



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 690 34 115 T2 2004.07.15**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 489 058 B1**

(21) Deutsches Aktenzeichen: **690 34 115.6**

(86) PCT-Aktenzeichen: **PCT/CA90/00267**

(96) Europäisches Aktenzeichen: **90 912 428.1**

(87) PCT-Veröffentlichungs-Nr.: **WO 91/002796**

(86) PCT-Anmeldetag: **20.08.1990**

(87) Veröffentlichungstag
der PCT-Anmeldung: **07.03.1991**

(97) Erstveröffentlichung durch das EPA: **10.06.1992**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **05.11.2003**

(47) Veröffentlichungstag im Patentblatt: **15.07.2004**

(51) Int Cl.7: **C12N 15/12**

**C12P 21/02, C07K 14/47, C12Q 1/68,
G01N 33/68, G01N 33/53, A61K 48/00,
A01K 67/027, C12P 21/08**

(30) Unionspriorität:

396894	22.08.1989	US
399945	24.08.1989	US
401609	31.08.1989	US

(73) Patentinhaber:

**HSC Research Development Corp., Toronto,
Ontario, CA; The Board of Regents Acting for and
on Behalf of the University of Michigan, Ann Arbor,
Mich., US**

(74) Vertreter:

**Grünecker, Kinkeldey, Stockmair &
Schwanhäusser, 80538 München**

(84) Benannte Vertragsstaaten:

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

(72) Erfinder:

**TSUI, Lap-Chee, Toronto, CA; RIORDAN, R., John,
Toronto, CA; COLLINS, S., Francis, Ann Arbor, US;
ROMMENS, M., Johanna, Willowdale, CA;
IANNUZZI, C., Michael, Ann Arbor, US; KEREM,
Bat-Sheva, Toronto, CA; DRUMM, L., Mitchell, Ann
Arbor, US; BUCHWALD, Manual, Toronto, CA**

(54) Bezeichnung: **CYSTIC FIBROSIS GEN**

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

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

Beschreibung

[0001] Die vorliegende Erfindung bezieht sich allgemein auf das Gen der zystischen Fibrose (ZF) und insbesondere auf das Identifizieren, Isolieren und Klonieren der DNA-Sequenz, die den normalen und mutanten ZF-Genen entsprechen, sowie auf ihre Transkripte und Genprodukte. Die vorliegende Erfindung bezieht sich auch auf Methoden für das Screenen auf und das Erfassen von ZF-Trägern, die ZF-Diagnose, das Screenen und die Diagnose der pränatalen ZF und die Gentherapie unter Verwendung rekombinanter Technologien und die Arzneimitteltherapie unter Verwendung der Informationen, die von der DNA, dem Protein und der Stoffwechselfunktion des Proteins abgeleitet werden.

[0002] Die zystische Fibrose (ZF) ist die häufigste, schwere, autosomale, rezessive, in der kaukasischen Bevölkerung vorkommende genetische Krankheit. Sie tritt bei ca. 1 unter 2000 Lebendgeburten in Nordamerika auf [Boat et al, *The Metabolic Basis of Inherited Disease*, sechste Ausgabe, Seite 2649–2680, McGraw Hill, NY (1989)]. Ungefähr jede zwanzigste Person ist ein Träger dieser Krankheit.

[0003] Obwohl die Krankheit zuerst in den späten dreißiger Jahren beschrieben worden ist, ist der grundlegende Defekt bisher unbekannt geblieben. Die Hauptsymptome der zystischen Fibrose umfassen die chronische Lungenerkrankung, die exokrine Pankreasinsuffizienz und erhöhte Schweißelektrolytspiegel. Den Symptomen nach ist die zystische Fibrose eine exokrine Krankheit. Obwohl in letzter Zeit Fortschritte bei der Analyse des Ionentransports durch die Apikalmembran des Epithels von Zellen von ZF-Patienten gemacht worden sind, ist es nicht klar, ob die anormale Regulierung von Chloridkanälen den primären Defekt der Krankheit darstellt. Aufgrund des fehlenden Wissens über den molekularen Mechanismus der Krankheit ist daher ein alternativer Ansatz gemacht worden, beim Versuch, die Natur des molekularen Defekts durch das direkte Klonieren des verantwortlichen Gens auf der Basis seiner Chromosomenposition besser zu verstehen.

[0004] Es besteht jedoch kein klarer Phänotyp, der einen Ansatz auf die genaue Natur der genetischen Basis der Krankheit hin steuert oder die Identifizierung des Gens der zystischen Fibrose gestattet. Die Natur des ZF-Defekts mit Bezug auf die genetischen Bevölkerungsdaten ist bisher nicht ohne weiteres offensichtlich geworden. Sowohl die Verbreitung der Krankheit als auch die klinische Heterogenität sind schon durch mehrere verschiedene Mechanismen erklärt worden: eine hohe Mutationsrate, einen heterozygoten Vorteil, den genetischen Drift, mehrere Loci und die Reproduktionskompensation.

[0005] Viele dieser Hypothesen können aufgrund des fehlenden Wissens bezüglich des grundlegenden Defekts nicht überprüft werden. Aus diesem Grund hat man sich bei alternativen Ansätzen für das Bestimmen und Charakterisieren des ZF-Gens auf den Versuch konzentriert, die Position des Gens durch Erbanalyse zu identifizieren.

[0006] In den 1950iger Jahren wurde die Analyse der Kopplung des ZF-Gens an antigene und Proteinmarker versucht, es ließen sich jedoch keine positiven Ergebnisse erzielen (Steinberg et al *Am. J. Hum. Genet.* 8: 162–176, (1956); Steinberg and Morton *Am. J. Hum. Genet.* 8: 177–189, (1956); Goodchild et al *J. Med. Genet.* 7: 417–419, 1976).

[0007] In letzter Zeit ist es möglich geworden, RFLPs zur Erleichterung der Kopplungsanalyse zu verwenden. Eine Kopplung eines RFLP-Markers an das ZF-Gen wurde zum ersten Mal im Jahre 1985 offenbart (Tsui et al. *Science* 230: 1054–1057, 1985), wobei eine Kopplung zwischen dem ZF-Gen und einem nicht charakterisierten Marker, nämlich DOCRI-917, festgestellt wurde. Dieser Zusammenhang wurde bei einer Analyse von 39 Familien an ZF leidenden Kindern festgestellt. Dabei wurde gezeigt, dass die Chromosomenposition zwar nicht bestimmt worden war, die Position des Gens der Krankheit jedoch auf ca. 1% des menschlichen Genoms oder ca. 30 Millionen Nucleotidbasenpaare beschränkt werden konnte.

[0008] Die Chromosomenposition der DOCRI-917-Sonde wurde unter Zuhilfenahme von Nagetier-Human-Hybridzelllinien, die verschiedene menschliche Chromosomenkomplemente enthalten, bestimmt. Es wurde gezeigt, dass DOCRI-917 (und deshalb das ZF-Gen) an das menschliche Chromosom 7 kartiert.

[0009] Weitere Studien der physischen und genetischen Verkopplung wurden durchgeführt, um zu versuchen, die Position des ZF-Gens zu bestimmen. Zengerling et al [*Am. J. Hum. Genet.* 40: 228–236 (1987)] beschreibt die Verwendung von somatischen Human-Maus-Zellhybriden, um weitere Einzelheiten des physischen Zusammenhangs zwischen dem ZF-Gen und den Markern zu erhalten, von denen bekannt ist, dass sie damit verknüpft sind. Diese Veröffentlichung zeigt, dass das ZF-Gen entweder der distalen Region der Bande q22 oder der proximalen Region der Bande g31 am Chromosom 7 zugeordnet werden kann.

[0010] Rommens et al [*Am. J. Hum. Genet.* 43: 645–663, (1988)] besprechen in genauen Einzelheiten das Isolieren vieler neuer 7g31-Sonden. Der dabei skizzierte Ansatz führte zum Isolieren von zwei neuen Sonden, nämlich D75122 und D75340, die nahe beieinander gelegen sind. Die Pulsfeldgelelektrophorese-Kartierung zeigt an, dass diese beiden RFLP-Marker zwischen zwei Markern liegen, von denen bekannt ist, dass sie das ZF-Gen flankieren, nämlich MET [White, R., Woodward S., Leppert M., et al. *Nature* 318: 382–384, (1985)] und D7S8 (Wainwright, B. J., Scrambler, P. J. und J. Schwidtko, *Nature* 318: 384–385 (1985)], d. h. sie liegen in der ZF-Genregion. Die Entdeckung dieser Marker bietet einen Ausgangspunkt für die Chromosomenwanderung und den Chromosomensprung.

[0011] Estivill et al. [Nature 326: 840–845 (1987)] offenbaren, dass die Position eines Kandidaten-cDNA-Gens bestimmt und dieses teilweise charakterisiert worden ist. Sie geben jedoch nicht die richtige Position des ZF-Gens an. In dieser Literaturangabe wird ein Kandidaten-cDNA-Gen stromabwärts von einer CpG-Insel beschrieben, wobei es sich um untermethylierte, an GZ-Nukleotid reiche Region stromaufwärts von vielen Vertebratengenen handelt. Die Chromosomenpositionierung des KandidatenLocus ist als XV2C-Region identifiziert worden. Diese Region ist in der Europäischen Patentanmeldung Nr. 88303645.1 beschrieben worden. In dieser tatsächlichen Region ist das ZF-Gen jedoch nicht enthalten.

[0012] Eine wesentliche Schwierigkeit beim Identifizieren des ZF-Gens ist durch den Mangel an zytologisch erfassbaren Chromosomen-Neuanordnungen oder – deletionen hervorgerufen worden, durch die alle vorherigen Erfolge beim Klonieren menschlicher Krankheitsgene durch das Wissen bezüglich der Kartenposition wesentlich erleichtert worden sind.

[0013] Derartige Neuanordnungen und Deletionen konnten zytologisch beobachtet werden und dadurch konnte eine physische Position an einem bestimmten Chromosom mit einer bestimmten Krankheit in Beziehung gesetzt werden. Des weiteren konnte diese zytologische Position mit einer Molekularposition in Verbindung gebracht werden auf der Basis einer bekannten Beziehung zwischen öffentlich verfügbaren DNA-Sonden und zytologisch sichtbaren Änderungen in den Chromosomen. Die Kenntnis der molekularen Position des Gens bezüglich einer bestimmten Krankheit würde das Klonieren und Sequenzieren dieses Gens durch routinemäßige Verfahren besonders dann gestatten, wenn das Genprodukt bekannt ist und der Kloniererfolg durch Immunoassay der Expressionsprodukte der klonierten Gene bestätigt werden kann.

[0014] Im Gegensatz dazu war im Stand der Technik weder die zytologische Position noch das Genprodukt des Gens der zystischen Fibrose bekannt. Durch die kürzliche Identifizierung von MET und D7S8, wobei es sich um Marker handelt, die das ZF-Gen flankieren, wobei jedoch ihre molekulare Position nicht bestimmt worden ist, haben die betreffenden Erfinder verschiedene neue Genklonierungsstrategien entwickelt, um dem ZF-Gen erfindungsgemäß nahe zu kommen. Die bei diesen Strategien angewendeten Methoden umfassen das Chromosomenspringen von den flankierenden Markern, das Klonieren von DNA-Fragmenten aus einer definierten physischen Region unter Anwendung der Pulsfeldgelelektrophorese, eine Kombination von Körperzellenhybrid- und molekularen Kloniertechniken, die zum Isolieren von DNA-Fragmenten von untermethylierten cPG-Inseln in der Nähe der ZF bestimmt sind, das Chromosomen-Mikrodissezieren und -klonieren und das Sättigungsklonieren einer großen Anzahl von DNA-Markern aus der 7g31-Region. Durch diese neuartigen Strategien waren die betreffenden Erfinder in der Lage, das Gen zu identifizieren, das für die zystische Fibrose verantwortlich ist, in den Fällen, in denen der Stand der Technik ungewisse oder in einem Fall sogar falsche Antworten lieferte.

[0015] Die Anwendung dieser genetischen und molekularen Klonierungsstrategien hat das Isolieren und cDNA-Klonieren des Gens der zystischen Fibrose auf der Basis seiner Chromosomenposition ohne den Vorteil der genomischen Neuanordnung als Wegzeiger gestattet. Das Identifizieren der normalen und mutanten Formen des ZF-Gens und der ZF-Genprodukte hat es gestattet, Screening- und diagnostische Tests auf ZF unter Verwendung von Nucleinsäuresonden und Antikörpern zum Genprodukt zu entwickeln. Durch die Wechselwirkung mit dem defekten Genprodukt und den Weg, auf dem dieses Genprodukt involviert ist, ist die Therapie durch normale Genproduktsupplementierung und Genmanipulation und – bereitstellung nun möglich.

