT TCCTGGGTGC TCTCTGCTGC TTCTCATGTG
 13261 TGCCACTGTG GGGACTCGGC CGCCACCCC ACCCGTGGT GTTACCTTGC GTGTGTAGTT
 13321 CTGTGAGCTC AGGGCTATGG TCTGCCAGAA CTAGGGGGCG TGGGGCCCCA GTACCAGCCC
 13381 AAGGCCTCCT CTCTGCAGGT ATGGACCTGT CTGCCAATCA GGACGAGGAG ACCGACCAGG
 13441 AGACCTTCCA GCTGGAGATC GACCGCGACA CCAAAAAGTG TGCCTTCCGT ACCCACACGG
 13501 GCAAGTACTG GACGCTGACG GCCACCGGGG GCGTGCAGTC CACCGCTCC AGCAAGTGAG
 13561 TGCCTCGCTC CCACCTGTCA CCGCCCCAC CACCTTGCCT GGGCTACCC GCCTGACCCT

Fig. 9 cont.

13621 GTCCCGCCAT CCCCCAGGAA TGCCAGCTGC TACTTTGACA TCGAGTGGCG TGACCGGGCGC
 13681 ATCACACTGA GGGCGTCCAA TGGCAAGTTT GTGACCTCCA AGAAGAATGG GCAGCTGGCC
 13741 GCCTCGGTGG AGACAGCAGG TAACACTAAA GCCCCAGTTC CCTGGAGCCG TCCTGGAGTC
 13801 CTGGAGGGTC TGGCCATGCC GTGGTCACTT GGTAGCCCCA GCCAAGGCCT GCTCTGTGCT
 13861 GGGCATCCCC CCGGACTGGC CCCGCACTGT CCTACCCTGG GGAAGTGTGT GTGACCCAG
 13921 CTCCTGGCCC TCCCTCTCTG GTCACCCAG CCTCCACCCC ACTCCCTGCC AGGAGGCTCA
 13981 CTGACTCCCC TCTTTCTGGG ACAGGGGACT CAGAGCTCTT CCTCATGAAG CTCATCAACC
 14041 GCCCCATCAT CGTGTTCGCG GGGGAGCATG GCTTCATCGG CTGCCGCAAG GTCACGGGCA
 14101 CCCTGGACGC CAACCGCTCC AGCTATGACG TCTTCCAGCT GGAGTTC AAC GATGGCGCCT
 14161 ACAACATCAA AGGCAGGTTT TCCTGTGGGC AGCTGCTGGG CAGGGAACCC CTCGGTCGGG
 14221 GCTGGGGTCA GTGCTGCGGG GAGCGCCCTC TGCATCCACA CTGGACCCCTG GCTTGGCTCA
 14281 GGGCCATTCC AGGCCCTAAA GGGACAGGTG TCTGATGGCC ACCAGGGGGC TCTGGGATGC
 14341 AAGCAGCCCC TTCCCTCTT GTCTGTGTGG TTGGGGGGAC TTACCTTGCC CACCTGACAG
 14401 AGAGGTGTGT GGAGGGGAGA GCAGGGAGGG AAGGAGACCA GGAAAGGGAG GGAGGAGGC
 14461 AGGGGAGGGG AGAGCCGGGA AGAGAGGAGA GCAGGGGTGG GGAGGTTTCT GGAAAGGGTG
 14521 TGCAAGGGGA GGACGCGCCT CGGTTATGGG ACTGGAGCCC CTTCCCAAGG GAGCCCCAA
 14581 CAATCCAGAG GTGCCTGTTA GGATTCAGAA CATGGTTTTT TTGTTTTTTT TTTTGAGACT
 14641 CACTCCCTCA CCCTGGCTGG ACTTGCACTG GCGTTATCTC GGCTCACTGC AACCTCTGCC
 14701 TCCTGGGTTT AAGTGATTCT CCTGCCTCAG CCTCCCAAGT AGCTGGGATT ACAGGTGTGC
 14761 ACCACCACGC CTGGCTAATT TTTATATTTT TAAATTTTAT TATTTATTTA TTTTGAGACC
 14821 CAGTCTGGAG TGCAGTGGCG TTATCTCGGC TCTCTGCAAC CTCTGCCTCC TGGGTTCAAG
 14881 CGATTCTCCT GCCTCAGCCT CCCGAGTAGC TGGGACTATG TGTGGGAGCC ACCATGCCTG
 14941 GCTAATTTTT TTGTATTTTT CATAGAGACG GGTTTCACCA TGTTGTCCAG GCTGGTCTTG
 15001 AATTCGTGGC CTCAAGTGAT CCGCCACCT CAGCCTCCCA CAGTGTGGG TTTATAGGTG
 15061 TGAGCCACCA CACCCGGCTA ATTGTTTTGT ATTTTGTAGT GAGACGGAGC TTCCTATGT
 15121 TGGCAAGGCT GGCTCGAACT CCTGACCTCA AGTGATCCGC CCACCTCAGC CTCCCAAAGT
 15181 GCTGGGATTA CAGGCGTGAG CCACTGCGGC CGAGCAGAAC ACGTTCTAGG ACCCTTGTTT
 15241 ATGTGTCCAT CATGGACAGG AGGACGTGCG GGCCATAGGG ACCCTGGCTC ATTCGGGAGC
 15301 CGGGACTGGA GGGTGGGGCG TCACCCTTGG GAACACCCGT GCCACCCCTC CGCTGCCCAG
 15361 GGTAGGGGTG GGGAGCCAGG CTTTGGGCCC CACTTGATAA AGTCCCCCTC CAGACTCCA
 15421 CAGGCAATA CTGGACGGTG GGCAGTACT CCGCGGTAC CAGCAGCGGC GACACTCCTG
 15481 TGGACTTCTT CTTGAGTTT TCGGACTATA ACAAGGTGGC CATCAAGGTG GCGGGGCGCT
 15541 ACCTGAAGGG CGACCACGCA GGCCTCCTGA AGGCCTCGGC GGAAACCGTG GACCCCGCCT
 15601 CGCTCTGGGA GTACTAGGGC CGGCCCGTCC TTCCCCGCCC CTGCCACAT GCGGGCTCCT
 15661 GCCAACCCCTC CCTGCTAACC CCTTCTCCGC CAGGTGGGCT CCAGGGCGGG AGGCAAGCCC
 15721 CCTTGCCCTT CAAACTGGAA ACCCCAGAGA AAACGGTGCC CCCACCTGTC GCCCCATAGG
 15781 ACTCCCCACT CTCCCCTCCG CCCGGGTTC CTAATCCCCT CCGGTCAGCG GCTGCGGCC
 15841 GGCCCTGGGA GGGATTTTCTG ATGCCCTGCT CCTCTTGTCT GCCACGGGGC GAGTCTGGCA
 15901 CCTCTTTCTT CTGACCTCAG ACGGCTCTGA GCCTTATTTT TCTGGAAGCG GCTAAGGGAC
 15961 GGTTGGGGGG TGGGACCCTG GCGGTGTAGT GTAACCTGGAA TCTTTTGCCT CTCCCAGCCA
 16021 CCTCCTCCCA GCCCCCAGG AGAGCTGGGC ACATGTCCCA AGCCTGTCTG TGGCCCTCCC
 16081 TGGTGCCTG TCCCCGAAAC CCCTGCTTGG GAAGGGAAGC TGTGCGGTGG GCTAGGACTG
 16141 ACCCTTGTTG TGTTTTTTTT GGTGGTGGCT GGAAACAGCC CTTCTCCAC GTGGCAGAGG
 16201 CTCAGCCTGG CTCCCTTCCC TGGAGCGGCA GGGCGTGACG GCCACAGGGT CTGCCCGCTG
 16261 CACGTTCTGC CAAGGTGGTG GTGGCGGGCG GGTAGGGGTG TGGGGGCCGT GTTCTCTCTG
 16321 TCTCTTTCTT TTCACCCTAG CCTGACTGCA AGCAGAAAAT GACCAAAATCA GTATTTTTTTT
 16381 TAATGAAATA TTATTGCTGG AGGCGTCCCA GGCAAGCCTG GCTGTAGTAG CGAGTGATCT
 16441 GCGGGGGGGC GTCTCAGCAC CCTCCCAGG GGGTGCATCT CAGCCCCCTC TTTCCGTCTT
 16501 TCCCGTCCAG CCCCAGCCCT GGGCCTGGGC TGCCGACACC TGGGCCAGAG CCCCTGCTGT
 16561 GATTGGTGCT CCCTGGGCCT CCGGGGTGGA TGAAGCCAGG CGTCGCCCCC TCCGGGAGCC
 16621 CTGGGGTGAG CCGCCGGGGC CCCCCTGCTG CCAGCCTCCC CCGTCCCCAA CATGCATCTC
 16681 ACTCTGGGTG TCTTGGTCTT TTATTTTTTT TAAGTGTCTT TTGTATAACT CTAAACGCCC
 16741 ATGATAGTAG CTTCAAACCTG GAAATAGCGA AATAAAATAA CTCAGTCTGC AGCCCCAGGC
 16801 CGGCCTGTGT GTGTCTTGGG GCTGAGGTGG GTGGGGGGGC TGAGGTGGGT GGGAGGGCTG
 16861 GCGGGACAGG TAGGCGCCCT GGCTCCCCAG CTCAGTGCTG GGAGTGTGCA GTGGGAGGGA
 16921 GGCCGTGGCT CCAGTGGGTG CTCCGGAGCT C

Fig. 9 cont.

SEQUENCE LISTING

<110> RESKE-KUNZ, A.B.
ROSS, Xiaolan
ROSS, Ralf
BROS, Matthias

<120> Regulatorische Sequenz zur spezifischen Expression in
dendritischen Zellen und deren Anwendung

<130> D 2299 PCT

<140>

<141>

<150> P 100 01 169.1-41

<151> 2000-01-13

<150> P 100 10 188.7-

<151> 2000-03-02

<160> 72

<170> PatentIn Ver. 2.1

<210> 1

<211> 3069

<212> DNA

<213> Homo sapiens

<400> 1

```
aagcttctag aaggggtata ttttgcccc aggcccaaa tcctggtctt tggactacag 60
tcaagacctt gcacagggct caggcctaaa aaaaagctgt tgaggaggggt gtgaggctgc 120
aacgttgccg tgagaagggg gtccccaggg agggcgccagc aggaagcccc aggggaagtgc 180
ctgggagagg gagggtgtgc agccagaagg agcagcccg agcctttggt tgttgactcc 240
tgctttgtaa gtggcagctc ggttgggtag gacaggggtcc gagtcctcac tcagggaact 300
gagtcagggt tgagctgagc agcccttcct ggtggcctca ctttcccttg cagagcctgc 360
tgcatgggtg ccagtggcca gcctgggagg agctcaggac tggcctcacc tcgtgccacg 420
ccctcgtagc aaggtggctc agatggcacc tgggttcccg aaggggcccag gaacacagcg 480
tccatgtccc catccttccc cggagggact tgggggtgggg cctgcaggaa ggatcatgtg 540
acttgttcaa gggcagcttg tcctgccccg gacacagagg tgcccatcgt aagcgagatg 600
caggcaaggt gaggacaggg catggtgccg agcaggaatg attttcgaaa atgctactct 660
gtggtcgggg acggtggctc actcctgtaa tccagtagt ttgggaggcc caggcgggca 720
gatcatgagg tcaggagttc aagaccagcc tggccaacat ggtgaaaccc cttctctact 780
gaaaatacaa aaaactaccc gggcgtggtg gtggatgcct gtaatcccag ctactcagga 840
ggctgaggca ggagaatcgc ttgaaccctg gaggtgaagg ttgcaatgag ccgagattgc 900
accactgcac tccgacctgg gcgacagaga aagactccgt ctgaaaaaca aacaaacaaa 960
accccgagat tctcttttcc cccgtccgga gctctatggc catctggagc tcgttctgtc 1020
cacaaggaca catttcctgg cagcagctgt ggaccagggc tcaactcactc acatctgtac 1080
ccaatctaga gcaggacaaa cacctatcac ctgcaactggc aatggacaga ggacagtggc 1140
agcccatca ccagaggcat tcaagccaag gcattttcta tcgtctatct atctatctat 1200
ctatctatct atctatctat ctatctatct atctatctag atgagatctt gctatgttgc 1260
ccaggttgaa ctctggcct caagcgatcc tcctgcttca gcctcccaaa gtgttgggat 1320
ttcaggtgtg aaccactgta cctggctgga gtgcagtggc actgtcatag ctactgcag 1380
cctccacctc ctgggctcaa ggaatcctct ttccctcagc tcctgagtag ctgggaccac 1440
aggcatgcac catcacacc agctaaacct tgatttttcc catagtatat ggctcagggc 1500
agagcaaaga agcacaacaa catggtggct tgcggttgag ggctgatggg aggggggtct 1560
ggccaaagcc caaaattaac agccacaacg tcagtgtctg tgggaggggt tgccaggggc 1620
gtgctgggtt tggggctcaa ggccctgccg gtaaaccat ttgaagcagg atggcaagaa 1680
ggtgcacacca tcttcccccg cgcactgaaa gccctggct gggattgcct ggggcagaca 1740
caggctcgga ccagccccag caatcccagt ttatcagcga gccggctgag ggcccgaggt 1800
tatctcagtg cccggcctga gaccttgtgg gcagcctgtg gtcagtctgg cattccaggg 1860
gccttttggc atgtggggaa tgtccaggaa aagcctcagc cttcggtgag gcgcagaaaa 1920
```

```

gggaagtgtc cctagagggg gtgggtgagg gcgtgggagg tgggtgtctgc aggggaatgtc 1980
cccttttggg gagggagatg gaggggttggg attctgagga tgggggggggg ggctgtagcc 2040
agcaccatgt ccctcctgtg tgaccagctc agagtcccat gaaattgggg cttggggaggg 2100
gaagggacac tggcctggga accagagacc tgggctgggtc tggctcacag ttgctggacc 2160
tctgtgatcc gtgtcaaaaa acgagaacac caattcctgt cctgcccact caccaccagg 2220
tgggacctga accctggcat cgccagcatt gggaatgtcg gccactgact caaccactct 2280
cccggagacc tatttgggcc acccgaggcg ggtgcctggg ccacacggag gggtcctggg 2340
ggtcttcagg gcagcggctg tggggctgaa gcctcaagga accacatctc tgcataggag 2400
ggccaggctg cagggcctcg ggagacaacc tagttgggcg tgttggggtc tgtgggtccc 2460
aggtcctggc ctaccgggt cccaccgcg ctgtcagctc ccagcctctt tccctcgtct 2520
gcctctgggc ttctgtaagg ctatgtgctc caaggccacc tctccaggc agccctcaga 2580
ccccacctt cgccagtag cgatctgcac cgtggtctct gaagtctcca tctgacctc 2640
aaacctcgct cgtccttgct cctagcgagg cttggggctg ggggtgtccga ggtggggggac 2700
atccgggggg gttaggtggc tggcgcgggg agccgggggt gtgaggggtg atgtcctcag 2760
gcggcgggcg tgcggggtgc ggcgaggaca ccggtggggg gagagcaccg gcggggcagc 2820
agcggggggc gcagcgccgg gtccctcggc ccggggcccc tcccgcgcgg agccaggggc 2880
gggacagggg ggcgtggcct ggtggcgctg acgtcacctc gcctataaaa tgtccggggc 2940
gccgctagct gggctttgtg gagcgctgcg gaggggtgct gcgggcccgc gcagccgaac 3000
aaaggagcag gggcgccgac gcagggaccc gccaccacc tcccggggcc gcgcagcgcc 3060
ctctcgtct 3069