[0016] Das Gen, das beim Krankheitsprozess der zystischen Fibrose involviert ist, das im Folgenden "ZF-Gen" genannt wird, und dessen funktionelle Äquivalente sind identifiziert, isoliert und cDNA-kloniert und dessen Transkripte und Genprodukte identifiziert und sequenziert worden. Eine Dreibasenpaar-Deletion, die zur Auslassung eines Phenylalaninrests im Genprodukt führt, wurde dahingehend bestimmt, dass es den Mutationen des ZF-Gens in ca. 70% der an ZF leidenden Patienten entspricht, wobei verschiedene Mutationen in den meisten, wenn nicht in allen verbleibenden Fällen eine Rolle spielen.

[0017] Im Zuge der Identifizierung und Sequenzierung des Gens und seiner Genprodukte können Nucleinsäuresonden und zum Genprodukt gebildete Antikörper in einer Reihe verschiedener Hybridisierungs- und immunologischer Assays zum Screenen und Erfassen des Vorliegens entweder eines normalen oder eines defekten ZF-Gens oder Genprodukts verwendet werden. Assaykitausrüstungen für ein derartiges Screenen und eine solche Diagnose können ebenfalls bereitgestellt werden.

[0018] Die Therapie des Patienten durch Supplementierung mit dem normalen Genprodukt, dessen Herstellung mit Hilfe genetischer und rekombinanter Techniken amplifiziert werden kann, oder dessen funktionellem Äquivalent ist nun eventuell möglich. Die Korrektur oder Modifikation des defekten Genprodukts durch Arzneimittelbehandlung ist nun eventuell möglich. Außerdem kann die zystische Fibrose durch Gentherapie durch Korrigieren des Gendefekts in situ oder mit Hilfe rekombinanter oder anderer Vehikel unter Bereitstellung einer DNA-Sequenz, die in der Lage ist, das normale Genprodukt an die Zellen des Patienten zu exprimieren, geheilt oder unter Kontrolle gehalten werden.

[0019] Einer Ausgestaltung der Erfindung gemäß wird ein gereinigtes DNA-Molekül bereitgestellt, das ein Gen der zystischen Fibrose umfasst, wobei das Gen eine DNA-Sequenz umfasst, die aus der Gruppe ausgewählt wird, die aus Folgenden besteht:

- (a) DNA-Sequenzen, die der DNA-Sequenz in **Fig. 1** von der Aminosäurerestposition 1 zur -position 1480 oder einer polymorphen Form der DNA-Sequenz entsprechen, in der die Nucleotidsubstitution in der DNA-Sequenz die wesentliche Funktion des dadurch codierten Polypeptids nicht beeinflusst;
- (b) DNA-Sequenzen, die für ein normales Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR), das die Sequenz der **Fig. 1** entsprechend von der Aminosäurerestposition 1 bis zur -position 1480 aufweist, oder eine polymorphe Form des Polypeptids codieren, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydrophathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst;
- (c) DNA-Sequenzen, die einem Fragment der DNA-Sequenz in **Fig. 1** entsprechen, das mindestens 16 sequentielle Nucleotide zwischen den Nucleotidsequenzpositionen 134 und 4573 umfasst, und die in der Lage sind, nach einem normalem oder einem defektem Gen oder Genprodukt der zystischen Fibrose zu suchen oder das Vorliegen derselben zu bestimmen, und
- (d) DNA-Sequenzen, die mindestens 16 sequentielle Nucleotide umfassen und für ein Fragment der Aminosäuresequenz der **Fig. 1** codieren, die in der Lage sind, nach einem normalem oder einem defektem Gen oder Genprodukt der zystischen Fibrose zu suchen oder das Vorliegen derselben zu bestimmen, und
- (e) DNA-Sequenzen, die für ein Epitop codieren, das durch mindestens 18 sequentielle Nucleotide in der DNA-Sequenz der **Fig. 1** zwischen den Aminosäurerestpositionen 1 und 1480 codiert ist.

[0020] Bei der DNA kann es sich beispielsweise um cDNA handeln.

[0021] Einer anderen Ausgestaltung der vorliegenden Erfindung gemäß wird ein gereinigtes DNA-Modul aus (a), (b) oder (c) wie oben definiert, bereitgestellt, das des Weiteren mindestens eine Mutation in der DNA-Sequenz enthält, die, falls sie in den Zellen des menschlichen Körpers exprimiert ist, mit einer geänderten Zellfunktion assoziiert ist, die mit der zystischen Fibrose in Beziehung steht, wobei die Mutation eine Deletion von 3 Nucleotiden einschließt, die für Phenylalanin in der Aminosäurerestposition 508 codieren.

[0022] Einer weiteren Ausgestaltung der Erfindung gemäß wird eine gereinigte Nucleinsäuresonde bereitgestellt, die eine DNA- oder RNA-Nucleotidsequenz umfasst, die der Sequenz den Merkmalen (c), (d) oder (e) gemäß, wie oben definiert, entspricht.

[0023] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein rekombinanter Kloniervektor bereitgestellt, der ein erfindungsgemäßes DNA-Molekül umfasst. Einer Ausgestaltung der Erfindung gemäß ist der Vektor funktionsfähig an eine Expressionskontrollsequenz in dem rekombinanten DNA-Molekül so verkoppelt, dass das normale ZFTR-Protein exprimiert werden kann oder, alternativ, mit der anderen ausgewählten mutanten DNA-Sequenz, so dass das mutante ZFTR-Polypeptid exprimiert werden kann. Die Expressionskontrollsequenz wird aus der Gruppe ausgewählt, die aus Sequenzen, die die Expression von Genen prokaryotischer oder eukaryotischer Zellen steuern, und ihren Viren und Kombinationen derselben besteht.

[0024] Bei einer anderen Ausgestaltung umfasst die Erfindung einen nichtmenschlichen Wirt, der mit einem erfindungsgemäßen Vektor transformiert wird.

[0025] Einer weiteren Ausgestaltung der Erfindung gemäß wird eine Methode für die Herstellung eines normalen oder mutanten Transmembrankonduktanzregulator-Polypeptids der zystischen Fibrose (ZFTR) bereitgestellt, die folgende Schritte umfasst:

- (a) Kultivieren einer Wirtszelle, die mit dem Vektor nach Anspruch 5 in einem Medium und unter Bedingungen transfiziert ist, die für die Expression eines normalen oder mutanten ZFTR-Polypeptids günstig sind, und
- (b) Isolieren des exprimierten normalen oder mutanten ZFTR-Polypeptids.

[0026] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein gereinigtes DNA-Molekül bereitgestellt, das eine DNA-Sequenz umfasst, die ein mutantes Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR), das eine Sequenz nach **Fig. 1** für Aminosäurerestpositionen 1 bis 1480 aufweist, oder eine polymorphe Form des Polypeptids codiert, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydrophathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, des Weiteren durch eine Deletion der drei Nucleotide gekennzeichnet, die Phenylalanin an der Aminosäurerestposition 508 codieren.

[0027] Einer weiteren Ausgestaltung der Erfindung gemäß wird gereinigtes Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR) bereitgestellt, dadurch gekennzeichnet, dass es eine Peptidmolmasse von 170.000 Dalton aufweist und aus der Gruppe ausgewählt wird, die aus Folgenden besteht:

- (a) einem normalen ZFTR-Polypeptid mit einer die Zellmembranionenkonduktanz beeinflussenden Aktivität, das die Sequenz nach **Fig. 1** für die Aminosäurepositionen 1 bis 1480 aufweist, oder eine polymorphe Form des normalen ZFTR-Polypeptids, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydrophathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, und

(b) einem mutanten ZFTR-Polypeptid mit einer zystischen Fibroseaktivität in menschlichen Zellen, das die Sequenz nach **Fig. 1** für die Aminosäurepositionen 1 bis 1480 aufweist, mit Ausnahme der Deletion von enylnalanin an der Aminosäurerestposition 508.

[0028] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein Polypeptid bereitgestellt, das durch die Expression einer erfindungsgemäßen DNA-Sequenz codiert ist, wobei das Polypeptid die immunologische oder biologische Aktivität eines normalen oder mutanten ZFTR-Polypeptids aufweist.

[0029] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein Verfahren bereitgestellt für das Isolieren des normalen oder mutanten erfindungsgemäßen ZFTR-Proteins aus das Protein enthaltenden Zellen, welches Verfahren folgende Schritte umfasst:

- (a) Löslichmachen des Proteins einer ausgewählten Zellmembran, in der entweder das normale oder das mutante ZFTR-Protein exprimiert ist, unter Bildung einer Lösung des ZFTR-Proteins,
- (b) Trennen des ZFTR-Proteins von der Lösung durch Kontaktieren der Lösung mit Antikörpern gegen das normale oder mutante ZFTR-Protein, wobei die Antikörper auf einem Substrat immobilisiert werden,
- (c) Spülen des Substrats zur Entfernung des Proteins, das nicht an den Antikörpern anhaftet,
- (d) Freisetzen des ZFTR-Proteins von den Antikörpern zur Isolierung des ZFTR-Proteins und
- (e) Reinigen des ZFTR-Proteins zur Entfernung von jeglichem anderem Mammalienprotein.

[0030] Einer weiteren Ausgestaltung der Erfindung gemäß wird eine Methode für das Screening einer Versuchsperson bereitgestellt, um zu bestimmen, ob diese Versuchsperson ein Träger der zystischen Fibrose (ZF) oder ZF-Patient ist, welche Methode folgende Schritte umfasst:

Durchführen einer Untersuchung für die Bestimmung, in einer einem Screenen zu unterwerfenden biologischen Probe, des Vorliegens von mindestens einem Mitglied der Gruppe, die aus Folgendem besteht: einem normalen ZF-Gen, normalen ZF-Genprodukten, einem mutanten ZF-Gen, mutanten ZF-Genprodukten und Mischungen derselben, einem normalen ZFTR-Polypeptid mit einer die Zelltransmembran-Ionenkonduktanz beeinflussenden Aktivität und der Sequenz nach **Fig. 1** für Aminosäurepositionen 1 bis 1480 oder einer polymorphen Form des normalen ZFTR-Polypeptids, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben, oder ihr hydropathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, einem mutanten ZFTR-Polypeptid mit einer zystischen Fibroseaktivität in menschlichen Zellen und der Sequenz nach **Fig. 1** für Aminosäurepositionen 1 bis 1480, mit der Ausnahme der Deletion von Phenylalanin in der Aminosäurerestposition 508, wobei es sich bei der Untersuchung um

- (i) eine Hybridisierung involvierende Untersuchung oder
 - (ii) eine immunologische Untersuchung handelt
- und/oder die unter Zuhilfenahme eines oder mehrerer der Folgenden erfolgt: Hybridisierung unter Zuhilfenahme spezifischer Olegonucleoide, des direkten DNA-Sequenzierens, von Restriktionsenzymen, der Diskriminierung auf der Basis der elektrophoretischen Mobilität in Gelen mit denaturierendem Mittel, RNase-Protektion, chemischer Spaltung oder einem durch Ligase-vermittelten Bestimmungsverfahren.

[0031] Der Hybridisierungsassay kann die Verwendung eines oder mehrerer erfindungsgemäßer Nucleotide umfassen, wobei keine zwei Nucleotide die gleichen sind, welche Nucleotide auf irgendeine herkömmliche Weise markiert sein können.

[0032] Ist die Assaymethode eine immunologische Methode und die biologische Probe umfasst ein normales Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR) der Versuchsperson und/oder ein mutantes ZFTR-Polypeptid der Versuchsperson: das normale ZFTR-Polypeptid, das die Sequenz nach **Fig. 1** für die Aminosäurepositionen 1 bis 1480 aufweist, oder eine polymorphe Form des Polypeptids, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydropathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, und das mutante ZFTR-Polypeptid, das die Sequenz nach **Fig. 1** für die Aminosäurepositionen 1 bis 1480 aufweist, mit Ausnahme der Deletion von Phenylalanin an der Aminosäurerestposition 508.

[0033] Die immunologische Assaymethode kann Antikörper umfassen, die für ein normales und/oder ein mutantes ZFTR-Polypeptid spezifisch sind. Die bei der Methode verwendeten Antikörper können monoklonal sein.

[0034] Einer anderen Ausgestaltung der Erfindung gemäß wird die Methode für das Screenen einer Versuchsperson der Erfindung gemäß an einem Fötus in utero durchgeführt.