```

<210> 2

<211> 1934

<212> DNA

<213> Homo sapiens

<400> 2

```

gtggcagccc catcaccaga ggcattcaag ccaaggcatt ttctatcgtc tatttatcta 60
tctatctatc tatctatcta tctatctatc tatctatcta tctagatgag atcttgctat 120
gttggccagg ttgaactcct ggcctcaagc gatcctcctg cttcagcctc ccaaagtgtt 180
gggattttcag gtgtgaacca ctgtacctgg ctggagtga gtggcactgt catagctcac 240
tgcagcctcc acctcctggg ctcaagggaat cctcttttct cagcctcctg agtagctggg 300
accacaggca tgcaccatca caccagcta aaccttgatt ttttccatag tatatggctc 360
agggcagagc aaagaagcac aaaaacatgt tggcttgagg ttgagggtg atgggagggg 420
gttctggcca aagcccaaaa ttaacagcca caacgtcagt gtctgtggga ggggttgcca 480
ggggcggtgc gggttctgggg ctcaaggccc tgccggtaaa ccattttgaa gcaggatggc 540
aagaagggtga caccatcttc ccccgcgcac tgaaagcccc tggctgggat tgcctggggc 600
agacacaggc tcggaccagc cccagcaatc ccagtttata agcagagccg ctgaggggcc 660
ggagttaatc cagtgcctgg cctgagacct tgtgggcagc ctgtgggtcat gctggcatc 720
caggggcctt ttggcattgt aggaatgtcc aggaagagcc tcagccttcg gtgaggcgca 780
gaaaaggga gtgtccctag aggggggtgg tgagggcgtg ggagggtgtg tctgcaggga 840
atgtccctt tgggggagga ggatggaggg ttgggattct gaggatgggg gggggggctg 900
tagccagcac catgtccctc ctgtgtgacc agctcagagt cccatgaaat tggggcttgg 960
gaggggaagg gacactggcc tgggaaccag agacctgggc tgggtctggc cacagttgct 1020
ggacctctgt gatccgtgtc aaaaaacgag aacaccaatt cctgtcctgc ccaactacca 1080
ccagggtggga cctgaaccct ggcatcgcca gcattgggaa tgtcggccac tgactcaacc 1140
actctcccgg agacctatct ggccacccg aggcgggtgc ctgggcccaca cggaggggtc 1200
ctggtgggtc tcagggcagc ggctgtgggg ctgaagcctc aaggaaaccac atctctgcat 1260
aggagggcca ggctgcaggg cctcgggaga caacctagtt gggcggtgtg ggggtctgtg 1320
gtcccaggtc ctggcctcac cgggtcccca ccgcgctgtc agctcccagc ctctttccct 1380
cgtctgcctc tgggcttctg taaggctatg tgctccaagg ccacctcctc caggcagccc 1440
tcagaccccc accttcgccc agtaccgatc tgacccgtgg tctctgaagt ctccatcgtg 1500
acctcaaacc tcgctcgctc ttgtccttag cgaggcttgg ggtcgggggtg tccgaggtg 1560
gggacatccg ggggggttag gtggtggcg cggggagccg ggggtgtgag ggggtgatgc 1620
ctcaggcggc ggcgtgcgg ggtgcggcga ggacaccgtt ggggtgagag caccggcggg 1680
gcagcagcgg gggcgcgagc gccgggtccc tcggcccggg gccctcccg cgcggagcca 1740
ggggcgggac aggggggcgt ggcctgggtg cgctgacgtc acctcgcta taaaatgtcc 1800
ggggcgccgc tagctgggct ttgtggagcg ctgcggaggg tgcgtgcggg ccgcggcagc 1860
cgaacaaagg agcagggggc ccgcccagc gacccgccac ccacctccc gggcgcgcca 1920
gcggcctctc gtct 1934

```

<210> 3
 <211> 1619
 <212> DNA
 <213> Homo sapiens

<400> 3

```

catcacaccc agctaaacct tgattttttc catagtatat ggctcagggc agagcaaaga 60
agcacaacaaa catgtttggct tgcggttgag ggctgatggg aggggggttct ggccaaagcc 120
caaaattaac agccacaacg tcagtgtctg tgggaggggt tgccaggggc gtgcgggttc 180
tggggctcaa ggccctgccg gtaaaccat ttgaagcagg atggcaagaa ggtgacacca 240
tcttcccccg cgcactgaaa gcccctggct gggattgcct ggggcagaca caggctcgga 300
ccagccccag caatcccagt ttatcagcga gccggctgag ggcccgaggt tatctcagtg 360
cccggcctga gaccttgttg gcagcctgtg gtcattgctgg cattccaggg gccttttggc 420
atgtggggaa tgtccaggaa aagcctcagc cttcgggtgag gcgcagaaaa gggaagtgtc 480
cctagagggg gtgggtgagg gcgtgggagg tgggtgtctgc agggaaatgtc ccctttgggg 540
gaggaggatg gagggttggg attctgagga tggggggggg ggctgtagcc agcaccatgt 600
ccctcctgtg tgaccagctc agagtcccat gaaattgggg cttgggaggg gaaggacac 660
tggcctggga accagagacc tgggctggtc tggctcacag ttgctggacc tctgtgatcc 720
gtgtcaaaaa acgagaacac caattcctgt cctgcccact caccaccagg tgggacctga 780
accctggcat cgccagcatt gggaatgtcg gccactgact caaccactct cccggagacc 840
tatttgggcc acccgaggcg ggtgcctggg ccacacggag gggctcctggg ggtcttcagg 900
gcagcggctg tggggctgaa gcctcaagga accacatctc tgcataggag ggccaggctg 960
cagggcctcg ggagacaacc tagttgggag tgttggggtc tgtgggtccc aggtcctggc 1020
ctcaccgggt cccacccgag ctgtcagctc ccagcctctt tccctcgtct gcctctgggc 1080
ttctgtaagg ctatgtgctc caaggccacc tcctccaggc agccctcaga ccccccactt 1140
cgccagctac cgatctgcac cgtggtctct gaagtctcca tcgtgacctc aaacctcgct 1200
cgtccttgct cctagcgagg cttgggggtcg ggggtgtccga ggtggggggac atccgggggg 1260
gttaggtggc tggcgcgggg agccgggggt gtgaggggtg atgtcctcag gcggcgggcg 1320
tgcggggtgc ggcgaggaca ccgggtgggg gagagcaccg gcgggggcagc agcgggggcc 1380
gcagcgccgg gtccctcggc ccggggcccc tcccgcgagg agccaggggc gggacagggg 1440
ggcgtggcct ggtggcgctg acgtcacctc gcctataaaa tgtccggggc gccgctagct 1500
gggctttgtg gagcgctgcg gaggggtgct gcggggcgcg gcagccgaac aaaggagcag 1560
gggcgcgcgc gcaggggacct gccaccacc tcccggggccc gcgcagcggc ctctcgtct 1619

```

<210> 4
 <211> 1449
 <212> DNA
 <213> Homo sapiens

<400> 4

```

gtgcgggttc tggggctcaa ggccctgccg gtaaaccat ttgaagcagg atggcaagaa 60
ggtgacacca tcttcccccg cgcactgaaa gcccctggct gggattgcct ggggcagaca 120
caggetcgga ccagccccag caatcccagt ttatcagcga gccggctgag ggcccgaggt 180
tatctcagtg cccggcctga gaccttgttg gcagcctgtg gtcattgctgg cattccaggg 240
gccttttggc atgtggggaa tgtccaggaa aagcctcagc cttcgggtgag gcgcagaaaa 300
gggaagtgtc cctagagggg gtgggtgagg gcgtgggagg tgggtgtctgc agggaaatgtc 360
ccctttgggg gaggaggatg gaggggtggg attctgagga tggggggggg ggctgtagcc 420
agcaccatgt ccctcctgtg tgaccagctc agagtcccat gaaattgggg cttgggaggg 480
gaaggggacac tggcctggga accagagacc tgggctggtc tggctcacag ttgctggacc 540
tctgtgatcc gtgtcaaaaa acgagaacac caattcctgt cctgcccact caccaccagg 600
tgggacctga accctggcat cgccagcatt gggaatgtcg gccactgact caaccactct 660
cccgagacc tatttgggcc acccgaggcg ggtgcctggg ccacacggag gggctcctgg 720
ggtcttcagg gcagcggctg tggggctgaa gcctcaagga accacatctc tgcataggag 780
ggccaggctg cagggcctcg ggagacaacc tagttgggag tgttggggtc tgtgggtccc 840
aggtcctggc ctcaccgggt cccacccgag ctgtcagctc ccagcctctt tccctcgtct 900
gcctctgggc ttctgtaagg ctatgtgctc caaggccacc tcctccaggc agccctcaga 960
ccccacactt cgccagctac cgatctgcac cgtggtctct gaagtctcca tcgtgacctc 1020
aaacctcgct cgtccttgct cctagcgagg cttgggggtcg ggggtgtccga ggtggggggac 1080
atccgggggg gttaggtggc tggcgcgggg agccgggggt gtgaggggtg atgtcctcag 1140
gcggcgggcg tgcgggggtg ggcgaggaca ccgggtgggg gagagcaccg gcgggggcagc 1200
agcgggggcc gcagcgccgg gtccctcggc ccggggcccc tcccgcgagg agccaggggc 1260
gggacagggg ggcgtggcct ggtggcgctg acgtcacctc gcctataaaa tgtccggggc 1320

```

```

gccgctagct gggctttgtg gagcgctgcg gaggggtgcgt gcgggcccgc gcagccgaac 1380
aaaggagcag gggcgccgcc gcagggaccc gccacccacc tcccggggcc gcgcagcggc 1440
ctctcgtct 1449

```

```

<210> 5
<211> 1240
<212> DNA
<213> Homo sapiens

```

```

<400> 5
ggcagcctgt ggtcatgctg gcattccagg ggccttttgg catgtgggga atgtccagga 60
aaagcctcag ccttcggtga ggcgagaaaa agggaaagtgt ccctagagggg ggtgggtgag 120
ggcgtgggag gtggtgtctg cagggaaatgt ccccttttggg ggaggaggat ggagggttgg 180
gattctgagg atgggggggg gggctgtagc cagcaccatg tccctcctgt gtgaccagct 240
cagagtccca tgaaattggg gcttgggagg ggaagggaca ctggcctggg aaccagagac 300
ctgggctggg ctggctcaca gttgctggac ctctgtgatc cgtgtcaaaa aacgagaaca 360
ccaattcctg tcctgcccac tcaccaccag gtgggacctg aaccttggca tcgccagcat 420
tggaatgtc ggccactgac tcaaccactc tcccggagac ctatttgggc cacccgaggc 480
gggtgcctgg gccacacgga ggggtcctgg tggctctcag ggcagcggct gtggggctga 540
agcctcaagg aaccacatct ctgcatagga gggccaggct gcagggcctc gggagacaac 600
ctagttgggc gtgttggggt ctgtgggtcc caggctcctgg cctcaccggg tccccaccgc 660
gctgtcagct cccagcctct ttccctcgtc tgctctggg cttctgtaag gctatgtgct 720
ccaaggccac ctctccagg cagccctcag acccccacct tcgccagta ccgatctgca 780
ccgtggtctc tgaagtctcc atcgtgacct caaacctcgc tcgtccttgc tcctagcgag 840
gcttgggggtc ggggtgtccg aggtggggga catccggggc ctgcccgggtg cggcgaggac 960
gagccggggg tgtgaggggt gatgtcctca ggccggcgcg ctgcccgggtg cggcgaggac 960
accgtggggg tgagagcacc ggccggggcag cagcgggggc cgcagcgccg ggtccctcgg 1020
cccggggccc ctcccgcgcg gagccagggg cgggacaggg gggcgtggcc tgggtggcgct 1080
gacgtcacct cgcctataaa atgtccgggg cgcgcgtagc tgggctttgt ggagcgctgc 1140
ggagggtgcg tgcgggcgcg ggcagccgaa caaaggagca ggggcgcgcg cgcagggacc 1200
cgccaccac ctcccggggc cgcgcagcgg cctctcgtct 1240

```

```

<210> 6
<211> 943
<212> DNA
<213> Homo sapiens

```

```

<400> 6
gacctgggct ggtctggctc acagttgctg gacctctgtg atccgtgtca aaaaacgaga 60
acaccaattc ctgtcctgcc cactcaccac cagggtgggac ctgaaccctg gcatcgccag 120
cattgggaat gtcgccact gactcaacca ctctcccga gacctatttg ggccaccga 180
ggcgggtgcc tgggccacac ggagggttcc tgggtggtctt cagggcagcg gctgtggggc 240
tgaagcctca aggaaccaca tctctgcata ggagggccag gctgcagggc ctcgggagac 300
aacctagttg ggcgtgttgg ggtctgtggg tcccagggtc tggcctcacc ggggtccccc 360
cgcgctgtca gctcccagcc tctttccctc gtctgctctt gggcttctgt aaggctatgt 420
gctccaaggc cactcctcc aggcagccct cagaccccca ccttcgcca gtaccgatct 480
gcaccgtggg ctctgaagtc tccatcgtga cctcaaacct cgtcgtcct tgctcctagc 540
gaggcttggg gtcggggtgt ccgaggtggg ggacatccgg ggggggttagg tggctggcgc 600
ggggagccgg ggttgtgagg ggtgatgtcc tcaggcggcg gcgctgcggg gtgcggcgag 660
gacaccggtg ggttgagagc accggcgggg cagcagcggg ggcgcagcg ccgggtccct 720
cggcccgggg cccctcccgc gcggagccag gggcgggaca ggggggcgtg gcttgggtgc 780
gctgacgtca cctcgcctat aaaatgtccg gggcgccgct agctgggctt tgtggagcgc 840
tgcgagggt gcgtgcgggc cgcggcagcc gaacaaagga gcaggggcgc cgcgcaggg 900
accgcaccac cactcccgg ggccgcgcag cggcctctcg tct 943

```

```

<210> 7
<211> 660
<212> DNA
<213> Homo sapiens

```

<400> 7

```
gcagggcctc gggagacaac ctagttgggc gtgttgggggt ctgtgggtcc caggtcctgg 60
cctcaccggg tccccaccgc gctgtcagct cccagcctct ttccctcgtc tgcctctggg 120
cttctgtaag gctatgtgct ccaaggccac ctccctccagg cagccctcag acccccacct 180
tcgcccagta ccgatctgca ccgtggtctc tgaagtctcc atcgtgacct caaacctcgc 240
tcgtccttgc tcctagcgag gcttgggggtc ggggtgtccg aggtggggga catccggggg 300
ggttaggtgg ctggcgcggg gagccgggggt tgtgaggggt gatgtcctca ggcggcgggc 360
ctgcggggtg cggcgaggac accggtgggg tgagagcacc ggcggggcag cagcgggggc 420
cgcagcgccg ggteccctcg cccggggccc ctcccgcgcg gagccagggg cgggacaggg 480
gggctgtggc tgggtggcgt gacgtcacct cgcctataaa atgtccgggg cgcgctagc 540
tgggctttgt ggagcgctgc ggaggggtgc tgcggggccg ggcagccgaa caaaggagca 600
ggggcgccgc cgcagggacc cgccaccacac ctcccggggc cgcgcagcgg cctctcgtct 660
```

<210> 8

<211> 370

<212> DNA

<213> Homo sapiens

<400> 8

```
catccggggg ggttaggtgg ctggcgcggg gagccgggggt tgtgaggggt gatgtcctca 60
ggcggcgggc ctgcggggtg cggcgaggac accggtgggg tgagagcacc ggcggggcag 120
cagcgggggc cgcagcgccg ggtccctcgg cccggggccc ctcccgcgcg gagccagggg 180
cgggacaggg gggcgtggcc tgggtggcgt gacgtcacct cgcctataaa atgtccgggg 240
cgccgctagc tgggctttgt ggagcgctgc ggaggggtgc tgcggggccg ggcagccgaa 300
caaaggagca ggggcgcgcg cgcagggacc cgccaccacac ctcccggggc cgcgcagcgg 360
cctctcgtct 370
```

<210> 9

<211> 191

<212> DNA

<213> Homo sapiens

<400> 9

```
gtgagtgggg acgctgcccc cgcctctcct ggtccgtgca caaagcgcac cccaccgcgc 60
cccctccagc ctcccgccct ttctcgctcg cggcgccgct gcggtccgga gcactgcccc 120
ttgcgcccc gctaggcacg cgggctaccc cgcctggagg gggcgaggag tggggctttg 180
cccatccgcg g 191
```