[0035] Einer weiteren Ausgestaltung der Erfindung gemäß, ein Verfahren für das Screening eines potentiellen Trägers der zystischen Fibrose (ZF) oder eines an zystischer Fibrose leidenden Patienten zur Identifizierung des Vorliegens einer identifizierten zystischen Fibrosemutation im ZF-Gen, das für ein Transmembrankonduktanzregulator-Polypeptid für zystische Fibrose (ZFTR) codiert mit der Sequenz nach **Fig. 1** für Aminosäureposition 1 bis Position 1480 oder eine polymorphe Form des Polypeptids aufweist, in der eine Aminosäuresubstitution in verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren

derselben oder ihr hydrophatisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, wobei das Verfahren folgende Schritte umfasst:

- (a) Das Hybridisieren einer DNA-Sonde nach Anspruch 3 auf die isolierte genome DNA vom potentiellen ZF-Träger oder dem potentiellen Patienten, wobei die DNA-Sonde die zystische Fibrosemutation in dem ZF-Gen umfasst, wobei die DNA-Sonde in der Lage ist, die zystische Fibrosemutation zu bestimmen und
- (b) Behandeln der genomischen DNA zur Bestimmung des Vorliegens oder der Abwesenheit der DNA-Sonde und Anzeigen des Vorliegens oder der Abwesenheit der zystischen Fibrosemutation einer vorbestimmten Art der Hybridisierung entsprechend.

[0036] Einer weiteren Ausgestaltung gemäß umfasst die Erfindung ein Verfahren für das Feststellen von Trägern der zystischen Fibrose oder an zystischer Fibrose leidenden Patienten, wobei das Verfahren aus Folgendem besteht: Kontaktieren eines ZF-Gens des Patienten oder Trägers mit einer Restriktionsendonuclease, Bestimmen des Vorliegens oder der Abwesenheit einer Restriktionsendonucleasestelle in dem ZF-Gen und Vergleichen des Musters der Restriktionsendonucleasestellen in dem ZP-Gen mit einem normalen ZF-Gen, das der Nucleotidsequenz entspricht, die für das Polypeptid zwischen dem Aminosäurerest in Position 1 und 1480 der **Fig. 1** codiert, wobei das Vorliegen zusätzlicher Stellen am normalen ZF-Gen oder die Abwesenheit von Stellen im normalen ZF-Gen auf ein mutantes ZF-Gen hinweist.

[0037] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein Verfahren bereitgestellt für das Feststellen von Trägern der zystischen Fibrose, wobei das Verfahren aus Folgendem besteht: Bestimmen der Differentialmobilität von Heteroduplex-PKR-Produkten in Polyacrylamidgelen durch Insertionen oder Deletionen in dem mutanten zystischen Fibrose-(ZF)-Gen im Vergleich mit der normalen ZF-Gen, das der Nucleotidsequenz entspricht, die für das Polypeptid zwischen der Aminosäurerestposition 1 und 1480 der **Fig. 1** codiert.

[0038] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein Versuchskit bereitgestellt für das Untersuchen auf das Vorliegen eines normalen ZF-Gens, das der Nucleotidsequenz entspricht, die für das Polypeptid zwischen der Aminosäurerestposition 1 und 1480 der **Fig. 1** codiert, oder eines mutanten ZF-Gens, das mindestens eine drei Basenpaare umfassende Deletion im normalen ZF-Gen aufweist, welche Deletion zur Deletion eines Phenylalanins an der Aminosäurerestposition 508 der **Fig. 1** führt, durch einen Immunoassay, der Folgendes umfasst:

- (a) einen Antikörper, der spezifisch an ein Genprodukt des normalen oder mutanten ZF-Gens nach Anspruch 9 bindet,
- (b) ein Reagensmittel für das Feststellen der Bindung des Antikörpers an das Genprodukt und
- (c) wobei der Antikörper und das Reagensmittel in Mengen vorliegen, die zur Durchführung des Immunoassays effektiv sind.

[0039] Einer weiteren Ausgestaltung der Erfindung gemäß, Kit für das Untersuchen auf das Vorliegen eines normalen Gens der zystischen Fibrose (ZF), das der Nucleotidsequenz entspricht, die das Polypeptid zwischen der Aminosäurerestposition 1 und 1480 der **Fig. 1** codiert, oder eines mutanten ZF-Gens, das mindestens eine drei Basenpaare umfassende Deletion im normalen ZF-Gen aufweist, welche Deletion zur Deletion eines Phenylalanins an der Aminosäurerestposition 508 der **Fig. 1** führt, durch eine Hybridisierung, die Folgendes umfasst:

- (a) eine Oligonucleotidsonde nach Anspruch 3, die spezifisch an das normale oder das mutante ZF-Gen bindet,
- (b) ein Reagensmittel für das Feststellen der Hybridisierung der Oligonucleotidsonde an dem normalen oder dem mutanten ZF-Gen und
- (c) wobei die Sonde und das Reagensmittel in Mengen vorliegen, die zur Durchführung des Hybridisierungsassays effektiv sind.

[0040] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein immunologisch aktiver gegen die zystische Fibrose wirkender polyklonaler oder monoklonaler Transmembrankonduktanzregulator-(ZFTR)-Polypeptid-Antikörper, der für ein erfindungsgemäßes ZFTR-Polypeptid spezifisch ist, bereitgestellt. Ein den erfindungsgemäßen monoklonalen Antikörper bildendes Hybridom ist eine weitere Ausgestaltung der Erfindung.

[0041] Einer weiteren Ausgestaltung gemäß umfasst die Erfindung eine Zusammensetzung für die Behandlung der zystischen Fibrose bei einem Patienten, die eine therapeutisch wirksame Menge eines gereinigten normalen erfindungsgemäßen ZFTR-Polypeptids umfasst. Der Träger kann ein Lungentensidprotein umfassen, um das Aufbringen der Zusammensetzung auf respiratorische Epithelzellen zu erleichtern.

[0042] Einer weiteren Ausgestaltung der Erfindung gemäß wird eine Zusammensetzung für die Gentherapie bei zystischer Fibrose bereitgestellt, die ein erfindungsgemäßes DNA-Molekül und ein Vehikel für die Verabreichung des DNA-Moleküls an eine Zelle eines an zystischer Fibrose leidenden Patienten umfasst. Bei dem Vehikel kann es sich beispielsweise um einen rekombinanten Vektor handeln.

[0043] Einer weiteren Ausgestaltung der Erfindung gemäß wird ein nichtmenschliches Tier bereitgestellt, das

ein heterologes Zellsystem umfasst, das einen erfindungsgemäßen rekombinanten Kloniervektor umfasst, welcher Vektor die Symptome der zystischen Fibrose in dem Tier induziert.

[0044] Einer weiteren Ausgestaltung der Erfindung gemäß, transgene Maus, die die Symptome der zystischen Fibrose aufweist, die durch die Schritte erzielt werden können, die Folgendes umfassen: Einführen eines zystischen Fibrosegens nach Anspruch 2, das für ein mutantes Transmembrankonduktanzregulatorprotein der zystischen Fibrose codiert, welches mindestens eine Deletion von Phenylalanin in der 508 in **Fig. 1** entsprechenden Position enthält, in einen Mausoozyten oder einen Mausembryo und gegebenenfalls Inaktivieren des endogenen Mausgens für zystische Fibrose.

[0045] Um die Erfindung besser zu verstehen, wird nun Bezug genommen auf die beiliegenden Zeichnungen in denen:

[0046] **Fig. 1** die Nucleotidsequenz des ZF-Gens und die Aminosäuresequenz des ZFTR-Proteins darstellt;

[0047] **Fig. 2** eine Restriktionskarte des ZF-Gens und der zur Chromosomenwanderung und zum Chromosomensprung zum Gen verwendeten schematischen Strategie zeigt;

[0048] **Fig. 3** eine Pulsfeldgelelektrophoresekarte der Region zeigt, die das ZF-Gen umfasst und es umgibt;

[0049] Die **Fig. 4A, 4B und 4C** die Erfassung konservierter Nucleotidsequenzen durch Kreuzspezieshybridisierung zeigen;

[0050] **Fig. 4D** eine Restriktionskarte sich überschneidender Segmente der Sonden E4.3 und H1.6 zeigt;

[0051] **Fig. 5** eine RNA-Blot-Hybridisierungsanalyse unter Verwendung von genomischen und cDNA-Sonden zeigt. Es wird die Hybridisierung zur Fibroblast-, (normalen und ZF-)Trachea-, Bauchspeicheldrüse-, Leber-, HL60-, T84- und Hirn-RNA gezeigt;

[0052] **Fig. 6** den Methylierungszustand der klonierten E4.3-Region am 5'-Ende des ZF-Gens zeigt;

[0053] **Fig. 7** eine Restriktionskarte der ZFTR-cDNA, die die Ausrichtung der cDNA auf die genomischen DNA-Fragmente zeigt;

[0054] **Fig. 8** eine RNA-Gelblotanalyse, die die Hybridisierung durch einen Teil der ZFTR-cDNA (Klon 10-1) an ein 6,5 kb-mRNA-Transkript in verschiedenen menschlichen Geweben zeigt;

[0055] **Fig. 9** eine DNA-Blothybridisierungsanalyse zeigt, die die Hybridisierung durch die ZFTR-cDNA-Klone zur genomischen, mit EcoRI und Hind III digerierten DNA veranschaulicht;

[0056] **Fig. 10** ein Primerextensionsexperiment zeigt, das die 5'- und 3'-Ende der ZFTR-cDNA charakterisiert;

[0057] **Fig. 11** ein Hydropathieprofil und die vorausgesagte Sekundärstrukturen des ZFTR zeigt;

[0058] **Fig. 12** eine Dot-Matrixanalyse interner Homologien im vorausgesagten ZFTR-Polypeptid zeigt;

[0059] **Fig. 13** ein schematisches Modell des vorausgesagten ZFTR-Proteins zeigt;

[0060] **Fig. 14** eine schematische Darstellung der Restriktionsfragmentlängen Polymorphismen (RFLP) zeigt, die mit dem ZF-Gen eng verknüpft sind, wobei das umgekehrte Dreieck die Position der F508 Drei-Basenpaardeletion anzeigt;

[0061] **Fig. 15** die Erfassung der F508 Mutation durch Oligonucleotidhybridisierung darstellt, wobei die Sonde N die normale Sequenz und die Sonde F die mutante ZF-Sequenz erfasst;

[0062] **Fig. 16** die Ausrichtung der am besten erhaltenen Segmente der verlängerten NBF des ZFTR mit vergleichbaren Regionen anderer Proteine darstellt;

[0063] **Fig. 17** die DNA-Sequenz um die F508 Deletion herum zeigt;

[0064] **Fig. 18** eine Darstellung des Nucleotidsequenziergels zeigt, die die DNA-Sequenz an der F508-Deletion zeigt;

[0065] **Fig. 19a und 19b** mit Coomassie-blau gefärbte Polyacrylamidgels auf die Elektrophorese von Protein aus Bakterienlysaten (JM101) hin zeigen, welche Bakterien durch die pGEX-Plasmide transformiert worden sind;

[0066] **Fig. 20** Immunblots von Bakterienlysaten zeigen, die das Fusionsprotein #1 (Tabelle 8) enthalten, mit Vorimmun- und Immunsere aus zwei verschiedenen Kaninchen;

[0067] **Fig. 21** ein Immunblot von T-84-Membranen unter Verwendung von Immunsere aus dem Kaninchen #1 aus **Fig. 20** zeigt; und

[0068] **Fig. 22** Immundotblots zeigt, die mit Vorimmun- und Immunsere aus einem Kaninchen sondiert sind, das mit dem KLH-Konjugat des Peptids #2 aus Tabelle 8 immunisiert worden ist.

[0069] Um die Erfindung noch besser verstehen zu können, werden nun Ausführungsformen derselben abschließend zur Veranschaulichung beschrieben.

1. DEFINITIONEN

[0070] Um eine Besprechung der verschiedenen Ausführungsformen der Erfindung und ein Verständnis der verschiedenen Elemente und Bestandteile, die beim Ausführen der Erfindung und Verwenden derselben verwendet werden, zu erleichtern, ist die folgende Definition von in der Beschreibung der Erfindung verwendeten Begriffen wie folgt:

ZF – zystische Fibrose

ZF-Träger – eine Person, die gesund zu sein scheint, deren Chromosome ein mutantes ZF-Gen enthalten, das auf die Nachkommen der Person übertragen werden kann.

ZF-Patient – eine Person, die ein mutantes ZF-Gen an jedem Chromosom derart trägt, dass sie die klinischen Symptome der zystischen Fibrose aufweist.

ZF-Gen – das Gen, dessen mutante Formen mit der zystischen Fibrosekrankheit verbunden sind. Diese Definition ist so zu verstehen, dass sie die verschiedenen Sequenzpolymorphismen einschließt, die existieren, bei denen Nucleotidsubstitutionen in der Gensequenz die wesentliche Funktion des Genprodukts nicht beeinflussen. Dieser Begriff bezieht sich vor allem auf eine isolierte Codiersequenz, kann jedoch einige oder alle der flankierenden Regulatorelemente und/oder Introne einschließen.

ZF-PI – zystische Fibrose-Pankreasinsuffizienz, die größte klinische Untergruppe von an zystischer Fibrose leidenden Patienten, die durch eine mangelnde Pankreasexokrinfunktion gekennzeichnet sind.

ZF-PS – zystische Fibrose-Pankreasinsuffizienz, eine klinische Untergruppe von an zystischer Fibrose leidenden Patienten mit einer ausreichenden Pankreasexokrinfunktion für die normale Verdauung von Nahrung.

ZFTR – Transmembrankonduktanzregulatorprotein der zystischen Fibrose, das durch das ZF-Gen kodiert ist. Diese Definition schließt das Protein ein, wie es aus menschlichen oder tierischen Quellen isoliert wird, wie es durch rekombinante Organismen gebildet und wie es chemisch oder enzymatisch synthetisiert wird. Diese Definition ist so zu verstehen, dass sie die verschiedenen polymorphen Formen des Protein einschließt, wobei Aminosäuresubstitutionen in den verschiedenen Regionen der Sequenz das wesentliche Funktionieren des Proteins oder dessen hydropathisches Profil oder sekundäre oder tertiäre Struktur nicht beeinflusst.

[0071] Die DNA-Standardnomenklatur wird zum Identifizieren der Basen verwendet.

[0072] Intronfreie DNA – ein Stück der DNA, dem interne nichtkodierende Segmente, beispielsweise cDNA, fehlen.

[0073] IRP-Locussequenz – (mit Protoonkogen-int-1-verbunden), ein in der Nähe des ZF-Gen gelegenes Gen.

[0074] Mutantes ZFTR – ein Protein, das ZFTR bezüglich der primären, sekundären und tertiären Struktur äußerst analog ist, bei dem jedoch eine geringe Anzahl von Aminosäuresubstitutionen und/oder -deletionen und/oder -insertionen zu einer Beeinträchtigung von dessen wesentlicher Funktion führen, so dass Organismen, deren Epithelzellen mutantes ZFTR anstatt ZFTR exprimieren, die Symptome zystischer Fibrose aufweisen.

[0075] mZF – ein Mausgen, das dem menschlichen ZF-Gen ortholog ist.

[0076] NBFs – Nucleotid-(ATP)-Bindungsfalten

[0077] ORF – offenes Leseraster

[0078] PKR – Polymerasekettenreaktion

[0079] Protein – es wird eine aus Einzelbuchstaben bestehende Nomenklatur zum Identifizieren der Aminosäuren verwendet

[0080] R-Domäne – eine hoch geladene zytoplasmatische Domäne des ZFTR-Proteins

[0081] RSV – Rous-Sarkom-Virus

[0082] TP – Tensidprotein

[0083] RFLP – Restriktionsfragmentlängenpolymorphismus.

2. ISOLIEREN DES ZF-GENS

[0084] Unter Verwendung der Chromosomenwanderung, des Chromosomensprungs und der cDNA-Hybridisierung sind DNA-Sequenzen, die > 500 Kilobasenpaare (kb) umfassen, aus einer Region am langen Arm des menschlichen, das Gen der zystischen Fibrose (ZF) enthaltenden Chromosoms 7 isoliert worden. Mehrere transkribierte Sequenzen und konservierte Segmente sind in dieser Region identifiziert worden. Eine bzw. eines Derselben entspricht dem ZF-Gen und umfasst ca. 250 kb der genomischen DNA. Überlappende komplementäre DNA-(cDNA)-Klone sind aus Epithelzellbibliotheken mit einem genomischen, einen Teil des zystischen Fibrosegens enthaltenden DNA-Segment isoliert worden. Die Nucleotidsequenz der isolierten cDNA ist in **Fig. 1** gezeigt. In jeder Reihe der entsprechenden Sequenzen ist die untere Reihe eine der Standardnomenklatur der Nucleotidsequenz entsprechende Liste. Die obere Reihe in jeder entsprechenden Reihe von Sequenzen ist eine Standard-Einzelbuchstaben Nomenklatur für die dem jeweiligen Codon entsprechende Aminosäure.

[0085] Transkripte einer Größe von ca. 6.500 Nucleotiden sind in betroffenen Geweben an ZF leidenden Patienten bestimmbar. Auf der Basis der isolierten Nucleotidsequenz besteht das vorausgesagte Protein aus zwei ähnlichen Regionen, von denen jede eine erste Domäne enthält, die Eigenschaften aufweist, die mit der Membranassoziation konsistent sind, sowie eine zweite Domäne, vor der angenommen wird, dass sie bei der ATB-Bindung eine Rolle spielt.

[0086] Eine 3 bp-Deletion, die zur Auslassung eines Phenylalaninrests in der Mitte der ersten vorausgesagten

Nucleotidbindungsdomäne (Aminosäureposition 508 des ZF-Genprodukts) führt, ist bei ZF-Patienten bestimmt worden. Diese Mutation in der normalen DNA-Sequenz von **Fig. 1** entspricht ca. 70% der Mutationen in an zystischer Fibrose leidenden Patienten. Extendierte Haplotypdaten auf der Basis von DNA-Markersubstanzen, die mit dem putativen Krankheitsgen eng verkoppelt sind, weisen darauf hin, dass der Rest des Pools mutanter ZF-Gene aus multiplen unterschiedlichen Mutationen besteht. Ein kleiner Satz dieser letzteren mutanten Allele (ca. 8%) kann einer Subgruppe von Patienten, die eine Pankreassuffizienz aufweisen, eine residuelle Pankreasesokrinfunktion verleihen.

2.1 CHROMOSOMENWANDERUNG UND -SPRUNG

[0087] Große Mengen der DNA, die die D7S122- und D75340-Verkoppelungsregionen von Rommens et al, supra, umgeben, wurden auf Kandidatengensequenzen hin untersucht. Außer den herkömmlichen Chromosomenwanderungsmethoden wurden Chromosomensprungtechniken zum Beschleunigen des Untersuchungsvorgangs angewendet. Von jedem Sprungendpunkt konnte eine neue bidirektionelle Wanderung eingeleitet werden. Das Anhalten der Sequenzwanderung durch "unklonierbare" Bereiche, die oft im Säugergenom anzutreffen sind, könnte durch Chromosomensprung umgangen werden.

[0088] Die verwendete Chromosomensprungbibliothek ist früher schon beschrieben worden [Collins et al, Science 235, 1046 (1987); Ianuzzi et al, Am. J. Hum. Genet, 44, 695 (1989)]. Die ursprüngliche Bibliothek wurde aus präparativem Pulsfeldgel zubereitet und sollte EcoRI-Teilfragmente von 70–130 kb enthalten; darauf folgende, mit dieser Bibliothek gemachte Erfahrungen weisen darauf hin, dass kleinere Fragmente ebenfalls repräsentiert sind und Sprunggrößen von 25–110 kb sind angetroffen worden. Die Bibliothek wurde auf sup-Wirt-MC1061 plattiert und durch Standardverfahren einem Screenen unterzogen (Maniatis et al). Positive Klone wurden in PBR \square 23Ava subkloniert und der Anfang und das Ende des Sprungs durch EcoRI- und Ava I-Digestion identifiziert, wie bei Collins, Genome analysis: A practical approach (IRL, London, 1988) Seiten 73–94) beschrieben. Für jeden Klon wurde ein Fragment vom Ende des Sprungs untersucht, um dessen Position am Chromosom 7 zu bestätigen. Die angrenzende Chromosomenregion, über die Chromosomenwanderung und der Chromosomensprung stattfand, betrug ca. 250 kb. Die Richtung der Sprünge wurde durch sorgfältige Wahl von Sonden beeinflusst, wie von Collins et al. und Ianuzzi et al, supra, beschrieben. Die gesamte klonierte Region, einschließlich der Sequenzen, die durch Verwendung der ZF-Gen-cDNA isoliert werden, beträgt ca. 500 kb.

[0089] Die schematische Darstellung der Strategie für die Chromosomenwanderung und den Chromosomensprung ist in **Fig. 2** dargestellt. In dieser Figur sind die ZF-Genexone durch römische Zahlen angezeigt. Waagrechte Linien über der Karte zeigen Wanderungsschritte an, während die Bogen über der Karte Sprungsschritte anzeigen. Die Figur geht von links nach rechts in jedem der sechs Ränge, wobei die Enden in der Richtung auf 7 cen und 7 qter, wie gezeigt, hinlaufen. Die Restriktionskarte für die Enzyme EcoRI, HindIII und BamHI ist über der ununterbrochenen Linie gezeigt und überspannt die gesamte klonierte Region. Restriktionsstellen, die durch Pfeile anstatt senkrechte Linien angezeigt sind, zeigen Stellen an, die nicht eindeutig positioniert worden sind. Zusätzliche Restriktionsstellen für andere Enzyme sind unter der Linie gezeigt. Hohlstellen in der klonierten Region sind durch || angezeigt. Diese kommen nur in dem durch cDNA-Klone des ZF-Transkripts erfassten Teil vor. Auf der Basis der Pulsfeldkartierung der Region ist es unwahrscheinlich, dass es sich dabei um große Hohlstellen handelt. Die wandernden Klone, wie sie durch waagrechte Pfeile über der Karte angezeigt sind, gehen in Richtung des Pfeils, der den bei jedem Klon erzielten Wanderungsfortschritt anzeigt. Cosmide Klone beginnen mit dem Buchstaben c, alle anderen Klone sind Phagen. Der Cosmid CF26 hat sich als Chimäre erwiesen; der gestrichelte Teil ist von einem anderen genomischen Fragment an einem anderen Chromosom abgeleitet. Die römischen Zahlen I bis XXIV zeigen die Position des Exons des ZF-Gens an. Die über der Linie gezeigten waagrecht Kästchen sind bei den Experimenten verwendete Sonden. Drei der Sonden stellen unabhängiges Subklonieren von vorher identifizierten Fragmenten zum Erfassen von Polymorphismen in dieser Region dar: H2.3A entspricht der Sonde XV2C (X. Estivill et al, Nature, 326: 840 (1987)), die Sonde E1 entspricht KN19 (Estivill supra) und die Sonde E4.1 entspricht Mp6d.9 (X. Estivill et al. Am. J. Hum. Genet. 44, 704 (1989)). G-2 ist ein Subfragment von E6, das eine transkribierte Sequenz erfasst. R161, R159 und R160 sind synthetische Oligonucleotide, die aus Teilen der IRP-Locussequenz konstruiert sind (B. Wainwright et al. EMBO J. 7: 1743 (1988)), was die Position dieses Transkripts auf der genomischen Karte anzeigt.

[0090] Da die beiden unabhängig voneinander isolierten DNA-Markersubstanzen D7S122 (pH131) und D7S340 (TM58) nur ca. 10 kb voneinander entfernt waren (**Fig. 2**), gingen die Wanderungen und Sprünge im Wesentlichen von einem einzigen Punkt aus. Die Richtung des Wanderns und Springens mit Bezug auf MET und D7S8 wurde daraufhin nachgewiesen unter Überqueren verschiedener Rohschnitt-Restriktionsendonuclease-Erkennungsstellen (wie beispielsweise denenigeri für Xho I, Nru I und Not I, man vergleiche **Fig. 2**) und mit Bezug auf die physikalische Langstreckenkarte von J. M. Rommens et al. Am. J. Hum. Genet., im Druck; A. M. Poustka et al. Genomics 2, 337 (1988); M. L. Drumm et al. Genomics 2, 346 (1988). Die Pulsfeld-Kartierungsdaten haben auch gezeigt, dass die NOT-I-Stelle, die durch die Erfinder der vorliegenden Erfindung iden-

tifiziert worden ist (man vergleiche **Fig. 2**, Position 113 kb), derjenigen entsprach, bei der vorher festgestellt worden ist, dass sie mit den IRP-Locus assoziiert ist (Estiville et al, 1987 supra). Da darauffolgende genetische Studien gezeigt haben, dass ZF wahrscheinlich zwischen IRP und D788 gelegen war [M. Farrall et al., *Am. J. Hum. Genet.*, 43, 471 (1988), B.-S. Keran et al. *Am. J. Hum. Genet.*, 44, 827 (1989)], wurden die Wanderungs- und Sprunganstrengungen ausschließlich auf das Klonieren dieses Intervalls hin fortgesetzt. Man wird sich jedoch im Klaren darüber sein, dass andere Kodierungsregionen, wie sie in **Fig. 2** identifiziert sind, beispielsweise G-2, ZF14 und ZF16, positioniert und eingehend untersucht worden sind. Derart eingehende Untersuchungen dieser anderen Regionen haben gezeigt, dass sie auf der Basis genetischer Daten und der Sequenzanalyse nicht das ZF-Gen darstellten. Angesichts des Mangels an Kenntnissen bezüglich der Position des ZF-Gens und seiner Charakteristik hat die umfangreiche und zeitraubende Untersuchung der mutmaßlichen, sich in der Nähe befindenden Kodierungsregionen die Ausrichtung der Suche nach dem ZF-Gen nicht vorangetrieben. Diese Untersuchungen waren notwendig, um die Möglichkeit auszuschalten, dass das ZF-Gen in diesen Regionen vorliegt.