<210> 10

<211> 1896

<212> DNA

<213> Homo sapiens

<400> 10

```
gtgtcagccc ttccccccag cccctcctcc ggcgtctgcc ccgcgctcga ggccgcggcc 60
tttgtgagca ggggggcggg tcgcctcgac tgggtcctcc ctgcgccctt ttgtcctgct 120
ccatctctgc agatgggaaa accagatgcg gcggggcggg ggaggggatc ggcttttgcg 180
gttcacccct gcagaggagc ccccgcgccg gcccccgggc ccagggtcgt ggtcccaggg 240
ttccccagga ggcggtctcc ctccctcgtc ccgcggcccg ggaacggcgt ggcgcggatg 300
gcggccctcc aggcaccccg cccttcgccc gccgcgcgcg tctccaggcc ggtgcgctga 360
gctccgctcc gggggcgacc gagggcggtc tcagcgcgag ccgggagacc ccaggccagc 420
tccctcggga agccctacc ctctggtgaa cccatccctt agggcgctcg ccgggaacaa 480
ttgactgcaa ctcgatccgt cccctccggg gcctcctgca ggcacggcta tttgcgaccg 540
gcctgtgcgc cactcttctc ctcttttctc agattctccc cgaggggctt tctcttcctt 600
ttggcgtcct cttccactcc tggcgagaaa gctgttcctt gcttgacagc agaggccccg 660
ctggaggagag gggcgtgggg ggatcttttt ccctggatag gaagtccgac ccaggccctt 720
ggagatggct ggacgggagc ccagtttgtg tcctcatccc tctcttgga aggagccctc 780
ctgacggccc cctccttggc tgtgggctct gcaggagggg gaaagacccc ccatggcagt 840
gggggtgagg gtggaggcct ggggtggggt atggccattt tgtgcctgaa gtttgctacc 900
ttaagtccca gagaggtgaa gctgcttgtc atgggtcaca cagcaagtga ccacaaggaa 960
```

```

gcagcatgca gaagtgggtg catttggctc cagaaacccc gtttcctgtt tcccacagcg 1020
catggcgggc agagctcctc acctgatcag caggcattga gccctactaa ggggacctgc 1080
tgagccatga gtgaaccagt ggggttcacg ttctcttggg gacattacga agtgtcaagg 1140
aaaaggcagc gggtagcagg tattagatgg ctgagtcctt tctgagcagg acatttgagc 1200
tggcctgaag gatgaataga aggtggcagg tatggggggg gtgctgtgtg gaccagcggt 1260
gtaggctgtg ggtgtctcct gtgcaaaggc cttatggcag gaaggtagag agaaagccca 1320
gtgtggctgg gaggaagcct gtgaggtctg ggccacggga ccttgggtga gagcttagtg 1380
cagggctatg gtccttgccct tgggggttgca gccacagcct cccgtgggga agaagtcgca 1440
gaagtaggtg gccttggccc tggggggccag gaagttgcag agagcctgag ggtgtttttg 1500
gcgagcctgg gcagataact cccctctcct gagaagcctg ctggggggcg ctggtgtgtg 1560
attcaggctg ataggggtgg aggaaccaga aattgcagcc aggagagggt gtggtgatgg 1620
ctcagcttgg agttgaaagg gatgtgggac ggcgcggggg tgggggtact tgggggccga 1680
ggttggggct cctctggcct cagatcttgg ggtccttata tcgtttctat gtcagccaac 1740
tcccgggggg cttttgggaa ctggagtctg caggggtggg cgtggggggc gtgtctgcag 1800
aggccatttt ggacgggcct gcgtggcaag ggggaggcgt ctgcttccga cacacatgct 1860
ggggaacgcc agccaggaag gggaggatcc gagctc 1896

```

<210> 11

<211> 356

<212> DNA

<213> Homo sapiens

<400> 11

```

ggggccgggtg ggggcagggg gctccccttt gaaaaacgct gcgccctgct cttcagcctc 60
agtttcccca tctgtaaagt gtgcattcca ggcctccttg ggctggccca gagctgccct 120
gggcagggggc tttcagccct tcacaccaga ggctgggtca gctgagttag tgctcctttg 180
tcctcccgcc atgcccaggg ctctccctg ctcccagga tccacaccta ccctggcccc 240
gaccctgggc catgcctcac cacacttct gtcttctctc acatctctca catctgggag 300
tcctctcccg ccagcctgtg cttgcatctg gggatccatg ccagggatag gacctg 356

```

<210> 12

<211> 550

<212> DNA

<213> Homo sapiens

<400> 12

```

ctgcagttag ccgagatcgt gccactgcac tccagcctgg gcaacagAAC gagaccctgt 60
cttaaaaaaa ttttttttgg cctggtgogg tggctcacgc ctgtaatgcc agcacttttg 120
gagtccaagg tgggtggatc acctggagggt cgggagtttg agaccagtct gaccaacatg 180
gagaaacccc atctctacta aaaacacaaa aattagccgg gcgtggtggc gcatgcctgt 240
aatcccagct actctggagg ctgaggcagg agaatccctt gaaccgggga ggcggaagtt 300
gcagtgagcc gagatcgac cattgcactc cagcctgggc aacaagagtg aaactccatc 360
tcaaaaaaca aaacaaaaca aaaaaaaca caaaaacaa caaacaaaaa actgtaatta 420
aaataattaa aagaaaaaag ttaaaaaaaa aaaaaagaga ggatggctca aaggctggaa 480
acagggaagg ccactgtgag gggaggacag gcagggctag acctctctgg aaggaggagg 540
gagaggtagc 550

```

<210> 13

<211> 1426

<212> DNA

<213> Homo sapiens

<400> 13

```

gggactgagt cgtggtacag atggcggcag ttgtgtggcc ctggggccca ggaatggctg 60
gggaagctccc tttcttctct ggcttggggg atgaggttag tgtaaatctc atggagttgc 120
caggaggatt aaggccgacg tgcattcaca cgtggccccg tggtgccgc cagttaagcc 180
ctgocgagcc attggaagtc tcattgagaa ggagcctccc ggcttccctg atttgttcca 240
ggcgggctgg taagcgcctt gcttatttgc aaggctgtgt gaggacgagc cctgtaaacc 300
ctaggacgag gctgtgctaa tgtggtttct gttccactgt cttcccttct ctttttgccc 360
ctctttctgg ggaacctgga gccctgatg cctctgttt tttcctgctg ctgggggtgga 420

```

```

cctagaacct ctccctccat cctttctgct gctgggggtgg gcctgcagaa gactgtcctc 480
ttacctgagt ccctaacttc cctccgatca gcggggggctt cagcggagcc ccagcgcat 540
ggacaggtgg tgctctgtgg ccggcacagg ccagggccca ggggtgcggcc cctgcctgta 600
ggtagacatg caggcccttc tgagcagaat gagggggcgg tgagggcagc tgtgcgggtg 660
tggctggcct gacggctccc tgcagccctg cggcttctgt gcagtggggg tggggcgggc 720
cagacctttg cagggcctct tggctggtgg agggctgagt gagcagaggc cccagccctc 780
gtgttccctg ggggtgtggc ttaggatggg cccggcaccc tggggaccca gccctgtctc 840
accttatcct cctcccgtgc ttggctgggg tgtggtggct gagccagccc acccagtgcg 900
gtggatgctc gtcggagggc agagggtggg ctaccggctt aaggggcccc agggaaacctg 960
gggggtggag cccaaggggt tccaagaagg ggcggggcca ggaaccata gacaagaggg 1020
tggggcaggg aaagggcgtg gcaagagggc caccacctc cgggccagag agccagccag 1080
accttactt tctcttgctg tgtgacctta ggcaaggaca tgccaccacc ggcttagtc 1140
tccctctttg tgaagtagga tgaagatccc cacctgtgga gcagatgtgg ggctgcatgg 1200
ttgggtcagg atggaacagg atggggaggc cgggcgtggg ggcttaccct tgtaatccca 1260
gcactttggg agggaaagg gggaggactg cttgagtcct ggagtttggg caacatagcg 1320
agaccacccc catctctaca aaataatgtt aaagttagcc aggtgtagtg gtgagtgctc 1380
gtggtcccac ctactgggga gactgaggca agaggatcct ttgagc 1426

```

<210> 14

<211> 477

<212> DNA

<213> Homo sapiens

<400> 14

```

ctgcagcctc cgcctcctgg gttcaagcaa ttctcctgcg tcagcctccc gagcagctgg 60
gattacaggg ccccggcacc acccccagct aatttttttt tttttgagac agagtctcgc 120
attgtcgccc aggtctggagt gcagtggcgc gatcttggct cactgcaagc tccacctccc 180
gggttcacgc cattctcctg cctcagcctc ccgagtagca gggactacag gcgcccgcga 240
ccgtgcccgg ctaatttttt tttttgtat tattaataga gacgggtttc accgtgttag 300
gatgatctcg atctcctgac ctctgatcc gcctgcctcg gcctcccaa gtgctgggat 360
tacaggcatg agccaccttg cctggctcct tttttgtat ttttagtaga gatgaggttt 420
caccatgttg gccaggtcgt tctcaaactc ctgacctcat gatccgccc cctcggc 477

```

<210> 15

<211> 473

<212> DNA

<213> Homo sapiens

<400> 15

```

ctgcagaaca cccctgtggt ggcaaataca tctctttctt tttgattcag ggtcttgctc 60
tggtgcccag gttggagtgc agtggcaca tcatagctca ctgtagcctc gacttctcgg 120
gctcaagtga tcctcccacc tcagcctccc aaatagctga gactacaggc acacaccacc 180
acatcaagtt aatttttttt tctatttggt acagacaggg tctggctatg ttgcctagga 240
tggtcttgaa ctcccggcct cacatgatcc tcctaactca gcctcccaa gtgctgaggt 300
tacagacaca agccaccatg tgcccagcca gtgaaccac aactttctag ctgttttttt 360
cttatacata agaagcatgc ggggtgccac ggcggctcac gcctgtaatc ccagcacttt 420
gggaagctga ggcgggtgga tcacctgagg tcatgagttc aagaccagcc ttg 473

```

<210> 16

<211> 311

<212> DNA

<213> Homo sapiens

<400> 16

```

caatcctcct gccttggggc tcccaaagtg ttgcgatcac aggcattgagc cctgtcctc 60
ggcccttatt atttctctaa ctggggaaag tcatgtggga acagatgtag cttgccttgg 120
cctctgaccg gccctgcctg cgttctctgg tgctctctgc tgcttctcat gtgtgccact 180
gtggggactc ggccgcccac cccacccctg ggtgttacct tgcgtgtgta gttctgtgag 240
ctcagggcta tggtctgcc aactagggg gcgtggggcc ccagtaccag cccaaggcct 300
cctctctgca g
311

```

<210> 17
<211> 82
<212> DNA
<213> Homo sapiens

<400> 17
gtgagtgcct cgctcccacc tgtcaccgcc cccaccacct tgctggggct accccgcctg 60
accctgtccc gccatcccc ag 82

<210> 18
<211> 245
<212> DNA
<213> Homo sapiens

<400> 18
gtaacactaa agccccagtt ccctggagcc gtcctggagt cctggagggt ctggccatgc 60
cgtgggtcact tggtagcccc agccaaggcc tgctctgtgc tgggcatccc cccggactgg 120
ccccgcactg tcctaccctg gggactgctg tgtgacccca gtcctgggcc ctccctctct 180
ggtcacccca gcctccaccc cactccctgc caggaggctc actgactccc ctctttctgg 240
gacag 245

<210> 19
<211> 1242
<212> DNA
<213> Homo sapiens

<400> 19
gcaggttctc ctgtgggagc ctgctgggca gggaaacctt cgggtcggggc tggggtcagt 60
gctgcgggga ggcacctctg catccacact ggaccttggc ttgggtcagg gccattccag 120
gccctaaagg gacaggtgtc tgatggccac caggggggctc tgggatgcaa gcagccccctt 180
tccctcttgt ctgtgtgggt ggggggactt accttgcaca cctgacagag aggtgtgtgg 240
aggggagagc agggaggga aaagggaggg agggagagcag gggaggggag 300
agccgggaag agaggagagc aggggtgggg aggtttctgg aaagggtgtg caaggggagg 360
acgcgcctcg gttatgggac tggagccctt tccaggagg accccaaca atccagagg 420
gcctgttagg attcagaaca tgggtttttt gtttgtttt ttgagactca ctccctcacc 480
ctggctggac ttgcagtggc gttatctcgg ctactgcaa cctctgcctc ctgggttcaa 540
gtgattctcc tgcctcagcc tcccaagtag ctgggattac aggtgtgcac caccacgcct 600
ggctaatttt tataatttta aaatttatta tttatttatt ttgagacca gtctggagt 660
cagtggcggt atctcggtc tctgcaacct ctgcctcctg gggtcaagcg attctcctgc 720
ctcagcctcc cgagtagctg ggactatgtg tgggagccac catgcctggc taattttttt 780
gtatttttca tagagacggg tttcaccatg ttgtccaggc tggctctgaa ttctgtggcct 840
caagtgatcc gccacactca gcctcccaca gtgctgggtt tataaggtgtg agccaccaca 900
cccggctaatt tgttttgtat ttttagtaga gacggagctt cactatgttg gcaaggctgg 960
ctcgaaactcc tgacctcaag tgatccgccc acctcagcct cccaaagtgc tgggattaca 1020
ggcgtgagcc actgcggccg agcagaacac gttctaggac ccttgttcat gtgtccatca 1080
tggacaggag gacgtgcggg ccatagggac cctggctcat tccggagccg ggactggagg 1140
gtggggcgtc acccttggga acaccctgc ccaccctccg ctgccaggg taggggtggg 1200
gagccaggct ttgggcccc cttgataaag tccctcccc ag 1242

<210> 20
<211> 161
<212> DNA
<213> Homo sapiens

<400> 20
agccccaggc cggcctgtgt gtgtcttggg gctgaggtgg gtgggggggc tgagggtgggt 60
gggagggtct gcgggacagg taggcgccct ggctccccag ctgagtgtg ggagtgtgca 120
gtgggaggga ggccgtggct ccagtgggtg ctccggagct c 161

<210> 21
<211> 211
<212> DNA
<213> Homo sapiens

<400> 21
cctcccgcgc ggagccaggg gcgggacagg ggggcgtggc ctggtggcgc tgacgtcacc 60
tcgcctataa aatgtccggg gcgccgctag ctgggctttg tggagcgctg cggaggggtgc 120
gtgcggggccg cggcagccga acaaaggagc aggggcgccg ccgcagggac ccgccaccca 180
cctcccgggg ccgcgcagcg gcctctcgtc t 211

<210> 22
<211> 155
<212> DNA
<213> Homo sapiens

<400> 22
cacctcgcct ataaaatgtc cggggcgccg ctagctgggc tttgtggagc gctgcggagg 60
gtgcgtgcgg gccgcggcag ccgaacaaag gagcaggggc gccgccgcag ggaccccgcca 120
cccacctccc ggggcgcgc agcggcctct cgtct 155

<210> 23
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 23
gctgacctgc tggattacat 20

<210> 24
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 24
cattatagtc aaggcatat cc 22

<210> 25
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 25
ggcaagcctg gctgtagtag 20

<210> 26
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: artificial
 sequence

<400> 26
 ccagagtgag atgcatgttg g 21

<210> 27
 <211> 2772
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (112)..(1590)