[0091] Es ist festgestellt worden, dass drei Regionen in 280 kb-Segment in den ursprünglich verwendeten amplifizierten genomischen Bibliotheken nicht ohne weiteres gewonnen werden konnten. Diese weniger klonierbaren Regionen befanden sich in der Nähe der DNA-Segmente H2.3A und X.6 und nur ein wenig über dem Cosmid cW44 hinaus an den Positionen 75–100 kb, 205–225 kb bzw. 275–285 kb in **Fig. 2**. Es wurde festgestellt, dass die rekombinanten Klone in der Nähe von H2.3A sehr unbeständig sind, wobei nach nur ein paar Durchgängen von Bakterienkulturen drastische Umgruppierungen stattfanden. Um die dabei entstehenden Leerstellen zu füllen, wurden primäre Wanderungsbibliotheken unter Zuhilfenahme spezieller Wirt-Vektorsysteme konstruiert, von denen berichtet worden ist, dass sie die Propagierung unbeständiger Sequenzen gestatten [A. R. Wyman, L. B. Wolfe, D. Botstein, *Proc. Nat. Acad. Sci. USA* 82, 2880 (1985); K. F. Wertman, A. R. Wyman, D. Botstein, *Gene* 49, 253 (1986); A. R. Wyman, K. F. Wertman, D. Barker, C. Helme, W. H. Petri, *Gene*, 49, 263 (1986)]. Obwohl die Region in der Nähe des Cosmids cW44 noch wiedergewonnen werden muss, wurde die Region in der Nähe von X.6 mit diesen Bibliotheken erfolgreich gerettet.

2.2 KONSTRUKTION VON GENOMISCHEN BIBLIOTHEKEN

[0092] Es wurden genomische Bibliotheken Verfahren entsprechend konstruiert, die von Maniatis, et al, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 1982) beschrieben und in Tabelle 1 aufgelistet sind. Diese umfassen acht Phagenbibliotheken, von denen eine von T. Maniatis (Fritsch et al, *Cell*, 19: 959 (1980)) bereitgestellt wurde; der Rest wurde als Teil dieser Arbeit von Maniatis et al, supra, beschriebenen Verfahren entsprechend konstruiert. Vier Phagenbibliotheken wurden in λ DASH (im Handel von Stratagene erhältlich) und drei in λ FIX (im Handel von Stratagene erhältlich) kloniert, wobei die Vektorarme von dem Hersteller bereitgestellt wurden. Eine λ DASH-Bibliothek wurde von teilweise mit Sau 3A digerierter DNA aus einem menschliches Chromosom 7 (4AF/102/K015) enthaltenden Human-Hamster-Hybrid (Rommens et al, *Am. J. Hum. Genet.* 43, 4 (1988)) und andere Bibliotheken aus der Teil-Sau 3A-, Total-BamHI- oder Total-EcoRI-Digestion von menschlichem peripherem Blut oder Lymphoblastoid-DNA konstruiert. Um einen Verlust an instabilen Sequenzen zu vermeiden, wurden fünf der Phagenbibliotheken auf den Rekombination defizitären Wirten DB1316 (recD-), CES 200 (recBC-) [Wyman et al, supra, Wertman et al supra, Wyman et al supra] oder TAP90 (Patterson et al *Nucleic Acids Res.* 15: 6298 (1987)) propagiert. Drei Cosmid-Bibliotheken wurden daraufhin konstruiert. Bei einer wurde der Vektor pCV108 (Lau et al *Proc. Natl. Acad. Sci. USA* 80: 5225 (1983) zum Klonieren von teilweise(durch Sau 3A) digerierter DNA aus 4AF/102/K015 verwendet [Rommens et al *Am. J. Hum. Genet.*, 43: 4 (1988)]. Eine zweite Cosmid-Bibliothek wurde durch Klonieren von teilweise (durch Mbo I) digerierter menschlicher Lymphoblastoid-DNA in den Vektor pWE-IL2R zubereitet, der durch Inserieren der durch RSV-(Rous-Sarkomvirus) Promotor getriebenen cDNA für die Interleukin-2-Rezeptor- α -Kette von M. Fordie und B. Howard geliefert) anstatt des Neorestistenz-Gens von pWE15 zubereitet wurde [Wahl et al *Proc. Natl. Acad. Sci. USA* 84: 2160 (1987)]. Eine zusätzliche Partial-Mbo-I-Cosmidbibliothek wurde im Vektor pWE-IL2-Sal zubereitet, der durch Inserieren eines Sal-I-Verkopplers in die Bam HI-Klonierstelle von pWE-EL2R geschaffen wurde (M. Drumm, unveröffentlichte Daten); das gestattet die Verwendung der teilweisen Eintragstechnik zum Ligieren der Sal-I- und Mbo-I-Enden unter Verhinderung von tandemartigen Insertionen (Zabarovsky et al *Gene* 42: 19 (1986)). Es wurden Cosmid-Bibliotheken in den *e. coli*-Wirtsstämmen DH1 oder 490A propagiert [Steinmetz, A. Winoto, K. Minard, L. Hood, *Cell* 28, 489 (1982)].

Tabelle 1
GENOMISCHE BIBLIOTHEKEN

<u>Vektor</u>	<u>Quelle menschlicher DNA</u>	<u>Wirt</u>	<u>Komplexität</u>	<u>Literatur-angabe</u>
λCharon 4A	Durch HaeII-AluI teilweise digerierte Gesamt-Humanleber-DNA	LE392	1 x 10 ⁶ (amplifiziert)	Lawn et al 19080
pCV108	Durch Sau3A teilweise digerierte DKI-DNA aus 4AF-KO15	DKI	3 x 10 ⁶ (amplifiziert)	
λdash	Durch Sau3A teilweise digerierte DNA aus 4AF-KO15	LE392	1 x 10 ⁶ (amplifiziert)	
λdash	Durch Sau3A teilweise digerierte Gesamthuman-Peripherieblut-DNA	DB1316	1,5 x 10 ⁶	
λdash	Durch BamHI teilweise digerierte Gesamthuman-Peripherieblut-DNA	DB1316	1,5 x 10 ⁶	
λdash	Durch EcoRI teilweise digerierte Gesamthuman-Peripherieblut-DNA	DB1316	8 x 10 ⁶	
λfix	Durch MboI teilweise digerierte Human-Lymphoblastoid-DNA	LE392	1,5 x 10 ⁶	
λFIX	Durch MboI teilweise digerierte Human-Lymphoblastoid-DNA	CE200	1,2 x 10 ⁶	
λFIX	Durch MboI teilweise digerierte Human-Lymphoblastoid-DNA	TAP90	1,3 x 10 ⁶	
pWE-IL2R	Durch MboI teilweise digerierte Human-Lymphoblastoid-DNA	490A	5 x 10 ⁶	
pWE-IL2R-Sal	Durch MboI teilweise digerierte Human-Lymphoblastoid-DNA	490A	1,2 x 10 ⁶	
GCh3A λlac (springend)	Durch EcoRI teilweise digerierte (24-110 kb) Human-Lymphoblastoid-DNA	MCI061	3 x 10 ⁶	Collins et al supra und Ianuzzi et al supra

[0093] Drei der Phagenbibliotheken wurden im E. coli Bakterienstamm LE392 propagiert und amplifiziert. Vier darauffolgende Bibliotheken wurden auf den an Rekombination defizitären Wirten DB1316 (recD⁻) oder CES200 (recB⁻) [Wyman 1985, supra; Wertman 1986, supra und Wyman 1986, supra] oder in einem Fall auf TAP90 [T. A. Petterson und M. Dean, Nucleic Acids Research 15, 6298 (1987)] plattiert.

[0094] Einzelkopie-DNA-Segmente (die von sich wiederholenden Elementen frei waren) in der Nähe der Enden jeder der Phagen- oder Cosmid-Insertion wurden gereinigt und als Sonden für das Bibliothekscreensen zum Isolieren sich überschneidender DNA-Fragmente durch Standardverfahren verwendet (Maniatis, et al,

supra).

[0095] $1-2 \times 10^6$ Phagenklone wurden auf 25–30 Petrischalen von 150 mm mit dem entsprechenden Indikator-Bakterienwirt plattiert und 10–16 h lang bei 37°C inkubiert. Für jede Platte wurden mit Nitrocellulose oder Nylonmembranen Duplikat-"Abhebungen" für jede Platte zubereitet, vorhybridisiert und unter den beschriebenen Bedingungen [Rommens et al, 1988, supra] hybridisiert. Es wurden Sonden mit ^{33}P auf eine spezifische Aktivität von $> 5 \times 10^8$ cpm/ μg unter Zuhilfenahme von Random-Startverfahren markiert [A. F. Feinberg und B. Vogelstein, Anal. Biochem. 132, 6 (1983)]. Die Cosmidbibliothek wurde auf Ampicillin-enthaltende Platten aufgebracht und auf ähnliche Weise einem Screenen unterzogen.

[0096] Es war oft möglich, DNA-Sonden, die hohe Hintergrundsignale abgeben, dadurch erfolgreicher zu verwenden, dass die gekochte Sonde vor Zugabe der Sonde zum Hybridisierungsbeutel 60 Minuten mit 250 $\mu\text{g}/\text{ml}$ einer Scherbeanspruchung unterzogener denaturierter Plazenta-DNA vor-reassoziert wurde.

[0097] Für jeden Wanderungsschritt wurde die Identität des klonierten DNA-Fragments durch Hybridisieren mit einem somatischen Zellenhybrid Abschnitt bestimmt, um seine Chromosomenposition zu bestätigen, sowie durch Restriktionskartierung und Southern-Blotanalyse, um seine Kolinearität mit dem Genom zu bestätigen.

[0098] Die gesamte kombinierte klonierte Region der isolierten genomischen DNA-Sequenzen und der überlappenden cDNA-Klone erstreckte sich über > 500 kb. Um sicherzustellen, dass die durch die Chromosomenwanderungs- und -sprungverfahren isolierten DNA-Segmente mit der genomischen Sequenz kolinear waren, wurde jedes Segment durch Folgendes untersucht:

- (a) eine Hybridisierungsanalyse mit somatischen Human-Nagetier-Hybridzelllinien, um die Position des Chromosoms 7 zu bestätigen,
- (b) eine Pulsfeldgelelektrophorese und
- (c) ein Vergleich der Restriktionskarte der klonierten DNA mit derjenigen der genomischen DNA.

[0099] Dementsprechend wurden Einzelkopie-Human-DNA-Sequenzen von jedem rekombinanten Phagen- und Cosmidklon isoliert und als Sonden bei jeder dieser Hybridisierungsanalysen, wie sie durch das Verfahren von Maniatis et al, supra, durchgeführt wurden, verwendet.

[0100] Während die Mehrzahl der Phagen- und Cosmidisolate die richtigen Wanderungs- und Sprungklone darstellten, gingen einige wenige aus Kloniertefakten und Kreuzhybridisierungssequenzen aus anderen Regionen im menschlichen Genom oder aus dem Hamstergenom in denjenigen Fällen hervor, in denen Bibliotheken aus einer Human-Hamster-Hybridzelllinie abgeleitet wurden. Die Bestätigung der richtigen Positionierung war bei Klonen, die durch Chromosomensprünge isoliert worden waren, besonders wichtig. Es wurden eine Vielzahl von Sprungklonen in Betracht gezogen und führten zu nichtschlüssigen Informationen, was die Richtung der Untersuchung vom Gen hinwegleitete.

2.3 BESTÄTIGUNG DER RESTRIKTIONSKARTE

[0101] Eine weitere Bestätigung der physikalischen Karte der überlappenden Klone insgesamt wurde durch Restriktionskartierungsanalyse über einen breiten Bereich unter Anwendung der Pulsfeldgelelektrophorese (J. M. Rommens et al. Am. J. Hum. Genet. im Druck, A. M. Pustka et al 1988, supra M. L. Drumm et al, 1988 supra erhalten.

[0102] Die **Fig. 3A bis 3E** veranschaulichen die Ergebnisse der Restriktionskartierungsstudie über eine breiten Bereich), wobei eine schematische Darstellung der Region in Abschnitt E angegeben ist. DNA aus der Human-Hamsterzelllinie 4AF/102/K015 wurde mit den Enzymen (A) Sal I, (B) Xho I, (C) Sfi I und (D) Nae I digeiert, durch Pulsfeldgelelektrophorese abgetrennt und auf Zetaprobe® (BioRad) übertragen. Für jedes Enzym wurde ein Einzel-Blot sequentiell mit den Sonden, die unter jedem Abschnitt der **Fig. A bis D** angezeigt sind, hybridisiert, wobei zwischen den Hybridisierungsvorgängen das Blot abgezogen wurden. Die Symbole für jedes Enzym der **Fig. 3E** sind: A, Nae I; B, Bse HII; F, Sfi I; L, Sal I; M, Mlu I; N, Not I; R, Nru I und X, Xho I. C entspricht der Kompressionszonenregion des Gels. DNA-Zubereitungen, Restriktionsdigestion und Kreuzfeldgelelektrophoresemethoden sind schon beschrieben worden (Rommens et al, im Druck, supra). Die Gele in **Fig. 3** wurden in 0,5 X TBE mit 7 Volt/cm für 20 Stunden gefahren wobei das Umschalten linear von 10–40 Sekunden für (A), (B) und (C) hochgefahren wurde, und mit 8 Volt/cm für 20 Stunden, wobei das Umschalten für (D) linear von 50–150 Sekunden hochgefahren wurde. Schematische Interpretationen des Hybridisierungsmusters sind unter jedem Abschnitt angegeben. Die Fragmentlängen sind in Kilobasen ausgedrückt und wurden durch Vergleich mit oligomelisierter Bakteriophagen- λ -DNA und Saccharomyces cerevisiae-Chromosomen großemäßig bestimmt.

[0103] H4.0, J44, EG1.4 sind genomische Sonden, die durch Wanderungs- und Sprungexperimente gebildet worden sind (Vergleiche **Fig. 2**). J30 ist durch vier aufeinanderfolgende Sprünge aus D7S8 isoliert worden (Collins et al, 1987, supra; Ianuzzi et al, 1989, supra; M. Dean, et al, zur Veröffentlichung unterbreitet). 10-1, B.75 und CE1.5/1.0 sind cDNA-Sonden, die sich über verschiedene Regionen des ZF-Transkripts erstrecken: 10-1 enthält die Exone I–VI, B.75 enthält die Exone V– XII und CE1.5/1.0 enthält die Exone XII–XXIV. In

Fig. 3E ist ein zusammengesetzte Karte des gesamten MET – D7S8-Intervalls gezeigt. Die umrandete Region zeigt das Segment an, das durch Wanderung und Springen kloniert worden ist und der schattierte Teil zeigt die Region an, über die sich der ZF-Transkript erstreckt. Die CpG-reiche Region, die mit dem D7S23-Locus assoziiert ist (Estivill et al, 1987, supra) ist an der in Klammern angegebenen Not I-Stelle angezeigt. Diese und andere in runden Klammer oder viereckigen Klammer angezeigten Stellen schneiden 4AF/102/K015 nicht, sind jedoch in menschlichen Lymphoblastzelllinien beobachtet worden.

2.4 IDENTIFIZIERUNG DES ZF-GENS

[0104] Auf der Basis der Ergebnisse der oben angegebenen Restriktionskartierung über einen großen Bereich ist bestimmt worden, dass das ZF-Gen als ganzes auf einem Sal I-Fragment von 380 kb enthalten ist.

[0105] Die Ausrichtung der Restriktionsstellen, die aus der Pulsfeldgenanalyse abgeleitet worden sind, zu denjenigen, die in den teilweise überlappenden genomischen DNA-Klonen identifiziert worden sind, zeigte, dass die Größe des ZF-Gens ca. 250 kb betrug.

[0106] Das aufschlussreichste Restriktionsenzym, das dem Ausrichten der Karte der klonierten DNA-Fragmente mit der Restriktionskarte über einen großen Bereich diente, war Xho I; alle 9 der mit den rekombinanten DNA-Klonen identifizierten Xho I-Stellen schienen für zumindest partielle Spaltung im genomischen DNA anfällig zu sein (man vergleiche die Karten in **Fig. 1** und **2**).

[0107] Des Weiteren identifizierte die Hybridisierungsanalyse mit Sonden, die vom 3'-Ende des ZF-Gens abgeleitet worden sind, 2 Sfi I-Stellen und bestätigte die Position einer vorausgesagten Nae I-Stelle.

[0108] Diese Befunde stützten die Schlussfolgerung weiterhin, dass die DNA-Segmente, die durch Chromosomenwanderung und -sprungtechniken isoliert worden waren, mit der wirklichen Sequenz kollinear waren.

2.5 KRITERIEN BEZÜGLICH DER IDENTIFIZIERUNG

[0109] Ein positives Ergebnis auf der Basis einer oder mehrerer der folgenden Kriterien weist darauf hin, dass ein kloniertes DNA-Segment eventuell Kandidatengensequenzen enthalten kann:

- (a) Erfassung von kreuzhybridisierten Sequenzen in anderen Spezies (denn viele Gene weisen eine evolutionäre Konservierung auf),
- (b) Identifizierung von CpG-Inseln, die oft das 5'-Ende von Vertebratengen markieren [A. F. Bird. Nature, 321, 209 (1986); M. Gardiner-Garden und M. Frommer, J. Mol. Biol. 196, 261 (1987)].
- (c) Untersuchung möglicher mRNA-Transkripte in betroffenen Geweben bei an ZF leidenden Patienten,
- (d) Isolierung entsprechender cDNA-Sequenzen,
- (d) Identifizierung offener Leseraster durch direkte Sequenzierung klonierter DNA-Segmente.

[0110] Die Kreuzspezieshybridisierung zeigte eine starke Sequenzkonservierung zwischen menschlicher und Rinder-DNA, wenn ZF14, E4.3 und H1.6 als Sonden verwendet wurden; die Ergebnisse derselben sind den **Fig. 4A, 4B** und **4C** gezeigt.

[0111] Genomische menschliche, Rinder-, Maus-, Hamster- und Hühner-DNAs wurden mit EcoRI (R), Hind III (H) und Pst I (P) digeriert, elektro-phoresiert und auf Zetabind® (BioRad) ein Blot hergestellt. Die Hybridisierungsverfahren nach Rommens et al, 1988, supra, wurden mit den bei 55°C strengsten Wäschen, 0.2X SSC und 0,1% SDS verwendet. Die für die Hybridisierung in **Fig. 4** verwendeten Sonden haben: (A) das gesamte Cosmid CF14, (B) E4.3, (C) H1.6 umfasst. Im Schema der **Fig. (D)** zeigt die schattierte Region den Bereich der Kreuzspezieskonservierung.

[0112] Die Tatsache, dass verschiedene Subgruppen von Banden in Rinder-DNA bei diesen beiden überlappenden DNA-Segmenten (H1.6 und E4.3) entdeckt worden sind, weist darauf hin, dass die konservierten Sequenzen an den Grenzen der überlappenden Region gelegen waren (**Fig. 4(D)**). Wenn diese DNA-Segmente zum Erfassen von RNA-Transkripten aus einer Reihe verschiedener Gewebe verwendet worden sind, wurde kein Hybridisierungssignal festgestellt. Beim Versuch, ein Verständnis der Kreuzhybridisierungsregion zu erhalten und mögliche offene Leseraster zu identifizieren, wurden DNA-Sequenzen des gesamten H1.6- und eines Teils des E4.3-Fragments bestimmt. Die Ergebnisse zeigen, dass mit Ausnahme einer lange Strecke der an CG-reichen Sequenz, die die Erkennungsstellen für zwei Restriktionsenzyme (Bss Hll und Sac II) enthält, die oft in Assoziation mit untermethylierten CpG-Inseln vorgefunden werden, nur kurze offene Leseraster vorlagen, die nicht ohne weiteres als Erklärung für die starken Kreuzspezies-Hybridisierungssignale dienen konnten.

[0113] Um den Methylierungsstatus dieser an CpG äußerst reichen Region, die durch Sequenzieren zu erkennen war, zu untersuchen, wurden Proben genomischer DNA, die aus Fibroblasten und Lymphoblasten zubereitet worden waren, mit den Restriktionsenzymen Hpa II und Msp I digeriert und durch Gelblothybridisierung analysiert. Das Enzym Hpa II schneidet die DNA-Sequenz 5'-CCGG-3' nur, wenn das zweite Cytosin unmethyliert ist, während Msp I diese Sequenz schneidet, gleichgültig, welcher Methylierungszustand vorliegt. Kleine

DNA-Fragmente wurden durch beide Enzyme gebildet, was darauf hinweist, dass diese an CpG-reiche Region in der Tat in genomischer DNA untermethyliert ist. Die Gelblothybridisierung mit dem E4.3-Segment (**Fig. 6**) bringt bei beiden Enzymen sehr kleine Hybridisierungsfragmente zum Vorschein, was auf das Vorliegen einer hypomethylierten CpG-Insel hinweist.

[0114] Die obigen Ergebnisse weisen stark auf das Vorliegen einer kodierenden Region an diesem Locus hin. Zwei DNA-Segmente (E4.3 und H1.6), die Kreuzspezies-Hybridisierungssignale aus diesem Bereich erfassten, wurden als Sonden zum Screenen von cDNA-Bibliotheken verwendet, die aus verschiedenen Geweben und Zelltyp hergestellt worden waren.

[0115] cDNA-Bibliotheken aus gezüchteten Epithelzellen wurden wie folgt zubereitet. Schweißdrüsenzellen, die aus einem nicht an ZF leidenden Individuum und von einem an ZF leidendem Patienten deriviert worden waren, wurden wie oben beschrieben bis zum ersten Durchgang gezüchtet (G. Collins et al, *In Vitro Cell. Dev. Biol.* 21, 592, 1985). Das Vorliegen nach außen rektifizierender Kanäle wurde in diesen Zellen bestätigt (J. A. Tabcharani, T. J. Jansen, J. R. Riordan, J. W. Hanrahan, *J. Memb. Biol.*, im Druck), die ZF-Zellen waren jedoch der Aktivierung durch zyklisches AMP gegenüber unempfindlich (T. J. Jensen, J. W. Hanrahan, J. A. Tabcharani, M. Buchwald und J. R. Riordan, *Pediatric Pulmonology*, Nachtrag 2, 100, 1988). RNA wurde von diesen durch die Methode von J. M. Chirgwin et al (*Biochemistry* 18, 5294, 1979) isoliert. Poly A + RNA wurde ausgewählt (H. Aviv und F. Leder, *Proc. Natl. Acad. Sci. USA* 69, 1408, 1972) und als Matrize für die Synthese von cDNA mit Oligo (dT) 12–18 als Primer verwendet. Der zweite Strang wurde Gubler und Hoffman (*Gene* 25, 263, 1983) entsprechend synthetisiert. Er wurde mit EcoRI-Methylase methyliert und die Enden wurden auf dieselbe Ebene wie T4-DANN-Polymerase gebracht. Phosphorylierte EcoRI-Koppler wurden an die cDNA ligiert und mit EcoRI einer Restriktion unterzogen. Die Entfernung von überschüssigen Kopplern und die partielle Größenfraktionierung wurden durch Biogel A-50-Chromatographie erzielt. Die cDNAs wurden in die EcoRI-Stelle des im Handel erhältlichen lambda ZAP ligiert. Rekombinanten wurden verpackt und in *E. coli* BB4 propagiert. Teile der verpackten Mischungen wurden amplifiziert und der Rest für das Screenen vor der Amplifizierung zurückbehalten. Die gleichen Verfahren wurden zum Konstruieren einer Bibliothek aus RNA verwendet, das aus vorkonfluierenden Kulturen der Kolonkrebzelllinie T-84 isoliert worden war (Dharmasathaphorn, K. et al. *Am. J. Physiol.* 246, G204, 1984). Die Anzahl unabhängiger Rekombinanten in den drei Bibliotheken waren wie folgt: 2×10^6 bei den Ncht-ZF-Schweißdrüsenzellen, $4,5 \times 10^6$ bei den ZF-Schweißdrüsenzellen und $3,2 \times 10^6$ bei den T-84-Zellen. Diese Phagen wurden mit 50.000 pro 15 cm-Platte plattiert und unter Zuhilfenahme von Nylonmembranen (Biodyne) wurden Plaqueabhebungen hergestellt und mit durch ^{32}p markierten DNA-Fragmenten unter Zuhilfenahme von DNA-Polymerase I und einer Zufallsmischung von Oligonucleotiden als Primer sondiert. Die Hybridisierungsbedingungen entsprachen G. M. Wahl und S. L. Berger (*Math. Enzymol.* 152, 415, 1987). Bluescript[®]-Plasmide wurden aus Klonen, die von Plaque gereinigt worden waren, durch Exzision mit M13-Hilfophage gerettet. Die Lungen- und Pankreasbibliotheken wurde von Clontech Lab Inc. erworben, die den Angaben nach Mengen von $1,4 \times 10^6$ bzw. $1,7 \times 10^6$ unabhängigen Klonen aufwiesen.