<400> 27
 gcggagggtg cgtgcgggcc gcggcagccg aacaaaggag caggggcgcc gccgcaggga 60
 cccgccaccc acctcccggg gccgcgcagc ggccctctcgt ctactgccac c atg acc 117
 Met Thr
 1
 gcc aac ggc aca gcc gag gcg gtg cag atc cag ttc ggc ctc atc aac 165
 Ala Asn Gly Thr Ala Glu Ala Val Gln Ile Gln Phe Gly Leu Ile Asn
 5 10 15
 tgc ggc aac aag tac ctg acg gcc gag gcg ttc ggg ttc aag gtg aac 213
 Cys Gly Asn Lys Tyr Leu Thr Ala Glu Ala Phe Gly Phe Lys Val Asn
 20 25 30
 gcg tcc gcc agc agc ctg aag aag aag cag atc tgg acg ctg gag cag 261
 Ala Ser Ala Ser Ser Leu Lys Lys Lys Gln Ile Trp Thr Leu Glu Gln
 35 40 45 50
 ccc cct gac gag gcg ggc agc gcg gcc gtg tgc ctg cgc agc cac ctg 309
 Pro Pro Asp Glu Ala Gly Ser Ala Ala Val Cys Leu Arg Ser His Leu
 55 60 65
 ggc cgc tac ctg gcg gcg gac aag gac ggc aac gtg acc tgc gag cgc 357
 Gly Arg Tyr Leu Ala Ala Asp Lys Asp Gly Asn Val Thr Cys Glu Arg
 70 75 80
 gag gtg ccc ggt ccc gac tgc cgt ttc ctc atc gtg gcg cac gac gac 405
 Glu Val Pro Gly Pro Asp Cys Arg Phe Leu Ile Val Ala His Asp Asp
 85 90 95
 ggt cgc tgg tcg ctg cag tcc gag gcg cac cgg cgc tac ttc ggc ggc 453
 Gly Arg Trp Ser Leu Gln Ser Glu Ala His Arg Arg Tyr Phe Gly Gly
 100 105 110
 acc gag gac cgc ctg tcc tgc ttc gcg cag acg gtg tcc ccc gcc gag 501
 Thr Glu Asp Arg Leu Ser Cys Phe Ala Gln Thr Val Ser Pro Ala Glu
 115 120 125 130
 aag tgg agc gtg cac atc gcc atg cac cct cag gtc aac atc tac agt 549

Lys Trp Ser Val His Ile Ala Met His Pro Gln Val Asn Ile Tyr Ser
 135 140 145

gtc acc cgt aag cgc tac gcg cac ctg agc gcg cgg ccg gcc gac gag 597
 Val Thr Arg Lys Arg Tyr Ala His Leu Ser Ala Arg Pro Ala Asp Glu
 150 155 160

atc gcc gtg gac cgc gac gtg ccc tgg ggc gtc gac tcg ctc atc acc 645
 Ile Ala Val Asp Arg Asp Val Pro Trp Gly Val Asp Ser Leu Ile Thr
 165 170 175

ctc gcc ttc cag gac cag cgc tac agc gtg cag acc gcc gac cac cgc 693
 Leu Ala Phe Gln Asp Gln Arg Tyr Ser Val Gln Thr Ala Asp His Arg
 180 185 190

ttc ctg cgc cac gac ggg cgc ctg gtg gcg cgc ccc gag ccg gcc act 741
 Phe Leu Arg His Asp Gly Arg Leu Val Ala Arg Pro Glu Pro Ala Thr
 195 200 205 210

ggc tac acg ctg gag ttc cgc tcc ggc aag gtg gcc ttc cgc gac tgc 789
 Gly Tyr Thr Leu Glu Phe Arg Ser Gly Lys Val Ala Phe Arg Asp Cys
 215 220 225

gag ggc cgt tac ctg gcg ccg tcg ggg ccc agc ggc acg ctc aag gcg 837
 Glu Gly Arg Tyr Leu Ala Pro Ser Gly Pro Ser Gly Thr Leu Lys Ala
 230 235 240

ggc aag gcc acc aag gtg ggc aag gac gag ctc ttt gct ctg gag cag 885
 Gly Lys Ala Thr Lys Val Gly Lys Asp Glu Leu Phe Ala Leu Glu Gln
 245 250 255

agc tgc gcc cag gtc gtg ctg cag gcg gcc aac gag agg aac gtg tcc 933
 Ser Cys Ala Gln Val Val Leu Gln Ala Ala Asn Glu Arg Asn Val Ser
 260 265 270

acg cgc cag ggt atg gac ctg tct gcc aat cag gac gag gag acc gac 981
 Thr Arg Gln Gly Met Asp Leu Ser Ala Asn Gln Asp Glu Glu Thr Asp
 275 280 285 290

cag gag acc ttc cag ctg gag atc gac cgc gac acc aaa aag tgt gcc 1029
 Gln Glu Thr Phe Gln Leu Glu Ile Asp Arg Asp Thr Lys Lys Cys Ala
 295 300 305

ttc cgt acc cac acg ggc aag tac tgg acg ctg acg gcc acc ggg ggc 1077
 Phe Arg Thr His Thr Gly Lys Tyr Trp Thr Leu Thr Ala Thr Gly Gly
 310 315 320

gtg cag tcc acc gcc tcc agc aag aat gcc agc tgc tac ttt gac atc 1125
 Val Gln Ser Thr Ala Ser Ser Lys Asn Ala Ser Cys Tyr Phe Asp Ile
 325 330 335

gag tgg cgt gac cgg cgc atc aca ctg agg gcg tcc aat ggc aag ttt 1173
 Glu Trp Arg Asp Arg Arg Ile Thr Leu Arg Ala Ser Asn Gly Lys Phe
 340 345 350

gtg acc tcc aag aag aat ggg cag ctg gcc gcc tcg gtg gag aca gca 1221
 Val Thr Ser Lys Lys Asn Gly Gln Leu Ala Ala Ser Val Glu Thr Ala
 355 360 365 370

ggg gac tca gag ctc ttc ctc atg aag ctc atc aac cgc ccc atc atc 1269
 Gly Asp Ser Glu Leu Phe Leu Met Lys Leu Ile Asn Arg Pro Ile Ile
 375 380 385

gtg ttc cgc ggg gag cat ggc ttc atc ggc tgc cgc aag gtc acg ggc 1317
 Val Phe Arg Gly Glu His Gly Phe Ile Gly Cys Arg Lys Val Thr Gly
 390 395 400

acc ctg gac gcc aac cgc tcc agc tat gac gtc ttc cag ctg gag ttc 1365
 Thr Leu Asp Ala Asn Arg Ser Ser Tyr Asp Val Phe Gln Leu Glu Phe
 405 410 415

aac gat ggc gcc tac aac atc aaa gac tcc aca ggc aaa tac tgg acg 1413
 Asn Asp Gly Ala Tyr Asn Ile Lys Asp Ser Thr Gly Lys Tyr Trp Thr
 420 425 430

gtg ggc agt gac tcc gcg gtc acc agc agc ggc gac act cct gtg gac 1461
 Val Gly Ser Asp Ser Ala Val Thr Ser Ser Gly Asp Thr Pro Val Asp
 435 440 445 450

ttc ttc ttc gag ttc tgc gac tat aac aag gtg gcc atc aag gtg ggc 1509
 Phe Phe Phe Glu Phe Cys Asp Tyr Asn Lys Val Ala Ile Lys Val Gly
 455 460 465

ggg cgc tac ctg aag ggc gac cac gca ggc gtc ctg aag gcc tcg gcg 1557
 Gly Arg Tyr Leu Lys Gly Asp His Ala Gly Val Leu Lys Ala Ser Ala
 470 475 480

gaa acc gtg gac ccc gcc tcg ctc tgg gag tac tagggccggc ccgtccttcc 1610
 Glu Thr Val Asp Pro Ala Ser Leu Trp Glu Tyr
 485 490

ccgccccctgc ccacatggcg gctcctgccca accctccctg ctaacccctt ctccgccagg 1670

tgggctccag ggcgggaggc aagccccctt gcctttcaaa ctggaaaccc cagagaaaac 1730

ggtgccccca cctgtcgccc ctatggactc cccactctcc cctccgcccg ggttccctac 1790

tccccctggg tcagcggtg cggcctggcc ctgggagggg tttcagatgc cctgccctc 1850

ttgtctgccca cggggcgagt ctggcacctc tttctttctga cctcagacgg ctctgagcct 1910

tattttctctg gaagcggcta agggacgggt gggggctggg agccctgggc gtgtagtgta 1970

actggaatct tttgcctctc ccagccacct cctcccagcc ccccaggaga gctgggcaca 2030

tgtcccaagc ctgtcagtgg ccctccctgg tgcaactgtcc ccgaaacccc tgcttgaggaa 2090

ggggaagctgt cgggagggct aggactgacc cttgtggtgt ttttttgggt ggtggctgga 2150

aacagcccct ctcccacgtg ggagaggctc agcctggctc ccttccctgg agcggcaggg 2210

cgtgacggcc acagggctctg cccgctgcac gttctgccaa ggtggtggtg gcgggcgggt 2270

aggggtgtgg gggccgtctt cctcctgtct ctttcctttc accctagcct gactggaagc 2330

agaaaatgac caaatcagta ttttttttaa tgaaatatta ttgctggagg cgtcccaggc 2390

aagcctggct gtagtagcga gtgatctggc ggggggcgtc tcagcaccct ccccaggggg 2450

tgcactcag cccctcttt cgtccttcc cgtccagccc cagccctggg cctgggctgc 2510

cgacacctgg gccagagccc ctgctgtgat tgggtctccc tgggcctccc ggggtggatga 2570

agccaggcgt cgccccctcc gggagccctg gggtagccg ccggggcccc cctgctgcca 2630

gcctcccccg tccccaacat gcatctcact ctgggtgtct tggcttttta ttttttgtaa 2690

gtgtcatttg tataactcta aacgcccacg atagtagctt caaactggaa atagcgaaat 2750

aaaataactc agtctgcgcg tg

2772

<210> 28

<211> 493

<212> PRT

<213> Homo sapiens

<400> 28

Met Thr Ala Asn Gly Thr Ala Glu Ala Val Gln Ile Gln Phe Gly Leu
1 5 10 15

Ile Asn Cys Gly Asn Lys Tyr Leu Thr Ala Glu Ala Phe Gly Phe Lys
20 25 30

Val Asn Ala Ser Ala Ser Ser Leu Lys Lys Lys Gln Ile Trp Thr Leu
35 40 45

Glu Gln Pro Pro Asp Glu Ala Gly Ser Ala Ala Val Cys Leu Arg Ser
50 55 60

His Leu Gly Arg Tyr Leu Ala Ala Asp Lys Asp Gly Asn Val Thr Cys
65 70 75 80

Glu Arg Glu Val Pro Gly Pro Asp Cys Arg Phe Leu Ile Val Ala His
85 90 95

Asp Asp Gly Arg Trp Ser Leu Gln Ser Glu Ala His Arg Arg Tyr Phe
100 105 110

Gly Gly Thr Glu Asp Arg Leu Ser Cys Phe Ala Gln Thr Val Ser Pro
115 120 125

Ala Glu Lys Trp Ser Val His Ile Ala Met His Pro Gln Val Asn Ile
130 135 140

Tyr Ser Val Thr Arg Lys Arg Tyr Ala His Leu Ser Ala Arg Pro Ala
145 150 155 160

Asp Glu Ile Ala Val Asp Arg Asp Val Pro Trp Gly Val Asp Ser Leu
165 170 175

Ile Thr Leu Ala Phe Gln Asp Gln Arg Tyr Ser Val Gln Thr Ala Asp
180 185 190

His Arg Phe Leu Arg His Asp Gly Arg Leu Val Ala Arg Pro Glu Pro
195 200 205

Ala Thr Gly Tyr Thr Leu Glu Phe Arg Ser Gly Lys Val Ala Phe Arg
210 215 220

Asp Cys Glu Gly Arg Tyr Leu Ala Pro Ser Gly Pro Ser Gly Thr Leu
225 230 235 240

Lys Ala Gly Lys Ala Thr Lys Val Gly Lys Asp Glu Leu Phe Ala Leu
245 250 255

Glu Gln Ser Cys Ala Gln Val Val Leu Gln Ala Ala Asn Glu Arg Asn
260 265 270

14

Val Ser Thr Arg Gln Gly Met Asp Leu Ser Ala Asn Gln Asp Glu Glu
 275 280 285

Thr Asp Gln Glu Thr Phe Gln Leu Glu Ile Asp Arg Asp Thr Lys Lys
 290 295 300

Cys Ala Phe Arg Thr His Thr Gly Lys Tyr Trp Thr Leu Thr Ala Thr
 305 310 315 320

Gly Gly Val Gln Ser Thr Ala Ser Ser Lys Asn Ala Ser Cys Tyr Phe
 325 330 335

Asp Ile Glu Trp Arg Asp Arg Arg Ile Thr Leu Arg Ala Ser Asn Gly
 340 345 350

Lys Phe Val Thr Ser Lys Lys Asn Gly Gln Leu Ala Ala Ser Val Glu
 355 360 365

Thr Ala Gly Asp Ser Glu Leu Phe Leu Met Lys Leu Ile Asn Arg Pro
 370 375 380

Ile Ile Val Phe Arg Gly Glu His Gly Phe Ile Gly Cys Arg Lys Val
 385 390 395 400

Thr Gly Thr Leu Asp Ala Asn Arg Ser Ser Tyr Asp Val Phe Gln Leu
 405 410 415

Glu Phe Asn Asp Gly Ala Tyr Asn Ile Lys Asp Ser Thr Gly Lys Tyr
 420 425 430

Trp Thr Val Gly Ser Asp Ser Ala Val Thr Ser Ser Gly Asp Thr Pro
 435 440 445

Val Asp Phe Phe Phe Glu Phe Cys Asp Tyr Asn Lys Val Ala Ile Lys
 450 455 460

Val Gly Gly Arg Tyr Leu Lys Gly Asp His Ala Gly Val Leu Lys Ala
 465 470 475 480

Ser Ala Glu Thr Val Asp Pro Ala Ser Leu Trp Glu Tyr
 485 490

<210> 29

<211> 2634

<212> DNA

<213> Mus musculus

<220>

<221> CDS

<222> (79)..(1557)

<400> 29

gagctaagcc gtgttgaaca aaggaggtcg ggcacagcta tccaagctcc cggggccacc 60

gggccgcct cgcgcacc atg acc gcc aac ggc acg gca gag gct gtg cag 111
 Met Thr Ala Asn Gly Thr Ala Glu Ala Val Gln
 1 5 10

att cag ttc ggg ctc atc agc tgc ggc aac aag tac ctg aca gcc gag 159
 Ile Gln Phe Gly Leu Ile Ser Cys Gly Asn Lys Tyr Leu Thr Ala Glu
 15 20 25

gcg ttc ggg ttc aag gtg aac gca tcc gct agt agc ttg aaa aag aag	207
Ala Phe Gly Phe Lys Val Asn Ala Ser Ala Ser Ser Leu Lys Lys Lys	
30 35 40	
cag atc tgg acg ctg gag caa cct ccc gat gag gcg ggc agc gcg gcc	255
Gln Ile Trp Thr Leu Glu Gln Pro Pro Asp Glu Ala Gly Ser Ala Ala	
45 50 55	
gtg tgt ctg cgc acg cac ctg ggt cgc tac ctg gcc gcc gac aag gac	303
Val Cys Leu Arg Thr His Leu Gly Arg Tyr Leu Ala Ala Asp Lys Asp	
60 65 70 75	
ggc aac gtg acc tgc gag cgc gag gtg ccc gac ggc gac tgc cgc ttt	351
Gly Asn Val Thr Cys Glu Arg Glu Val Pro Asp Gly Asp Cys Arg Phe	
80 85 90	
ctc gtc gtg gcg cac gac gac ggc cgc tgg tgc ctg cag tcc gag gct	399
Leu Val Val Ala His Asp Asp Gly Arg Trp Ser Leu Gln Ser Glu Ala	
95 100 105	
cac cgg cgc tac ttt ggc ggc acc gag gac cgc ctg tcc tgc ttc gcg	447
His Arg Arg Tyr Phe Gly Gly Thr Glu Asp Arg Leu Ser Cys Phe Ala	
110 115 120	
cag agc gtg tgc ccg gcc gag aag tgg agc gtg cac atc gcc atg cac	495
Gln Ser Val Ser Pro Ala Glu Lys Trp Ser Val His Ile Ala Met His	
125 130 135	
ccg cag gtt aac atc tac agc gtt acc cgc aag cgc tac gcg cat ctg	543
Pro Gln Val Asn Ile Tyr Ser Val Thr Arg Lys Arg Tyr Ala His Leu	
140 145 150 155	
agc gcg cgg ccg gcc gac gag atc gcg gta gac cgc gac gtg cct tgg	591
Ser Ala Arg Pro Ala Asp Glu Ile Ala Val Asp Arg Asp Val Pro Trp	
160 165 170	
ggc gtc gac tgc ctc atc acc ttg gcc ttc cag gac caa cgc tac agt	639
Gly Val Asp Ser Leu Ile Thr Leu Ala Phe Gln Asp Gln Arg Tyr Ser	
175 180 185	
gtg cag acg tcc gac cac cgc ttc ctg cgc cac gac ggg cgc ctt gtg	687
Val Gln Thr Ser Asp His Arg Phe Leu Arg His Asp Gly Arg Leu Val	
190 195 200	
gca cgg ccg gag ccc gcc acg ggc ttc acg ctg gag ttc cgc tcc ggc	735
Ala Arg Pro Glu Pro Ala Thr Gly Phe Thr Leu Glu Phe Arg Ser Gly	
205 210 215	
aag gtg gcc ttt cgc gac tgc gaa ggt cgc tac ctg gct ccg tcc ggg	783
Lys Val Ala Phe Arg Asp Cys Glu Gly Arg Tyr Leu Ala Pro Ser Gly	
220 225 230 235	
ccc agc ggc acc ctc aag gct ggc aag gcc acc aag gtg ggc aaa gat	831
Pro Ser Gly Thr Leu Lys Ala Gly Lys Ala Thr Lys Val Gly Lys Asp	
240 245 250	
gag ctc ttc gcc ctg gaa cag agc tgc gct gag gtg gtg ctg cag gcg	879
Glu Leu Phe Ala Leu Glu Gln Ser Cys Ala Glu Val Val Leu Gln Ala	
255 260 265	
gcc aac gag ggg aac gtg tcc acg cgc cag gga atg gac ctg tca gcc	927
Ala Asn Glu Gly Asn Val Ser Thr Arg Gln Gly Met Asp Leu Ser Ala	