[0116] Nach dem Screenen von 7 verschiedenen Bibliotheken, von denen jede 1×10^5 – 5×10^6 unabhängige Klone enthielt, wurde ein einziger Klon (als 10–1 identifiziert) mit H1.6 aus einer cDNA-Bibliothek isoliert, die aus den gezüchteten Epithelzellen von Schweißdrüsen bei einem unbetroffenen (nicht-ZF)-Individuum hergestellt worden waren.

[0117] Die DNA-Sequenzanalyse zeigte, dass 10-1 eine Insertion einer Größe von 920 bb und ein potentielles langes offenes Leseraster (OLR) enthielt. Da ein Ende der Sequenz eine perfekte Sequenzidentität mit H1.6 gemeinsam hatte, wurde der Schluss gezogen, dass der cDNA-Klon wahrscheinlich aus dieser Region abgeleitet worden war. Die gemeinsame DNA-Sequenz war jedoch nur 113 bp lang (vergleiche **Fig. 1** und 7). Wie weiter unten erklärt, entspricht diese Sequenz in der Tat dem 5'-nächsten Exon des putativen ZF-Gens. Die kurze Sequenzüberlappung ist daher eine Erklärung der schwachen Hybridisierungssignale beim Bibliothekscreenen und der Unfähigkeit, Transkripte in RNA-Genblotanalysen zu erfassen. Außerdem wurde die Orientierung der Transkriptionseinheit auf der Basis der Ausrichtung der genomischen DNA-Sequenz mit dem angenommenen OLR von 10-1 zögernd bestimmt.

[0118] Da das entsprechende Transkript durch RNA-Gelblothybridisierungsexperimente schätzungsweise auf eine Länge von ca. 6500 Nucleotide bestimmt worden war, war ein weiteres cDNA-Bibliothekscreenen erforderlich, um den Rest der kodierenden Region zu klonieren. Durch mehrere aufeinanderfolgende Screeningvorgänge mit cDNA-Bibliotheken, die aus der Kolon-Karzinomzelllinie T84, normalen und ZF-Schweißdrüsenzellen, der Pankreas und Lungen Erwachsener gebildet worden waren, wurden 18 zusätzliche Klone isoliert (**Fig. 7** wie weiter unten in weiteren Einzelheiten besprochen). Die DNA-Sequenzanalyse zeigte, dass keiner dieser cDNA-Klone der Länge des beobachteten Transkripts entspricht, dass es jedoch möglich war, eine Konsens-Sequenz auf der Basis überlappender Regionen abzuleiten. Zusätzliche cDNA-Klone, die den 5'- und 3'-Enden des Transkripts entsprachen, wurden aus 5'- und 3'-Primärausdehnungsexperimenten abgeleitet. Zusammen überspannen diese Klone insgesamt ca. 6,1 kb und enthalten ein OLR, das fähig ist, für ein Polypeptid von 1480 Aminosäureresten zu kodieren (**Fig. 1**).

[0119] Außergewöhnlicherweise wurde beobachtet, dass die meisten der hier isolierten cDNA-Klone Sequen-

zinsertionen an verschiedenen Positionen der Restriktionskarte von **Fig. 7** enthielten. Die Karte gibt Einzelheiten der genomischen Struktur des ZF-Gens an. Exon-/Introngrenzen sind dort angegeben, wo alle isolierten cDNA-Klone schematisch im oberen Teil der Figur dargestellt sind. Viele dieser extra Sequenzen entsprechen eindeutigweise Intronregionen, die während der Konstruktion der cDNA revers transkribiert werden, wie auf das Ausrichten mit genomischen DNA-Sequenzen hin gezeigt.

[0120] Da die Anzahl rekombinanter cDNA-Klone für das beim Bibliothekscreenen erfasste ZF-Gen viel geringer war, als aufgrund der großen Anzahl von Transkripten, die durch RNA-Hybridisierungsexperimente schätzungsweise bestimmt worden war, zu erwarten gewesen wäre, scheint es wahrscheinlich, dass die Klone, die aberrante Strukturen enthielten, bevorzugt zurückgehalten wurden, während die richtigen Klone während des Propagierens verloren gingen. Im Einklang mit dieser Interpretation wurde ein geringes Wachstum bei der Mehrzahl der rekombinanten Klone beobachtet, die bei dieser Studie isoliert wurden, gleichgültig, welcher Vektor auch verwendet wurde.

[0121] Die Verfahren, die angewendet wurden, um die 5'- und 3'-Ende der cDNA zu erhalten, waren denjenigen ähnlich, die schon beschrieben worden sind (M. Frohman et al. Proc. Nat. Acad. Sci, USA, 85, 8998–9002, 1988). Für die 5'-endigen Klone wurden Proben der Gesamtpankreas- und T84 Poly A+ RNA revers transkribiert unter Zuhilfenahme eines Primers (10b), der für Exon 2 spezifisch ist, und zwar auf ähnliche Weise wie es schon für die Primerextensionsreaktion beschrieben worden ist, mit Ausnahme der Tatsache, dass ein radioaktiver Markierungsstoff in die Reaktion eingeschlossen wurde. Die von einer Agaroseperlsäule der ersten Strangsynthese abgenommenen Fraktionen wurden durch Polymerasekettenreaktion (PKR) eluierter Fraktionen untersucht. Die verwendeten Oligonucleotide lagen innerhalb der 10-1-Sequenz (145 Nucleotide entfernt) und nur 5' vom Extensionsprimer. Die frühesten, das PKR-Produkt ergebenden Fraktionen wurden zu einem Pool vereinigt und durch Verdampfen konzentriert und daraufhin mit terminaler Desoxynucleotidyltransferase (BRL Labs.) und dATP, wie vom Lieferanten (BRL Labs.) empfohlen geschwänzt. Eine zweite Strangsynthese wurde daraufhin mit Taq-Polymerase (Cetus AmpliTaq®) unter Zuhilfenahme eines Oligonucleotids durchgeführt, das eine geschwänzte Kopplersequenz, nämlich 5'CGGAATTCTCGAGATC(T)₁₂3', enthielt.

[0122] Die Amplifizierung durch ein (PKR)-Fixierungsexperiment wurde daraufhin durchgeführt, unter Zuhilfenahme der Kopplersequenz und eines Primers, der gerade innerhalb des Extensionsprimers lag, der die EcoRI-Restriktionsstelle an seinem 5'-Ende aufwies. Auf die Restriktion mit den Enzymen EcoRI und Bgl II und die Agarosegelreinigung hin, wurden größtmäßig ausgewählte Produkte durch Standardverfahren (Maniatis et al, supra) in das Plasmid Bluescript KS kloniert, das von Stratagene erhältlich ist. Im Wesentlichen enthielten alle gewonnenen Klone Insertionen von weniger als 350 Nucleotiden. Um die 3'-endigen Klone zu erhalten, wurde Erststrang-cDNA unter Reverstranskription von 2 µg T84 Poly A + RNA zubereitet unter Zuhilfenahme des oben beschriebenen geschwänzten Koppleroligonucleotids unter Bedingungen, die denjenigen der Primerextension ähnlich waren. Die Amplifizierung durch PKR wurde daraufhin mit dem Koppleroligonucleotid und drei verschiedenen Oligonucleotiden durchgeführt, die bekannten Sequenzen des Klons T16-4.5 entsprachen. Eine Reaktion im Vorbereitungsmaßstab (2 × 100 µl) wurde mit einem dieser Oligonucleotide mit der Sequenz 5'ATGAAGTCCAAGGATTTAG3' durchgeführt.

[0123] Dieses Oligonucleotid befindet sich ca. 70 Nucleotide stromaufwärts von einer Hind III-Stelle innerhalb der bekannten Sequenz von T16-4.5. Auf die Restriktion des PKR-Produkts mit Hind III und Xho I folgte die Agarosegelreinigung zum größtmäßigen Auswählen einer Bande bei 1.0–1.4 kb. Diese Produkt wurde daraufhin in das von Stratagene erhältliche Plasmid Bluescript KS kloniert. Ca. 20% der so erhaltenen Klone wurde an den 3'-Endteil von T16-4.5 hybridisiert. 10/10 der aus diesen Klonen isolierten Plasmide wiesen identische Restriktionskarten mit Insertionsgrößen von ca. 1,2 kb auf. Alle diese PK Reaktionen wurden über 30 Zyklen in von einem Lieferanten von Enzymen vorgeschlagenem Puffer durchgeführt.

[0124] Ein 157 nt vom 5'-Ende des 10-1-Klons entfernter Extensionsprimer wurde zum Identifizieren des Ausgangspunkts des putativen ZF-Transkripts verwendet. Der Primer wurde mit γ (³²P)ATP bei 5000 Curie/mMol und T4-Polynucleotidkinase endmarkiert und durch Spinsäulengelfiltration gereinigt. Der radiomarkierte Primer wurde daraufhin mit 4–5 µg Poly A + RNA, die aus T-84-Kolonkrebszellen zubereitet worden war, in 2X Reverstranskriptasepuffer 2 Stunden bei 60°C reassoziert. Auf die Verdünnung und Zugabe von AMV-Reverstranskriptase (Life Sciences, Inc.) hin wurde eine einstündige Inkubation bei 41°C durchgeführt. Die Probe wurde daraufhin auf 0,4 M NaOH und 20 mM EDTA eingestellt und schließlich neutralisiert, mit NH₄OAC auf einen pH-Wert von 4,6 eingestellt, mit Phenol extrahiert, mit Ethanol ausgefällt, nochmals in Puffer mit Formamid gelöst und auf einem Polyacrylamid-Sequenziergel analysiert. Die Einzelheiten dieser Methoden sind schon beschrieben worden (Meth. Enzymol. 152, 1987, Verfasser. S. L. Berger, A. R. Kimmel, Academic Press, N. Y.).

[0125] Die Ergebnisse des Experiments zur Primerextension unter Zuhilfenahme eines Extensionsoligonucleotidprimers, der 157 Nucleotide vom 5'-Ende von 10-1 entfernt beginnt, ist im Abschnitt A der **Fig. 10** gezeigt. Der mit Hae III (BRL Labs.) digerierte endmarkierte X174 Bakteriophage wird als Größenmarker verwendet. Zwei Hauptprodukte werden bei 216 und 100 Nucleotiden beobachtet. Die 100 Nucleotiden in 10–1 entspre-

chende Sequenz entspricht einer an GZ äußerst reichen Sequenz (11/12), was darauf hinweist, dass es sich dabei um eine Reverstranskriptaseleerstelle handeln könnte. Die 5'-fixierten PKR-Ergebnisse sind in Abschnitt B der **Fig. 10** gezeigt. Das 1,4%ige Agarosegel, das auf der linken Seite gezeigt ist, wurde dann einem Blottingverfahren unterzogen und auf die Zetaprobe[®]-Membran (Bio-Rad Lab) übertragen. Die DNA-Gelblothybridisierung mit radiomarkiertem 10-1 ist auf der rechten Seite gezeigt. Es ist ersichtlich, dass die 5'-Extensionsprodukte größtenteils zwischen 170–280 nt liegen, wobei das Hauptprodukt bei ca. 200 Nucleotide liegt. Die PKR-Steuerbahn zeigt ein Fragment von 145 Nucleotiden. es wurde unter Verwendung der Testoligomere innerhalb der 10-1-Sequenz erhalten. Die gezeigten Größenmarker entsprechen Größen von 154, 220/210, 298, 344, 394 Nucleotiden (von BRL Lab erworbene 1 kb-Leiter).

[0126] Die unten in Abschnitt B von **Fig. 10** gezeigte schematische Darstellung skizziert das Verfahren, das zum Erzielen der Doppelstrang cDNA verwendet wird, die zum Amplifizieren und Klonieren unter Bildung der in **Fig. 7** gezeigten Klone PA3-5 und TB2-7 verwendet wird. Die zum Charakterisieren des 3'-Endes angewendeten PKR-Fixierexperimente sind in Abschnitt C gezeigt. Wie in der schematischen Darstellung unter **Fig. 10C** gezeigt, wurden drei Primer, deren relative Position zueinander bekannt war, zum Amplifizieren mit revers transkribierter T84 RNA, wie beschrieben, verwendet. Diese Produkte wurden auf einem 1%igen Agarosegel getrennt und es wurde, wie oben beschrieben, damit ein Blot auf Nylonmembran durchgeführt. Die DNA Blothybridisierung mit dem 3'-Teil des Klons T16-4.5 ergab Banden, die Größen aufwiesen, die dem Abstand zwischen dem verwendeten spezifischen Oligomer und dem 3'-Ende des Transkripts entsprachen. Diese Banden in den Bahnen 1, 2A und 3 sind in **Fig. 10** unterhalb des Abschnitts C schematisch dargestellt. Die Bande in Bahn 3 ist schwach, da nur 60 Nucleotide dieses Segments mit der verwendeten Sonde überlappen. In der schematischen Darstellung und wie in der Bahn 2b gezeigt, ist auch das Produkt angegeben, das durch Restriktion des PKR-Fixierprodukts zum Erleichtern des Klonierens zur Bildung des in **Fig. 7** gezeigten Klons THZ-4 gebildet wird.