270	275	280	
aat cag gat gaa gag acc gat cag gag acc ttc cag ctg gag atc gac			975
Asn Gln Asp Glu Glu Thr Asp Gln Glu Thr Phe Gln Leu Glu Ile Asp			
285	290	295	
cgc gac aca aga aag tgt gcc ttt cgc acc cac acg ggc aag tac tgg			1023
Arg Asp Thr Arg Lys Cys Ala Phe Arg Thr His Thr Gly Lys Tyr Trp			
300	305	310	315
aca ctg acg gcg acc gga ggt gtg caa tcc act gcg tcc acc aag aac			1071
Thr Leu Thr Ala Thr Gly Gly Val Gln Ser Thr Ala Ser Thr Lys Asn			
	320	325	330
gcc agc tgc tac ttt gac atc gag tgg tgt gac cgc cgg atc act ctg			1119
Ala Ser Cys Tyr Phe Asp Ile Glu Trp Cys Asp Arg Arg Ile Thr Leu			
	335	340	345
aga gcc tcc aac ggc aag ttt gtg acc gcc aag aaa aat ggc cac gtg			1167
Arg Ala Ser Asn Gly Lys Phe Val Thr Ala Lys Lys Asn Gly His Val			
	350	355	360
gcc gcc tcg gtg gag aca gca ggg gac tcg gaa ctc ttc ctc atg aag			1215
Ala Ala Ser Val Glu Thr Ala Gly Asp Ser Glu Leu Phe Leu Met Lys			
	365	370	375
ctg att aac cgc ccc atc att gcg ttc cgg ggg gaa cac ggg ttc att			1263
Leu Ile Asn Arg Pro Ile Ile Ala Phe Arg Gly Glu His Gly Phe Ile			
380	385	390	395
gcg tgc cgc aag gtc acg ggc act ctg gat gcc aac cgt tcc agt tac			1311
Ala Cys Arg Lys Val Thr Gly Thr Leu Asp Ala Asn Arg Ser Ser Tyr			
	400	405	410
gat gtc ttc cag ttg gaa ttc aat gac ggc gcc tac aac atc aaa gac			1359
Asp Val Phe Gln Leu Glu Phe Asn Asp Gly Ala Tyr Asn Ile Lys Asp			
	415	420	425
tcc acg ggc aag tac tgg acg gtg ggt agt gat tcc tcg gtc acc agc			1407
Ser Thr Gly Lys Tyr Trp Thr Val Gly Ser Asp Ser Ser Val Thr Ser			
	430	435	440
agc agc gac acc cct gtg gat ttc ttc ctt gag ttc tgt gac tac aat			1455
Ser Ser Asp Thr Pro Val Asp Phe Phe Leu Glu Phe Cys Asp Tyr Asn			
	445	450	455
aag gtg gct ctc aag gtg ggc ggc cgc tac ctg aag ggg gac cac gct			1503
Lys Val Ala Leu Lys Val Gly Gly Arg Tyr Leu Lys Gly Asp His Ala			
460	465	470	475
ggg gtc ctg aag gcc tgc gcg gag act atc gac ccc gcc tca ctc tgg			1551
Gly Val Leu Lys Ala Cys Ala Glu Thr Ile Asp Pro Ala Ser Leu Trp			
	480	485	490
gag tac tagggccacc tgccctctgca gccgctctcg tcagtcctcc tggtatcctt 1607			
Glu Tyr			
actcatcggg tggcctgcag caggtggcaa accccttgcc tttcaaactg gaaacccaag 1667			
agaaaacggg gcccttgctg tcaccctctg tggaccctt ttccctaact cactgctccc 1727			
catggggtcgg tggctgcaga ctgtccccag gaggactctg gttccctctg tccccttctt 1787			

tccatgggga actctggcac ctttcttctg acctcagtca actctgagcc ttatttcccc 1847
 ccaggaagtg gcctaggaga agctacaggg cctagggact taccctgagc ttgtaactgg 1907
 aagaccccggt ccctatcccc gctcccgccc ccaccccacc ccacccctgc tctggcccca 1967
 gcctctggag gccagccttt tggcgggact gaagccgggc atggccaacc ttgccacaa 2027
 gtgtttttct ggatcttggc tggaaggcag tctgtcccat cctgcagtgt ttgggcctgg 2087
 ctctttgact caaagctagc taggtggcac tccgtgtcgc tctgcacat tctggaaggg 2147
 gcgggcctct caccacctc attccttttc cccctggcct gactggaagc agaaaaatga 2207
 ccaaatacagt attttttttt tttctttaag gaaatgttac tgttgaaagg ccctaggcaa 2267
 gcctgccttg ttggttgtag tctgtgagtg tcttgggggg agatgcttg ctctgtccc 2327
 tgccctccca gcggttccct cctccctcc tgcccgacca cccagctct ggctctgtga 2387
 ttggtgctcc acgtctccag acacctcggg gctcctgggc ggagaaagcc gatgtgcccc 2447
 tccttgggag cctgagtaa acctcagggg gccctttccc aatcaccct ccaccgacct 2507
 ctcaacacca tgcattctac tctgggtgta ctgcctcaca tttatttttt tgtaactgtc 2567
 atttctataa ctctgaagac ccatgatagt aagctttgaa ctggaaaata aagtaaaatc 2627
 aagtctg 2634

<210> 30

<211> 493

<212> PRT

<213> Mus musculus

<400> 30

Met Thr Ala Asn Gly Thr Ala Glu Ala Val Gln Ile Gln Phe Gly Leu
 1 5 10 15
 Ile Ser Cys Gly Asn Lys Tyr Leu Thr Ala Glu Ala Phe Gly Phe Lys
 20 25 30
 Val Asn Ala Ser Ala Ser Ser Leu Lys Lys Lys Gln Ile Trp Thr Leu
 35 40 45
 Glu Gln Pro Pro Asp Glu Ala Gly Ser Ala Ala Val Cys Leu Arg Thr
 50 55 60
 His Leu Gly Arg Tyr Leu Ala Ala Asp Lys Asp Gly Asn Val Thr Cys
 65 70 75 80
 Glu Arg Glu Val Pro Asp Gly Asp Cys Arg Phe Leu Val Val Ala His
 85 90 95
 Asp Asp Gly Arg Trp Ser Leu Gln Ser Glu Ala His Arg Arg Tyr Phe
 100 105 110
 Gly Gly Thr Glu Asp Arg Leu Ser Cys Phe Ala Gln Ser Val Ser Pro
 115 120 125
 Ala Glu Lys Trp Ser Val His Ile Ala Met His Pro Gln Val Asn Ile
 130 135 140

Tyr	Ser	Val	Thr	Arg	Lys	Arg	Tyr	Ala	His	Leu	Ser	Ala	Arg	Pro	Ala	145	150	155	160
Asp	Glu	Ile	Ala	Val	Asp	Arg	Asp	Val	Pro	Trp	Gly	Val	Asp	Ser	Leu	165	170	175	
Ile	Thr	Leu	Ala	Phe	Gln	Asp	Gln	Arg	Tyr	Ser	Val	Gln	Thr	Ser	Asp	180	185	190	
His	Arg	Phe	Leu	Arg	His	Asp	Gly	Arg	Leu	Val	Ala	Arg	Pro	Glu	Pro	195	200	205	
Ala	Thr	Gly	Phe	Thr	Leu	Glu	Phe	Arg	Ser	Gly	Lys	Val	Ala	Phe	Arg	210	215	220	
Asp	Cys	Glu	Gly	Arg	Tyr	Leu	Ala	Pro	Ser	Gly	Pro	Ser	Gly	Thr	Leu	225	230	235	240
Lys	Ala	Gly	Lys	Ala	Thr	Lys	Val	Gly	Lys	Asp	Glu	Leu	Phe	Ala	Leu	245	250	255	
Glu	Gln	Ser	Cys	Ala	Glu	Val	Val	Leu	Gln	Ala	Ala	Asn	Glu	Gly	Asn	260	265	270	
Val	Ser	Thr	Arg	Gln	Gly	Met	Asp	Leu	Ser	Ala	Asn	Gln	Asp	Glu	Glu	275	280	285	
Thr	Asp	Gln	Glu	Thr	Phe	Gln	Leu	Glu	Ile	Asp	Arg	Asp	Thr	Arg	Lys	290	295	300	
Cys	Ala	Phe	Arg	Thr	His	Thr	Gly	Lys	Tyr	Trp	Thr	Leu	Thr	Ala	Thr	305	310	315	320
Gly	Gly	Val	Gln	Ser	Thr	Ala	Ser	Thr	Lys	Asn	Ala	Ser	Cys	Tyr	Phe	325	330	335	
Asp	Ile	Glu	Trp	Cys	Asp	Arg	Arg	Ile	Thr	Leu	Arg	Ala	Ser	Asn	Gly	340	345	350	
Lys	Phe	Val	Thr	Ala	Lys	Lys	Asn	Gly	His	Val	Ala	Ala	Ser	Val	Glu	355	360	365	
Thr	Ala	Gly	Asp	Ser	Glu	Leu	Phe	Leu	Met	Lys	Leu	Ile	Asn	Arg	Pro	370	375	380	
Ile	Ile	Ala	Phe	Arg	Gly	Glu	His	Gly	Phe	Ile	Ala	Cys	Arg	Lys	Val	385	390	395	400
Thr	Gly	Thr	Leu	Asp	Ala	Asn	Arg	Ser	Ser	Tyr	Asp	Val	Phe	Gln	Leu	405	410	415	
Glu	Phe	Asn	Asp	Gly	Ala	Tyr	Asn	Ile	Lys	Asp	Ser	Thr	Gly	Lys	Tyr	420	425	430	
Trp	Thr	Val	Gly	Ser	Asp	Ser	Ser	Val	Thr	Ser	Ser	Ser	Asp	Thr	Pro	435	440	445	
Val	Asp	Phe	Phe	Leu	Glu	Phe	Cys	Asp	Tyr	Asn	Lys	Val	Ala	Leu	Lys	450	455	460	
Val	Gly	Gly	Arg	Tyr	Leu	Lys	Gly	Asp	His	Ala	Gly	Val	Leu	Lys	Ala	465	470	475	480

Cys Ala Glu Thr Ile Asp Pro Ala Ser Leu Trp Glu Tyr
 485 490

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
 sequence

<400> 31

gccaccatga ccgccaacgg

20

<210> 32

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
 sequence

<400> 32

tgtgtcgcgg tcgatctcca

20

<210> 33

<211> 4100

<212> DNA

<213> Homo sapiens

<400> 33

aagcttctag aaggggtata ttttgccccc agggcccaaa tcctgggtctt tggactacag 60
 tcaagacctg gcacaggggt caggcctaaa aaaaagctgt tgaggagggg gtgaggctgc 120
 aacgttgagg tgagaagggg gtccccaggg agggcgagc aggaagcccc aggggaagtgc 180
 ctgggagagg gaggttggtc agccagaagg agcagccgc agcctttggg tgttgactcc 240
 tgctttgtaa gtggcagttc gggtgggtag gacaggggtc gagtcctcac tcagggaact 300
 gagtccaggg tgagctgagc agcccttccg ggtggcctca ctttccctcg cagagcctgc 360
 tgcattggtg ccagtggcca gcctgggagg agctcaggac tggcctcacc tcgtgccacg 420
 cctcgtacc aaggtgggtc agatggcacc tgggttcccg aaggggccag gaacacagcg 480
 tccatgtccc catccttccc cggagggact tggggtgggg cctgcaggaa ggatcatgtg 540
 acttggttcaa gggcagcttg tcctgccccg gacacagagg tgcccatcgt aagcgagatg 600
 caggcaaggt gaggacaggg catggtgccc agcaggaatg attttcgaaa atgctactct 660
 gtggtcgggc acggtgggtc actcctgtaa tcccagttat ttgggagggc caggcgggca 720
 gatcatgagg tcaggagttc aagaccagcc tggccaacat ggtgaaaccc cttctctact 780
 gaaaatacaa aaaactaccc gggcgtgggtg gtggatgcct gtaatcccag ctactcagga 840
 ggctgaggga ggagaatcgc ttgaaccttg gaggtgaagg ttgcaatgag ccgagattgc 900
 accactgcac tccgacctgg gcgacagaga aagactccgt ctgaaaaaca aacaaacaaa 960
 accccgagat tctcttttcc cccgtccgga gctctatggc catctggagc tcgttctgtc 1020
 cacaaggaca catttccctg cagcagctgt ggaccagggc tcaactcactc acatctgtac 1080
 ccaatctaga gcaggacaaa cacctatcac ctgcaactggc aatggacaga ggacagtggc 1140
 agcccatca ccagaggcat tcaagccaag gcattttcta tcgtctatct atctatctat 1200
 ctatctatct atctatctat ctatctatct atctatctag atgagatctt gctatgttgc 1260
 ccaggttgaa ctctggcct caagcgatcc tcctgcttca gcctcccaaa gtgttgggat 1320
 ttcaggtgtg aacctgtga cctggctgga gtgcagtggc actgtcatag ctactgcag 1380
 cctccacctc ctgggctcaa ggaatcctct tcctcagcc tcctgagtag ctgggaccac 1440
 aggcattgcac catcacaccc agctaaacct tgattttttc catagtatat ggctcagggc 1500

```

agagcaaaga agcacaaaaa catgttggtc tgcggttgag ggctgatggg aggggggttct 1560
ggccaaagcc caaaattaac agccacaacg tcagtgtctg tgggaggggt tgccaggggc 1620
gtgcgggttc tggggctcaa ggccctgccg gtaaaccat ttgaagcagg atggcaagaa 1680
ggtgacacca tcttcccccg cgcactgaaa gcccctggct gggattgcct ggggcagaca 1740
caggctcggg ccagccccag caatcccagt ttatcagcga gccggctgag ggcccggagt 1800
tatctcagtg cccggcctga gaccttgtgg gcagcctgtg gtcattgctg cattccaggg 1860
gccttttggc atgtggggaa tgtccaggaa aagcctcagc cttcgggtgag gcgcagaaaa 1920
gggaagtgtc cctagagggg gtgggtgagg gcgtgggagg tgggtgtctgc agggaaatgtc 1980
ccctttgggg gaggaggatg gaggttgagg attctgagga tggggggggg ggctgtagcc 2040
agcaccatgt ccctcctgtg tgaccagctc agagtcccat gaaattgggg ctggggaggg 2100
gaagggacac tggcctggga accagagacc tgggctgggtc tggctcacag ttgctggacc 2160
tctgtgatcc gtgtcaaaaa acgagaacac caattcctgt cctgcccact caccaccagg 2220
tgggacctga accctggcat cgccagcatt gggaatgtcg gccactgact caaccactct 2280
cccgagagac tatttgggac acccgaggcg ggtgcctggg ccacacggag gggctcctgg 2340
ggtcttcagg gcagcggctg tggggctgaa gcctcaagga accacatctc tgcataggag 2400
ggccaggctg cagggcctcg ggagacaacc tagttgggag tgttgggggtc tgtgggtccc 2460
aggtcctggc ctccacgggt ccccaccgag ctgtcagctc ccagcctctt tccctcgtct 2520
gcctctgggc ttctgtaagg ctatgtgtct caaggccacc tctccaggc agccctcaga 2580
ccccacctt cgcccagtac cgatctgcac cgtggtctct gaagtctcca tctgtacctc 2640
aaacctcgct cgtccttgct cctagcgagg cttggggtcg ggtgttccga ggtgggggac 2700
atccgggggg gttaggtggc tggcgcgggg agccgggggt gtgaggggtg atgtcctcag 2760
gcggcggcgc tgcggggtgc ggcgaggaca ccgggtgggt gagagcaccg gcggggcagc 2820
agcggggggc gcagcgccgg gtccctcggc ccggggcccc tcccgcgagg agccaggggc 2880
gggacagggg gggtggcctt ggtggcgctg acgtcacctc gcctataaaa tgtccggggc 2940
gccgctagct gggctttgtg gagcgctgcg gaggtgtcgt gcgggcccgc gcagccgaac 3000
aaaggagcag gggcgccgac gcagggacc gccaccacc tcccgggggc gcgcagcgc 3060
ctctcgtcta ctgccaccat gaccgccaac ggcacagccg aggcggtgca gatccagttc 3120
ggcctcatca actgcggcaa caagtacctg acggccgagg cgttcgggtt caaggtgaac 3180
gcgtccgcca cagcctgaa gaagaagcag atctggacgc tggagcagcc cctgacgag 3240
gcgggcagcg cggcctgtg cctgcgcagc cacctgggac gctacctggc ggcggacaag 3300
gacggcaacg tgacctgcga gcgcgaggtg cccggtcccg actgccgtt cctcatcgtg 3360
gcgcacgacg acggtcgtg gtgcgtgcag tccgaggcgc accggcgcta cttcggcggc 3420
accgaggacc gctgtcctg cttcgcgcag acgggtcccc cgccgagaag tggagcgtgc 3480
acatcgccat gcacctcag gtcaacatct acagcgtcac ccgtaagcgc tacgcgcacc 3540
tgagcgcgcg gccggccgac gagatcgccg tggaccgcga cgtgccctgg ggcgtcgact 3600
cgctcatcac cctcgcttc caggaccagc gctacagcgt gcagaccgcc gaccaccgct 3660
tctgcgcca cgacggggcg cttggtggcg gcccgagcc ggccactggc tacacgctgg 3720
agttccgctc cggcaagggt gccttccgag actgcgaggg ccgttacctg gcgcgctcg 3780
ggcccgaggc cagctcaag gcgggcaagg ccaccaagg gggcaaggac gagctctttg 3840
ctctggagca gagctgcgc caggtcgtgc tgcaggcggc caacgagagg aacgtgtcca 3900
cgcgccaggg tgagtgggga cgtgcctcgt gctcgtgcac aaagcgacc 3960
ccaccgcgc cctccagcc tcccgccctt tctcgtcgc ggcgcgctg cggctcggag 4020
cactgcccac tgcgcccccg ctaggcacgc gggctacccc gcctggaggg ggcgaggagt 4080
ggggctttgc ccatccgcgg 4100

```

<210> 34

<211> 2843

<212> DNA

<213> Homo sapiens

<400> 34

```

caatcctcct gccttggggc tcccaaagtg ttgcgatcac aggcattgagc ccctgtccct 60
ggcccttatt atttctctaa ctggggaaag tcatgtggga acagatgtag cttgccttgg 120
cctctgaccg gccctgcctg cgttcctggg tgctctctgc tgcttctcat gtgtgccact 180
gtggggactc ggccgcccac cccaccctgt ggtgttacct tgcgtgtgta gttctgtgag 240
ctcagggtca tggctgtcca gaactagggg gcgtggggcc ccagtaccag cccaaggcct 300
cctctctgca ggtatggacc tgtctgcca gtgtgccttc cgtaccacca cgggcaagta 360
ccagctggag atcgaccgag acaccaaaaa gtgtgccttc cgtaccacca cgggcaagta 420
ctggacgctg acggccaccg ggggctgtga gtccaccgcc tccagcaagt gaggcctcg 480
ctccacctg tcaccgcccc caccaccttg cctgggctac cccgctgac cctgtcccgc 540
catccccag gaatgccagc tgctactttg acatcgagtg gcgtgaccgg cgcatacac 600
tgagggcgtc caatggcaag tttgtgacct ccaagaagaa tgggcagctg gccgcctcgg 660

```

```

tggagacagc aggtaacact aaagccccag ttccctggag ccgtcctgga gtccctggagg 720
gtctggccat gccgtgggtca cttggtagcc ccagccaagg cctgctctgt gctgggcatc 780
cccccgact gggcccgcac tgtcctaccc tggggactgc tgtgtgaccc cagctcctgg 840
ccctccctct ctggtcaccc cagcctccac ccactccct gccaggaggc tcaactgactc 900
ccctctttct gggacagggg actcagagct cttcctcatg aagctcatca accgccccat 960
catcgtgttc cgcggggagc atggcttcat cggtgcgc aaggtcacgg gcaccctgga 1020
cgccaaccgc tccagctatg acgtcttcca gctggagtgc aacgatggcg cctacaacat 1080
caaaggcagg ttctcctgtg ggcagctgct gggcagggaa cccctcggtc ggggctgggg 1140
tcagtgtgc ggggagcgcc ctctgcatcc acactggacc ctggcttggc tcagggccat 1200
tccaggccct aaagggacag gtgtctgatg gccaccaggg ggctctggga tgcaagcagc 1260
ccctttccct cttgtctgtg tggttggggg gacttacctt gccacactga cagagagggtg 1320
tgtggagggg agagcagggg gggaaaggga ccaggaaaagg gagggaggag agcagggggag 1380
gggagagccg ggaagagagg agagcagggg tggggagggt tctggaaagg gtgtgcaagg 1440
ggaggacgag cctcgtttat gggactggag ccccttccca ggaggacccc caacaatcca 1500
gaggtgcctg ttaggattca gaacatgggt tttttgtttg tttttttgag actcactccc 1560
tcaccctggc tggacttgca gtggcgttat ctcggtcac tgcaacctct gcctcctggg 1620
ttcaagtgat tctcctgcct cagcctccca agtagctggg attacagggtg tgcaccacca 1680
cgcctggcta atttttatat ttttaaaatt tatttttat ttattttgag acccagtctg 1740
gagtgcagtg gcgttatctc ggctctctgc aacctctgcc tccctgggttc aagcgattct 1800
cctgcctcag cctcccgagt agctgggact atgtgtggga gccaccatgc ctggctaatt 1860
ttttgtatt tttcatagag acgggtttca ccatgttgtc caggctgggtc ttgaattcgt 1920
ggcctcaagt gatccgcccc cctcagcctc ccacagtgtc gggtttatag gtgtgagcca 1980
ccacaccgg ctaattgttt tgtattttta gtagagacgg agcttacta tgttggcaag 2040
gctggctcga actcctgacc tcaagtgatc cgccacctc agcctccaa agtgctggga 2100
ttacaggcgt gagccactgc ggccgagcag aacacgttct aggaccctg ttcattgtgc 2160
catcatggac aggaggacgt gcggggcata gggaccctgg ctcatccgg agccgggact 2220
ggagggtggg gcgtcacctc tgggaacacc cgtgccacc ctccgctgcc cagggtaggg 2280
gtggggagcc aggttttggg cccacttga taaagtcccc tcccagact ccacaggcaa 2340
atactggacg gtgggcagtg actccgcggt caccagcagc ggcgacactc ctgtggactt 2400
cttcttcgag ttctgcgact ataacaagggt ggccatcaag gtggcgggc gctacctgaa 2460
gggcgaccac gcaggcgtcc tgaaggcctc ggcggaaacc gtggacccc cctcgctctg 2520
ggagtactag ggccggcccc tccctccccg cccctgccca catggcggct cctgccaacc 2580
ctccctgcta accccttctc cgccagggtg gctccagggc gggaggcaag ccccttgcc 2640
tttcaaactg gaaacccag agaaaacggg gccccacct gtcgccccta tggactcccc 2700
actctccct ccgcccgggt tccctactcc cctcgggtca gcggtgcgg cctggccctg 2760
ggagggattt cagatgcccc tgccctcttg tctgccacgg ggcgagtctg gcacctctt 2820
cttctgacct cagacggctc tga
2843

```

<210> 35

<211> 934

<212> DNA

<213> Homo sapiens

<400> 35

```

ccacctctc ccagcccccc aggagagctg ggcacatgtc ccaagcctgt cagtggccct 60
ccctgggtgca ctgtccccga aacccctgct tgggaaggga agctgtcggg tgggctagga 120
ctgacctttg tgggtgtttt ttgggtgggt gctggaaaca gcccctctcc cacgtggcag 180
aggctcagcc tggctccctt ccctggagcg gcagggcgtg acggccacag ggtctgccc 240
ctgcacgttc tgccaagggt gtggtggcg gcggttaggg gtgtgggggc cgtcttcctc 300
ctgtctcttt cctttcaccc tagcctgact ggaagcagaa aatgaccaa tcaagtattt 360
ttttaatgaa atattattgc tggaggcgct ccaggcaagc ctggctgtag tagcgagtga 420
tctggcgggg ggcgtctcag caccctcccc aggggggtgca tctcagcccc ctctttccgt 480
ccttcccgte cagccccagc cctgggcctg ggctgccgac acctgggcca gagccctgc 540
tgtgattggg gctccctggg cctcccggtt ggaatgaagc aggcgtcgcc cctccggga 600
gccttggggg gagcgcgggt ggccccctg ctgccagcct ccccgctccc caacatgcat 660
ctcactctgg gtgtcttggg cttttatttt ttgtaagtgt catttgtata actctaaacg 720
cccatgatag tagcttcaaa ctggaaatag cgaaataaaa taactcagtc tgcagcccca 780
ggccggcctg tgtgtgtctt ggggctgagg tgggtggggg ggctgagggt ggtgggaggg 840
ctggcgggac aggtaggcgc cctggctccc cagctcagtg ctgggagtg gtagtgggag 900
ggaggccgtg gctccagtgg gtgtccgga gctc
934

```

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 36
aagcttctag aaggggtata

20

<210> 37
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 37
agacgagagg ccgctgcgcg

20

<210> 38
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 38
gtggcagccc catcaccaga

20

<210> 39
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 39
catcacaccc agctaaacct

20

<210> 40
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 40
gtgcgggttc tggggctcaa

20

<210> 41
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 41
ggcagcctgt ggatcatgctg 20

<210> 42
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 42
gacctgggct ggtctggctc 20

<210> 43
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 43
gcagggcctc gggagacaac 20

<210> 44
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 44
catccggggg ggtaggtgg 20

<210> 45
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 45
gtgagtggggg acgctgcc 18

<210> 46
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 46
ccgcggatgg gcaaagcc 18

<210> 47
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 47
gtgtcagccc ttccccccag 20

<210> 48
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 48
gagctcggat cctccccttc 20

<210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 49
gggcccgggtg ggggcagggt 20

<210> 50
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial

sequence

<400> 50

caggtcctat ccctggcatg

20

<210> 51

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 51

ctgcagtgag ccgagatcgt

20

<210> 52

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 52

ggtacctctc cctcctcctt

20

<210> 53

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 53

gggactgagt cgtggtacag

20

<210> 54

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 54

gctcaaagga tcctcttgcc

20

<210> 55

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 55

ctgcagcctc cgctcctgg

20

<210> 56

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 56

gccgaggcgg gcggatcatg

20

<210> 57

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 57

ctgcagaaca cccctgtgtt

20

<210> 58

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 58

caaggctggg cttgaactca

20

<210> 59

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial
sequence

<400> 59

caatcctcct gccttgggcc

20

<210> 60

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 60

ctgcagagag gaggccttgg

20

<210> 61

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 61

gtgagtgcct cgctcccacc

20

<210> 62

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 62

ctgggggatg gcgggacagg

20

<210> 63

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 63

gtaacactaa agccccagtt

20

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 64

ctgtcccaga aagaggggag

20

<210> 65

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 65
gcaggttctc ctgtgggcag 20

<210> 66
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 66
ctggggaggg gactttatca 20

<210> 67
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 67
agccccaggc cggcctgtgt 20

<210> 68
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 68
gagctccgga gcacccactg 20

<210> 69
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificial
sequence

<400> 69
aagcttctag aaggggtata ttttg 25

<210> 70

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 70

cgttcacctt gaacccgaa

19

<210> 71

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificial sequence

<400> 71

gcagccgaac aaaggagcag g

21

<210> 72

<211> 16951

<212> DNA

<213> Homo sapiens

<400> 72

```

aagcttctag aaggggtata ttttgcccc aggccccaaa tcctggtctt tggactacag 60
tcaagacctg gcacagggct caggcctaaa aaaaagctgt tgaggagggg gtgaggctgc 120
aacgttgccg tgagaagggg gtccccaggg agggcgagc aggaagcccc agggaggtgc 180
ctgggagagg gaggttggtc agccagaagg agcagcccg agcctttggt tgttgactcc 240
tgctttgtaa gtggcagtct ggttgggtag gacaggggtcc ggtcctcac tcagggaact 300
gagtcagggt tgagctgagc agcccttcct ggtggcctca cttccctctg cagagcctgc 360
tgcatgggtg ccagtgagca gcctgggagg agctcaggac tggcctcacc tcgtgccacg 420
ccctcgtagc aaggtggctc agatggcacc tgggttcccg aaggggccag gaacacagcg 480
tccatgtccc catccttccc cggagggact tggggtgggg cctgcaggaa ggatcatgtg 540
acttggtcaa gggcagcttg tcctgccccg gacacagagg tgcccatcgt aagcgagatg 600
caggcaaggt gaggacaggg catggtgccg agcaggaatg attttcgaaa atgtactctt 660
gtggtcgggc acggtggctc actcctgtaa tcccagtag ttgggaggcc caggcgggca 720
gatcatgagg tcaggagttc aagaccagcc tggccaacat ggtgaaacct cttctctact 780
gaaaatacaa aaaactaccc gggcgtgggt gtggatgcct gtaatcccag ctactcagga 840
ggctgaggca ggagaatcgc ttgaacctg gaggtgaagg ttgcaatgag ccgagattgc 900
accactgcac tccgacctg ggcacagaga aagactccgt ctgaaaaaca acaaaacaaa 960
accccgagat tctcttttcc ccgctccgga gctctatggc catctggagc tcgttctgtc 1020
cacaaggaca catttcctgg cagcagctgt ggaccagggc tcaactcact acatctgtac 1080
ccaatctaga gcaggacaaa cacctatcac ctgcactggc aatggacaga ggacagtggc 1140
agcccatcat ccagaggcat tcaagccaag gcattttcta tcgtctatct atctatctat 1200
ctatctatct atctatctat ctatctatct atctatctat atgagatctt gctatggtgc 1260
ccaggttgaa ctcttggcct caagcgatcc tcctgcttca gcctcccaa gtgttgggat 1320
ttcaggtgtg aacctgtgta cctggctgga gtgcagtggc actgtcatag ctactgcag 1380
cctccacctc ctgggctcaa ggaatcctct ttccctagcc tcctgagtag ctgggaccac 1440
aggcatgcac catcacaccc agctaaacct tgatttttcc catagtatat ggctcagggc 1500
agagcaaaga agcacaacaaa catgttgggt tgcggttgag ggctgatggg agggggttct 1560
ggccaaagcc caaaattaac agccacaacg tcagtgtctg tgggaggggt tgccaggggc 1620
gtgcgggttc tggggctcaa ggccctgccg gtaaaccat ttgaagcagg atggcaagaa 1680
ggtgacacca tcttcccccg cgcactgaaa gccctgggt gggattgcct ggggcagaca 1740
caggctcgga ccagccccag caatcccagt ttatcagcga gccggctgag ggcccggagt 1800

```

tatctcagtg	cccgccctga	gaccttgtgg	gcagcctgtg	gtcatgctgg	cattccaggg	1860
gccttttggc	atgtggggaa	tgtccaggaa	aagcctcagc	cttcgggtgag	gcgcagaaaa	1920
gggaagtgtc	cctagagggg	gtgggtgagg	gcgtgggagg	tgggtgtctgc	agggaaatgtc	1980
ccctttgggg	gaggaggatg	gaggggttggg	attctgagga	tggggggggg	ggctgtagcc	2040
agcaccatgt	ccctcctgtg	tgaccagctc	agagtcccat	gaaattgggg	cttgggaggg	2100
gaaggacac	tggcctggga	accagagacc	tgggctggtc	tggctcacag	ttgtgggacc	2160
tctgtgatcc	gtgtcaaaaa	acgagaacac	caattcctgt	cctgcccact	caccaccagg	2220
tgggacctga	accctggcat	cgcagcatt	gggaatgtcg	gccactgact	caaccactct	2280
cccgagagacc	tatttgggccc	acccgaggcg	ggtgccctggg	ccacacggag	gggtcctggg	2340
ggctctcagg	gcagcggctg	tggggctgaa	gcctcaaggga	accacatctc	tgcataggag	2400
ggccaggctg	cagggcctcg	ggagacaacc	tagttgggcg	tgttgggggtc	tgtgggtccc	2460
aggtcctggc	ctcaccgggt	ccccaccgcg	ctgtcagctc	ccagcctctt	tccctcgtct	2520
gcctctgggc	ttctgttaagg	ctatgtgctc	caaggccacc	tcctccaggc	agccctcaga	2580
ccccacctt	cgcccagtac	cgatctgcac	cgtggctctc	gaagtctcca	tcgtgacctc	2640
aaacctcgct	cgctccttgc	cctagcgagg	cttggggctcg	gggtgtccga	ggtgggggac	2700
atccgggggg	gttaggtggc	tggcgcgggg	agccgggggt	gtgaggggtg	atgtcctcag	2760
gcggcgggcg	tgcgggggtg	ggcgaggaca	ccggtggggg	gagagcaccg	gcggggcagc	2820
agcggggggc	gcagcgccgg	gtccctcggc	ccggggcccc	tcccgcgcg	agccaggggc	2880
gggacagggg	ggcgtggcct	ggtggcgctg	acgtcacctc	gcctataaaa	tgtccggggc	2940
gccgctagct	gggctttgtg	gagcgctgcg	gaggggtgct	gcgggcccgc	gcagccgaac	3000
aaaggagcag	gggcgccgcc	gcagggacct	gccaccacc	tcccgggggc	gcgcagcggc	3060
ctctcgtcta	ctgccaccat	gaccgccaac	ggcacagccg	aggcggtgca	gatccagttc	3120
ggcctcatca	actgcggcaa	caagtacctg	acggcgagg	cgttcggggt	caaggtgaac	3180
gcgtccgcga	gcagcctgaa	gaagaagcag	atctggacgc	tggagcagcc	ccctgacgag	3240
gcggcgagcg	cgccgctgtg	cctgcgcagc	cacctggggc	gctacctggc	ggcggaacaag	3300
gacggcaacg	tgacctgcga	gcgcgaggtg	cccggtcccg	actgccgttt	cctcatcgctg	3360
gcgcacgacg	acggctcgctg	gtcgctgcag	tccgaggcgc	accggcgcta	cttcggcgggc	3420
accgaggacc	gcctgtcctg	cttcgcgcag	acgggtgtccc	ccgcccagaa	gtggagcgctg	3480
cacatcgcca	tgcacctca	ggtcaacatc	tacagcgtca	cccgtaaagc	ctacgcgcac	3540
ctgagcgcg	ggccggccga	cgagatcgcc	gtggaccgcg	acgtgccctg	gggcgtcgac	3600
tcgctcatca	ccctcgccct	ccaggaccag	cgctacagcg	tgcagaccgc	cgaccaccgc	3660
ttcctgcgc	acgacgggcg	cctgggtggcg	cgccccgagc	cggccactgg	ctacacgctg	3720
gagttccgct	ccggcaagggt	ggccttccgc	gactgcgagg	gccgttacct	ggcgccgctg	3780
gggcccagcg	gcacgctcaa	ggcgggcaag	gccaccaagg	tgggcaaggga	cgagctcttt	3840
gctctggagc	agagctgcgc	ccaggctcgtg	ctgcaggcgg	ccaacgagag	gaacgtgtcc	3900
acgcgccagg	gtgagtgggg	acgctgcccc	cgctctcct	ggtccgtgca	caaagcgcac	3960
cccaccgcg	ccctccagc	ctcccgccct	ttctcgctcg	cggcgccgct	gcgggtccgga	4020
gcaactgcca	ttgogccccc	gctaggcacg	cgggctaccc	cgcttgagg	gggcgaggag	4080
tggggctttg	cccatcctcg	ggtgcgcgtg	cccactccc	accccgggct	gggatcatgg	4140
gctcccctag	ccccgcggga	gtcgcaaccg	tacgtgcacc	ctcctaacc	ccccccccgc	4200
ccaatcttgg	ctctcccccac	gcgcgctgat	ccctcggtac	agctgtccca	gctcttgccg	4260
agtgggagcc	cctcaccat	ttcctcgtgc	ccctccccc	gccggctgtc	ctggagctcg	4320
gggggtccgag	gcaagggtcg	ccaccgcga	gggcgcgcct	ccacccccac	cggcagcctt	4380
tcgcggggcg	gatgggggag	gttagccagg	cctttgatcc	cggcgggggc	cgcctccacc	4440
tcccgcgtg	cccgcccttc	ctccacctcc	tcccgcgtgc	gggcgggggtc	ggcctccgct	4500
gggtggggggg	ggcgcggggt	gtcagccctt	ccccccagcc	cctcctcccg	cgtctgcccc	4560
gcgctcgagg	ccgcggccct	tgtgagcagg	ggggcggggtc	gcctcgactg	ggctcctcct	4620
cgcccccttt	gtcctgctcc	atctctgcag	atgggaaaac	cagatgcggc	ggggcggggg	4680
aggggatcgg	cttttgcggt	tcacccctgc	agaggagccc	cccgccgcgc	ccccggggccc	4740
agggctcggg	tcccaggggt	ccccaggagg	cggtctccct	cctcgtcgcc	gcggcccggg	4800
aacggcggtg	cgcggtggc	ggccctccag	gcaccccgcc	cttcgcccgc	cgcgcgcgtc	4860
tccaggcccg	tgcgtgagc	tccgctccgc	gggcgaccga	gggcgggttc	agcgcgagcc	4920
gggagacccc	aggccagctc	cctcggaag	ccctaccct	ctggtgaacc	catcccctag	4980
ggcgctcgcc	gggaacaatt	gactgcaact	cgatcgctcc	cctccggggc	ctcctgcagg	5040
cacggctatt	tgcgaccggc	ctgtgcgcca	ctctttccct	tctttctcag	attctccccg	5100
aggggctttc	tcttcccttt	ggcgctcctc	tccactcctg	gcggagaagc	tgttcccttg	5160
ttgcagacag	aggccccgct	ggagggagg	gcgtgggggg	atctttttcc	ctggatagga	5220
agtccgaccc	caggccttgg	agatggctgg	acgggagccc	agtttgtgtc	ctcatccctc	5280
ctcttggaag	gagccctcct	gacggccccc	ctcttggtg	tgggctctgc	aggaggggga	5340
aagaccccc	atggcagtg	gggtgagggt	ggaggcctgg	gtggggtaat	ggccattttg	5400
tgctgaagt	ttgctacctt	aagtcccaga	gaggtgaagc	tgcttgcat	gggtcacaca	5460
gcaagtgacc	acaaggaagc	agcatgcaga	agtgggtgca	tttggctcca	gaaaccccg	5520
ttcctgtttc	ccacagcgca	tggcgggcag	agctcctcac	ctgatcagca	ggcattgagc	5580

cctactaagg	ggacctgctg	agccatgagt	gaaccagtg	ggttcacgtt	ctcttgggga	5640
cattacgaag	tgtcaaggaa	aaggcagcgg	gtacagggtta	ttagatggct	gagtcctttc	5700
tgagcaggac	atttgagctg	gcctgaagga	tgaatagaag	gtggcaggta	tggggggggg	5760
gctgtgtgga	ccagcgttgt	aggctgtggg	tgtctcctgt	gcaaaggctc	tatggcagga	5820
aggtagacag	aaagcccagt	gtggctggga	ggaagcttgt	gaggtctggg	ccacgggacc	5880
ttgggtgaga	gcttagtgca	gggctatggg	ccttgccctg	gggttgacgc	cacagcctcc	5940
cgtggggaag	aagtcgcaga	agtaggtggc	cctggccctg	ggggccagga	agttgcagag	6000
agcctgaggg	tgtttttggc	gagcctgggc	agataactcc	cctctcctga	gaagcttgct	6060
gggggcgtct	ggtgtgtgat	tcaggctgat	agggtgggag	gaaccagaaa	ttgcagccag	6120
gagaggtggt	ggtgatggct	cagcttgagg	ttgaaaggga	tgtgggacgg	cgcggggggtg	6180
ggggtacttg	ggggccgagg	ttggggctcc	tctggcctca	gatcttgggg	tccttatctc	6240
gtttctatgt	cagccaactc	ccggggggct	tttgggaact	ggagtctgca	ggggtgggcg	6300
tggggggcgt	gtctgcagag	gccatttttg	acgggcctgc	gtggcaaggg	ggaggcgtct	6360
gcttccgaca	cacatgctgg	ggaacgccag	ccaggaaggg	gaggatccgg	acttagctgg	6420
caggggggat	gtgaatcatc	tctgcaggcc	tgcacgggcg	ggcccgggtg	gggcaggggtg	6480
ctcccctttg	aaaaacgctg	cgccctgctc	ttcagcctca	gtttcccat	ctgtaaaagt	6540
tgcattccag	gcctccttgg	gctggcccag	agctgccctg	ggcaggggct	ttcagccctt	6600
cacaccagag	gctgggtcag	ctgagtga	gctcctttgt	cctcccggca	tgccccaggc	6660
tcctccctgc	tccccaggat	ccacacctac	cctggccccg	acctggggcc	atgcctcacc	6720
acacttctctg	tcttctctca	catctctcac	atctgggagt	cctctcccg	cagcctgtgc	6780
ttgcatcttg	ggatccatgc	cagggatagg	acctgtctc	cctgctggct	tgggacatgc	6840
cctctcccag	ccactctgag	gagggctgac	tgaggaaggg	ctcgtcaagc	tgggctttgc	6900
aggatgtgta	agagttctcc	ccacgagggc	cagggcattc	tgagccag	gaatggcttg	6960
tataagtagt	cagagtcatt	tgaatggcc	aagtagttgc	aggaatcgac	tggaaaccgg	7020
ggtggtaagg	tgaagccata	gaacattcca	ggccccctcc	cctaaatgag	atggaaggag	7080
tgctgtttt	gaacaagcca	gggcatctgg	ggaccgttaa	ggcctggggg	tgggtgatggg	7140
gactggagg	tgtgaggcaa	ccaggggggtg	cccctctgag	ccaccacaga	cagaatggtt	7200
cagaaggcca	ggcacactgg	ctctcgctg	taatccaagc	actttgggag	gccaaggagg	7260
gaagatggct	tggccaagg	agggaagatg	gcttcagccc	aggagttaa	ggccagcctg	7320
gaaaacatgg	caaaacctct	gtctacaaaa	aatacaaaaa	attagccagg	catggtgatg	7380
cacgcctgtc	gtcccagcta	ctcaggaggc	ttagatggga	ggatcacttg	agccggggag	7440
tttgaggctg	cagttagccg	agatcgtgcc	actgcactcc	agcctgggca	acagaacgag	7500
acctgtctt	aaaaaaat	tttttggcct	ggtgcgggtg	ctcacgcctg	taatgccagc	7560
actttgggag	tccaagggtg	gtggatcacc	tggaggtcgg	gagtttgaga	ccagtctgac	7620
caacatggag	aaaccccctc	tctactaaaa	acacaaaaat	tagccggggc	tgggtggcgca	7680
tgctgtaat	cccagctact	ctggaggctg	aggcaggaga	atcccttgaa	cccgggaggc	7740
ggaagtgtca	gtgagccgag	atcgacccat	tgcactccag	cctgggcaac	aagagtga	7800
ctccatctca	aaaaacaaaa	aaaaacaaaa	aaaaacaaaa	aaacaaacaa	acaaaaaact	7860
gtaattaaaa	taattaaaag	aaaaaagtta	aaaaaaaaaa	aaagagagga	tggctcaaag	7920
gctggaaaaca	gggaaggcca	ctgtgagggg	aggacaggca	gggctagacc	tctctggaag	7980
gaggagggag	aggtaaccgt	gggcccggca	caggagactc	ttaatctcct	ggctcccggg	8040
ggcttctgtg	ggcctgtgac	tcaggatttc	tgtgcctctg	ttgatgaaaa	gaaaaatcct	8100
gaagaaaaca	ctgttcccat	agtaacacag	ctctaattag	agactccaag	gccaagcgag	8160
ggccttggag	ccaggagggg	ccctgctcta	ctcgggggac	gacacccctt	tcccctatcc	8220
ctccctgctg	gggtgtggcc	tgatgtctac	gtggcaccag	gcccctaaat	cctttcaaca	8280
tcctggcgga	gcggagaggc	tgggcccctt	tgcaggtgg	ggaaactgag	gaagggtgca	8340
gggaaggcac	tcatccaagg	gtgcctggcc	ccccgogggtg	ggtgcctcac	tggctggcct	8400
aaacctttgc	atcaaaccag	ttccaggatt	gaacgcaaca	ggctgatggg	gactgagtcg	8460
tggtagacat	ggcggcagtt	gtgtggccct	gggcccagga	atggctggga	agctcccttt	8520
cttctctggc	ttgggggatg	aggttagtgt	aaatttcatg	gagttgccag	gaggattaa	8580
gccgacgtgc	attcacacgt	ggccccgtgg	ctgcgcag	ttaagccctg	cgcagccatt	8640
ggaagtctca	ttgagaagga	gcctcccggc	ttcctgcatt	tgttccaggc	gggctggtaa	8700
gcgcctgtct	tatttgcaag	gctgtgtgag	gacgagccct	gtaaacccta	ggacgaggct	8760
gtgctaagt	ggtttctgtt	ccactgtctc	cccttctcct	tttgccctc	tttctgggga	8820
acctggagcc	cctgatgccc	tctgtttttt	cctgtctgctg	gggtggacct	agaacctctc	8880
ctccatcct	ttctgtgct	gggggtggcc	tgcagaagac	tgtcctctta	cctgagtcct	8940
taacttccct	ccgatcagcg	ggggcttcag	cggagcccca	gcgcattgga	cagggtggtgc	9000
tctgtggccg	gcacaggcca	gggcccagg	tgcggccct	gcctgtaggt	agacatgcag	9060
gcccttctga	gcagaatgag	ggggcggtga	gggcagctgt	gcgggtgtgg	ctggcctgac	9120
ggctccctgc	agccctgcgg	cttctgtgca	gtgggggtgg	ggcggggccag	acctttgcag	9180
ggcctcttgg	ctgggtggag	gctgagtga	cagaggcccc	agccctcgtg	ttccctgggg	9240
tgtggcctta	ggatggtccc	ggcacccctg	ggaccagcc	cctgctcacc	ttatcctcct	9300
ccgtgcttg	gctggggtgt	ggtggctgag	ccagcccacc	cagtgcgggtg	gatgctcgtc	9360

ggagggcaga	gggtgggcta	cgggcttaag	gggccccagg	gaacctgggg	ggtggagccc	9420
aaggggttcc	aagaaggggc	ggggccagga	acccatagac	aagaggggtg	ggcagggaaa	9480
gggctgtgca	agagggccac	ccacctccgg	tccagagagc	cagccagacc	cttactttct	9540
cttgctgtgt	gaccttaggc	aaggacatgc	caccaccggc	cttagtctcc	ctctttgtga	9600
agtaggatga	agatccccac	ctgtggagca	gatgtggggc	tgcattggtg	ggtcaggatg	9660
gaacaggatg	gggaggccgg	gcgtgggtgc	ttaccctgt	aatcccagca	ctttgggagg	9720
caaaggtggg	aggactgctt	gagtcctgga	gtttgggcaa	catagcgaga	ccacccccat	9780
ctctacaaaa	taatgttaaa	gttagccagg	tgtagtgggt	agtgcctgtg	gtcccaccta	9840
ctggggagac	tgaggcaaga	ggatcctttg	agcccaggag	gtggaggctg	tagagagcca	9900
tgcttggggc	acttgcactc	cagcctgggc	aacagagcaa	gaccctatct	ctaaaaaaa	9960
aatggaattg	aggatgtgtc	ctctgggtgt	gggctctacc	acccaccagc	ctccctctcc	10020
tgtgggtgct	gcttgccacc	caccgctgag	gtccctgcag	catcaagggt	cccaaacgaa	10080
gacactctcc	agctctgagc	caggcaccca	tacagagccc	ggcagacgcc	tctgctgctg	10140
cagttcttag	aatccagagc	gcgtggggaa	tgtgaatttg	tgtgctgga	gccaaacatc	10200
ccacctccag	gttttagctg	gaggaatcca	tgtggatatg	caagcaatgt	agttataatg	10260
aatagcaatt	gtgggctggg	cgcagtggct	catacctgta	atcccagcac	tttgggaggc	10320
tgaggtgggt	ggatcacctg	aggtcaggag	tttgagaccg	gcctggccac	ataatgaaac	10380
tccgtctcta	ctaaaaatac	aaaaaaatta	gccgggcata	gtggcggtg	cctgtaatcc	10440
cagctactca	ggaggctgag	gcaggagaat	cacttgaaac	ccaggaggca	gaggtttcag	10500
tgaggcgaga	tctgtccatt	gcactccagc	ccgggagaca	cagcgagact	ctctttctca	10560
aaaaaagaat	agcagctttg	ctatagctta	ggggagcaaa	acccgaaagc	caccttaagt	10620
ttcccgaaga	tagagatccc	tggggtccta	ttctgagaca	gagtcttgct	ctgtcgccca	10680
ggctggagtg	catggcgga	tctcagctca	ctgcagctc	cgcctctgg	gttcaagcaa	10740
ttctcctgcg	tcagcctccc	gattacaggc	'gcccggcacc	acccccagct	10800	
aatttttttt	tttttgagac	agagtctcgc	attgtcgccc	aggctggagt	gcagtggcgc	10860
gatcttggct	cactgcaagc	tccacctccc	gggttcacgc	cattctcctg	cctcagcctc	10920
ccgagtagca	gggactacag	gcgcccgcga	ccgtgcccgg	ctaatttttt	ttttttgtat	10980
tattaataga	gacgggtttc	accgtgttag	gatgatctcg	atctcctgac	ctcgtgatcc	11040
gcctgcctcg	gcctcccaaa	gtgctgggat	tacaggcatg	agccaccttg	cctggctcctt	11100
ttttttgtat	tttttagtaga	gatgaggttt	caccatgttg	gccaggctgg	tctaaactcc	11160
tgacctcatg	atccgcccgc	ctcggcctcc	caaagtgcgt	ggattacagg	tgtggggcac	11220
tgtgcccggc	cagtcccatt	ctacttttta	gggacatgga	atcatgtaaa	gatgccatac	11280
agaaaacagt	atgcaaatac	agagacagag	gacatctgtg	catgtttctt	tctttcttcc	11340
tttttttttt	tttttttttt	tttgagatgg	attcttgcct	tgtcgcccag	gctggagtgc	11400
agtgggtgaa	tctcagctca	ctgcaacctc	tgcctcccag	gttcaagtga	ttctcctgcc	11460
tcagacttcc	aagtagctgg	gattacaggc	atgccccacc	atgccagct	aatttttgta	11520
tttttagtag	aggcggggtt	tcaccatgtt	ggcaaggctg	gtcttgaact	catgacctca	11580
ggtgatccac	ccgcctcagc	ttcccaaagt	gctgggatta	caggcgtgag	ccgcctgggc	11640
aaccgcgatg	cttcttatgt	ataagaaaaa	aacagctaga	aagtgtgtgg	ttcactggct	11700
gggcacatgg	tggcttgtgt	ctgtaacctc	agcactttgg	gaggctgagt	taggaggatc	11760
atgtgaggcc	gggagttcaa	gaccatccta	ggcaacatag	ccagaccctg	tctgtaccaa	11820
atagaaaaaa	aaattaactt	gatgtgggtg	tgtgtgcctg	tagtctcagc	tatttgggag	11880
gctgaggtgg	gaggatcact	tgagcccagg	aagtgcaggc	tacagtgagc	tatgattgtg	11940
ccactgcact	ccaacctggg	caacagagca	agaccctgaa	tcaaaaagaa	agagatgtat	12000
ttgccaacac	aggggtgttc	tgcagtgggg	tgggaagggg	acgtgtagga	cttcatagtt	12060
tccttttttt	tgttttagag	acagggttta	aaaaaaaaatt	tttttttttt	ttttgaggcg	12120
gaatttcatt	cttggtgccc	aggctggagt	gatctcggct	cactgccatc	tccgcctccc	12180
aggttgaagt	gattctcctg	cctcagcctc	aggcaatttt	gtatttttag	tagagacggg	12240
gtgactccat	gttggtcagg	ctggtctcaa	actcccagac	tccgggtgatc	tgccctgcctt	12300
ggcctcccaa	actgctggga	ttacagggtg	gagccactgt	gcctggcccc	tttttttttt	12360
ttttaagtag	ggtcttgcta	agttgtccag	gctggtcttg	aactcctggg	ctcaaagtat	12420
cctcctgact	tggcctccca	aaaggctggg	attgcaggag	atatgagcca	ccgcacccgg	12480
cccttttata	agttttctaa	aggaaaggaa	catgtaccat	cgtgttttat	tttaattttta	12540
catagaagca	ggaaccaatg	catgaacccc	cacagacggc	cccccaagat	gcgctgttcc	12600
gctgtgctgc	agcccatatt	gtccttttct	ataccatttt	tggttgtatt	tgaagcaaat	12660
ccctgacatc	agggcatttc	atcccaaaat	acttccatct	gtgtttctaa	aaaaacgagg	12720
ccattttctt	atgtaaccac	aacgtgatca	acagtaatca	atgcttaaca	tctaattctg	12780
ggccacgtta	agaccctcct	gattggctca	agaatgtctt	ttgtgcagtt	ggtttcttga	12840
ttctttttgt	tttttttgag	acagggtctc	actgtgtcac	ccaggctgga	gtgcagtggg	12900
gcaatcacgg	ctcactgcag	cctcagcctc	ctgcctcagc	gatcctcctg	cctcagcctc	12960
ccaagtagct	gggaccacag	gtttgcacca	ccactcccag	ctaattaaaa	aaagtaattt	13020
gtagagacgg	gggtgggggtg	ggggctctgc	tatgttgccc	aggctgggtc	caaactccta	13080
gactcaaaca	atcctcctgc	cttggcctcc	caaagtgttg	cgatcacagg	catgagcccc	13140

tgccctggc	ccctattatt	tctctaactg	gggaaagtca	tgtgggaaca	gatgtagctt	13200
gccttggcct	ctgaccggcc	ctgcctgogt	tcctgggtgc	tctctgctgc	ttctcatgtg	13260
tgccactgtg	gggactcggc	cgcccacccc	accccggtgt	gttaccttgc	gtgtgtagtt	13320
ctgtgagctc	agggctatgg	tctgccagaa	ctagggggcg	tggggcccca	gtaccagccc	13380
aaggcctcct	ctctgcaggt	atggacctgt	ctgccaatca	ggacgaggag	accgaccagg	13440
agaccttcca	gctggagatc	gaccgcgaca	ccaaaaagtg	tgccttccgt	acccacacgg	13500
gcaagtactg	gacgctgacg	gccaccgggg	gcgtgcagtc	caccgcctcc	agcaagttag	13560
tgctcgcctc	ccacctgtca	ccgccccac	caccttgctt	gggtaccccc	gcctgacctt	13620
gtcccgccat	ccccaggaa	tgccagctgc	tactttgaca	tcgagtggcg	tgaccggcgc	13680
atcacactga	gggcgtccaa	tggcaagttt	gtgacctcca	agaagaatgg	gcagctggcc	13740
gcctcgggtg	agacagcagg	taacactaaa	gccccagttc	cctggagccg	tcctggagtc	13800
ctggagggtc	tggccatgcc	gtggtcactt	ggtagcccca	gccaaaggct	gctctgtgct	13860
gggcatcccc	ccggactggc	cccgactgtt	cctaccctgg	ggactgctgt	gtgacccag	13920
ctcctggccc	tcctctcttg	gtcaccccag	cctccacccc	actccctgcc	aggaggctca	13980
ctgactcccc	tctttctggg	acaggggact	cagagctctt	cctcatgaag	ctcatcaacc	14040
gccccatcat	cgtgttccgc	ggggagcatg	gcttcatcgg	ctgccgcaag	gtcacgggca	14100
ccctggacgc	caaccgctcc	agctatgacg	tcttccagct	ggagtccaac	gatggcgctt	14160
acaacatcaa	aggcaggttc	tcctgtgggc	agctgctggg	cagggaaacc	ctcggtcggg	14220
gctggggctca	gtgctgcggg	gagcgccctc	tgcattccaca	ctggaccctg	gcttggctca	14280
gggcatattc	aggccctaaa	gggacaggtg	tctgatggcc	accagggggc	tctgggatgc	14340
aagcagcccc	tttccctctt	gtctgtgtgg	ttggggggac	ttaccttgcc	cacctgacag	14400
agaggtgtgt	ggaggggaga	gcagggaggg	aaggagacca	ggaaaggagg	ggaggagagc	14460
aggggagggg	agagccggga	agagaggaga	gcaggggtgg	ggaggtttct	ggaaagggtg	14520
tgcaagggga	ggacgcgcct	cgggttatggg	actggagccc	cttcccagga	ggacccccaa	14580
caatccagag	gtgcctgtta	ggattcagaa	catggttttt	ttgtttgttt	ttttgagact	14640
cactccctca	ccctggctgg	acttgacagt	gcgttatctc	ggctcactgc	aacctctgcc	14700
tcctgggttc	aagtgtattc	cctgcctcag	cctcccaagt	agctgggatt	acaggtgtgc	14760
accaccacgc	ctggctaatt	tttatatttt	taaaatttat	tatttattta	ttttgagacc	14820
cagtctggag	tgcatggcg	ttatctcggc	tctctgcaac	ctctgcctcc	tgggttcaag	14880
cgattctcct	gcctcagcct	cccagtagtc	tgggactatg	tgtgggagcc	accatgcctg	14940
gctaattttt	ttgtattttt	catagagacg	ggtttcacca	tgttgtccag	gctggctctg	15000
aattcgtggc	ctcaagtgat	ccgcccacct	cagcctccca	cagtgtctgg	tttataggtg	15060
tgagccacca	caccgggcta	attgttttgt	attttttagta	gagacggagc	ttcactatgt	15120
tggcaaggct	ggctcgaact	cctgacctca	agtgatccgc	ccacctcagc	ctcccaaagt	15180
gctgggatta	caggcgtgag	ccactgcggc	cgagcagaac	acgttctagg	acccttgttc	15240
atgtgtccat	catggacagg	aggacgtgcg	ggccataggg	accctggctc	attccggagc	15300
cgggactgga	gggtggggcg	tcacctttgg	gaacacccgt	gcccaccctc	cgctgcccag	15360
ggtaggggtg	gggagccagg	ctttggggcc	cacttgataa	agtccctctc	ccagactcca	15420
caggcaataa	ctggacgggtg	ggcagtgaat	ccgcgggtcac	cagcagcggc	gacactcctg	15480
tggacttctt	cttcgagttc	tgcgactata	acaaggtggc	catcaagggtg	ggcggggcgt	15540
acctgaaggg	cgaccacgca	ggcgtcctga	aggcctcgcc	ggaaaccgtg	gaccccgctt	15600
cgctctggga	gtactagggc	cgcccgctcc	ttcccgccgc	ctgcccacat	ggcggtctct	15660
gccaaccctc	cctgctaacc	ccttctccgc	caggtgggct	ccagggcggg	aggcaagccc	15720
ccttgccctt	caaaactggaa	acccagagaa	aaacgggtgc	cccacctgtc	gcccctatgg	15780
actcccact	ctcccctccg	cccgggttcc	ctactccctt	cgggtcagcg	gctgcggcct	15840
ggccctggga	gggattttcag	atgcccctgc	cctcttgtct	gccacggggc	gagtctggca	15900
cctctttctt	ctgacctcag	acggctctga	gccttatctt	tctggaagcg	gctaaggggac	15960
ggttgggggc	tgggaccctg	ggcgtgtagt	gtaactggaa	tcttttgctt	ctcccagcca	16020
cctcctccca	gccccccagg	agagctgggc	acatgtccca	agcctgtcag	tggccctccc	16080
tgggtgactg	tccccgaaac	ccctgcttgg	gaaggggaagc	tgtcgggtgg	gctaggactg	16140
acccttgtgg	tgtttttttg	ggtggtggct	ggaaacagcc	cctctcccac	gtggcagagg	16200
ctcagcctgg	ctcccttccc	tggagcggca	gggcgtgacg	gccacagggt	ctgcccgcctg	16260
cacgttctgc	caagggtggg	gtggcggggc	ggtaggggtg	tggggggcgt	cttcctcctg	16320
tctctttcct	ttcacccctag	cctgactgga	agcagaaaat	gaccaaataca	gtattttttt	16380
taatgaaata	ttattgctgg	aggcgtccca	ggcaagcctg	gctgtagtag	cgagtgtatc	16440
ggcggggggc	gtctcagcac	cctccccagg	gggtgcattc	cagccccctc	tttccgtcct	16500
tcccgctccag	ccccagccct	gggcctgggc	tgccgacacc	tgggcccagag	cccctgctgt	16560
gattggtgct	ccctgggcct	cccgggtgga	tgaagccagg	cgtcgccccc	tccgggagcc	16620
ctgggggtgag	ccgcgggggc	ccccctgctg	cagacctccc	ccgtcccca	catgcatctc	16680
actctgggtg	tcttgggtct	ttattttttg	taagtgtcat	ttgtataact	ctaaacgccc	16740
atgatagtag	cttcaaactg	gaaatagcga	aataaaaataa	ctcagtctgc	agccccaggc	16800
cggcctgtgt	gtgtcttggg	gctgaggtgg	gtgggggggc	tgaggtgggt	gggagggtctg	16860
gcgggacagg	taggcgcctt	ggctccccag	ctcagtgtctg	ggagtgtgca	gtgggaggga	16920

ggccgtggct ccagtgggtg ctccggagct c

16951