[0127] Die DNA-Blothybridisierungsanalyse von genomischer DNA, die mit EcoRI- und Hind III-Enzymen digeriert ist, die mit Anteilen von cDNAs sondiert sind, die das gesamte Transkript überspannen, weist daraufhin, dass das Gen mindestens 24 Exone enthält, die mit den römischen Zahlen I bis XXIV bezeichnet sind (vergleiche **Fig. 9**). Diese entsprechen den in **Fig. 7** gezeigten Zahlen 1 bis 24. Die Größe jeder Bande ist in kb angegeben.

[0128] In **Fig. 7** zeigen die offenen Kästchen die ungefähren Positionen der 24 Exone an, die durch Isolieren von > 22 Klonen aus dem Screenen von cDNA-Bibliotheken und PKR-Fixierexperimenten, die zum Klonieren der 5'- und 3'-Enden konzipiert worden sind, identifiziert worden sind. Die Längen, in kb, der von jedem Exon erfassten genomischen EcoRI-Fragmente sind ebenfalls angezeigt. Die schattierten Kästchen in **Fig. 7** zeigen das Vorliegen von Intrasequenzen und die punktierten Kästchen zeigen andere Sequenzen an. In der unteren linken Ecke in der Nähe des geschlossenen Kästchens ist die relative Position des Klons H1.6 angezeigt, der zum Erfassen des ersten cDNA-Klons 10-1 unter 10⁶ Phagen der normalen Schweißdrüsenbibliothek verwendet worden ist. Wie in den **Fig. 4(D)** und **7** gezeigt, überlappt der genomische Klon H1.6 teilweise mit einem EcoRI-Fragment von 4,3 kb. Alle der gezeigten cDNA-Klone wurden zu genomischer DNA hybridisiert und/oder einer feinen Restriktionskartierung unterzogen. Beispiele der innerhalb der cDNAs und in den entsprechenden genomischen Fragmenten vorkommenden Restriktionsstellen sind angegeben.

[0129] Mit Bezug auf **Fig. 9** umfasst die Hybridisierungsanalyse Sonden, d. h. die cDNA-Klone 10-1 für Abschnitt A, T16-1 (3'-Anteil) für Abschnitt B, T16-4.5 (mittlerer Anteil) für Abschnitt C und T16-4.5 (3'-Endanteil) für Abschnitt D. In Abschnitt A der **Fig. 9** erfasst die cDNA-Sonde 10-1 die genomischen Banden für die Exone I bis VI. Der durch Nru I-Restriktion gebildete 3'-Anteil von T16-1 erfasst die Exone IV bis XIII, wie in Abschnitt B gezeigt. Diese Sonde überlappt Teil 10-1 teilweise. Die Abschnitte C und D zeigen jeweils genomische Banden, die durch die mittleren und 3'-endigen EcoRI-Fragmente des Klons T16-4.5 erfasst werden. Zwei EcoRI-Stellen kommen innerhalb der cDNA-Sequenz vor und spalten die Exone XIII und XIX. Wie durch die Exone in Klammern angezeigt, entsprechen zwei genomische EcoRI-Banden jedem dieser Exone. Es wurde die Kreuzhybridisierung zu anderen genomischen Fragmenten beobachtet. Der Ursprung dieser durch N angezeigten Banden ist nicht das Chromosom 7, denn sie sind in Human-Hamsterhybriden, die das menschliche Chromosom 7 enthalten, nicht vorgekommen. Es wird angenommen, dass die durch XI in Klammern angezeigte schwache Bande in Abschnitt D durch die Kreuzhybridisierung von Sequenzen aufgrund der internen Homologie mit der cDNA hervorgerufen wird.

[0130] Da 10-1 bei der Gelblothybridisierung von RNA aus der Kolonkrebiszelllinie T-84 eine starke Bande erfasste, wurde diese cDNA zum Screenen der aus dieser Quelle konstruierten Bibliothek verwendet. Es wurden fünfzehn Positive erhalten, aus denen die Klone T6, T6/20, T11, T16-1 und T13-1 gereinigt und sequenziert wurden. Das erneute Screenen der gleichen Bibliothek mit einem Bam HI-EcoRI-Fragment von 0,75 kb aus dem 3'-Ende von T16-1 ergab T16-4.5. Ein EcoRI-Fragment von 1,8 kb aus dem 3'-Ende von T16-4.5 ergab T8-B3 und T12a, wobei letzteres ein Polyadenylierungssignal und einen Polyadenylierungsschwanz enthielt. Gleichzeitig wurde eine menschliche Lungen-cDNA-Bibliothek einem Screenen unterworfen. Es wurden viele Klone isoliert, einschließlich derjenigen, die hier mit der Vorsilbe "CDL" gezeigt sind. Eine Pankreasbibliothek

wurde ebenfalls einem Screenen unterzogen, was zum Klon CDPJ5 führte.

[0131] Um Kopien dieses Transkripts aus einem an ZF leidenden Patienten zu erhalten, wurde eine cDNA-Bibliothek von der RNA von Schweißdrüsenepithelzellen eines Patienten einem Screening mit dem Bam HI-EcoRI-Fragment von 0,75 kb aus dem 3'-Ende von T16-1 und den Klonen C16-1 und C1-1/5 unterzogen, die alle, mit Ausnahme von Exon I überspannten. Diese beiden Klone wiesen beide bei Exon 10 eine 3 bp-Deletion auf, die bei keinem anderen Exon enthaltenden Klon vorhanden ist. Mehrere Klone, einschließlich CDLS26-1 aus der Lungenbibliothek und aus T84 isolierte T6/20 und T13-1 wurden von teilweise verarbeiteten Transkripten abgeleitet. Dies wurde durch genomische Hybridisierung und durch Sequenzieren über die Exon-Intron-Grenzen hinweg bei jedem Klon bestätigt. T11 enthielt auch eine zusätzliche Sequenz an jedem Ende. T16-4.5 enthielt eine kleine Insertion in der Nähe der Grenze zwischen den Exonen 10 und 11, die keiner Intronsequenz entsprach. Die Klone CDLS16A, 11a und 13a aus der Lungenbibliothek enthielten auch fremde Sequenzen unbekanntes Ursprungs. Der Klon C16-1 enthielt auch eine kurze Insertion, die einem Teil des γ -Transposons von *E. coli* entsprach. Dieses Element wurde bei anderen Klonen nicht entdeckt. Die 5'-Klone PA3-5, die aus Pankreas-RNA und TB2-7 gebildet sind; das unter Zuhilfenahme der PCR-Fixierungstechnik aus T84-RNA gebildet ist, weisen identische Sequenzen auf, mit Ausnahme eines einzigen Nucleotidunterschieds bezüglich der Länge beim 5'-Ende, wie in **Fig. 1** gezeigt. Der 3'-Klon THZ-4, der aus T84-RNA erhalten wird, enthält die 3'-Sequenz des Transkripts der genomischen Sequenz dieser Region entsprechend.

[0132] Eine kombinierte Sequenz, die die mutmaßliche Kodierregion des ZF-Gens darstellt, wurde aus sich überlappenden cDNA-Klonen gebildet. Da die meisten der cDNA-Klone anscheinenderweise aus unverarbeiteten Transkripten abgeleitet waren, wurden weiter Studien durchgeführt, um die Authentizität der kombinierte Sequenz sicherzustellen. Jeder cDNA-Klon wurde zuerst bezüglich der Lokalisierung bezüglich des Chromosoms 7 durch Hybridisierungsanalyse mit einer ein einziges menschliches Chromosom 7 enthaltenden Hybride einer somatischen Human-Hamsterzelle und durch Pulsfeldgelelektrophorese getestet. Eine Restriktionsenzym-Feinkartierung wurde ebenfalls für jeden Klon durchgeführt. Während sich überlappende Regionen bei den meisten der Klone eindeutig identifizierbar waren, enthielten viele Regionen mit einzigartigen Restriktionsmustern.

[0133] Zum weiteren Charakterisieren dieser cDNA-Klone wurden sie als Sonden bei Gelhybridisierungsexperimenten mit durch EcoRI oder Hind III digerierter menschlicher genomischer DNA verwendet. Wie in **Fig. 9** gezeigt, konnten fünf bis sechs verschiedene Restriktionsfragmente mit den 10-1 cDNA- und eine ähnliche Anzahl von Fragmenten mit anderen cDNA-Klonen erfasst werden, was auf das Vorliegen multipler Exone für das putative ZF-Gen hinweist. Bei den Hybridisierungsstudien wurden auch diejenigen cDNA-Klone identifiziert, die unverarbeitete Intron-Sequenzen aufwiesen, denn sie zeigten eine bevorzugte Hybridisierung an eine Subgruppe genomischer DNA-Fragmente. Für die bestätigten cDNA-Klone wurden ihre entsprechenden genomischen DNA-Segmente isoliert und die Exone und Exon/Intron-Grenzen sequenziert. Wie in **Fig. 7** angegeben, wurden insgesamt 24 Exone identifiziert. Auf der Basis dieser Informationen und der Ergebnisse physikalischer Kartierungsexperimente, wurde der Genlocus schätzungsweise dahingehend bestimmt, dass er sich über 250 kb am Chromosom 7 erstreckt.

2.6 DIE SEQUENZ

[0134] **Fig. 1** zeigt die Nucleotidsequenz der klonierten, für CFTR kodierenden cDNA in Verbindung mit der deduzierten Aminosäuresequenz. Die erste Basenposition entspricht dem ersten Nucleotid im 5'-Extensionklon PA3-5, der um ein Nucleotid länger ist als TB2-7. Die Pfeile zeigen die Position der Transkriptions-Initierungsstelle bei der Primerextensionsanalyse an. Auf das Nucleotid 6129 folgt ein Poly(dA)-Trakt. Die Positionen von Exonverbindungsstellen sind durch senkrechte Linien angezeigt. Potentielle membranüberspannende Segmente wurde mit Hilfe der Algorithmen von Eisenberg et al. *J. Mol. Biol.* 179: 125 (1984) festgestellt. Potentiell membranüberspannende Segmente, wie sie in **Fig. 11** analysiert und gezeigt sind, sind in **Fig. 1** mit Kästchen umgeben. In **Fig. 11** ist der durchschnittliche Hydropathieindex (Kyte und Doolittle. *J. Molec. Biol.* 157: 105, (1982)) von 9 Restpeptiden als Funktion der Aminosäurezahl aufgezeichnet. Die entsprechenden Positionen von Merkmalen der nach Garnier et al. *J. Molec. Biol.* 157, 165 (1982)) vorausgesagten Sekundärstruktur sind im unteren Abschnitt aufgezeigt. Aminosäuren, die putative ATP-Bindungsfalten umfassen, sind in **Fig. 1** unterstrichen. Mögliche Stellen für die Phosphorylierung durch Proteinkinasen A (PKA) oder C (PKC) sind durch offene bzw. geschlossenen Kreise angezeigt. Das offene Dreieck befindet sich über den 3 bp (CTT), die bei ZF gelöscht sind (man vergleiche die Diskussion weiter unten). Die cDNA-Klone in **Fig. 1** wurden durch die Didesoxy-Kettenterminationsmethode unter Zuhilfenahme von mit ³⁵S markierten Nucleotiden durch den automatischen DNA-Sequenzierer Genesis 2000[®] von Dupont sequenziert.

[0135] Die kombinierte DNA-Sequenz überspannt 6129 Basenpaare, ausschließlich des Poly(A)-Schwanzes am Ende der untranslatierten 3'-Region und enthält ein OLR, das in der Lage ist, für ein Polypeptid von 1480 Aminosäuren zu kodieren (**Fig. 1**). Zu Beginn dieses OLR (Basenposition 133–135), liegt ein ATG- (AUG) Triplet vor. Da die dieses Kodon (5'-AGACCAUGCA-3') umgebende Nucleotidsequenz die vorgeschlagenen

Merkmale der Konsenssequenz (KS) A/GCCAUGG(G) einer eukaryotischen Translationsinitiationsstelle mit einem stark konservierten A in der 3'-Position aufweist, ist es äußerst wahrscheinlich, dass dieses AUG dem ersten Methioninkodon für das putative Polypeptid entspricht.

[0136] Um die Sequenz zu erhalten, die dem 5'-Ende des Transkripts entspricht, wurde ein Primerextensionsexperiment, wie oben beschrieben, durchgeführt. Wie in **Fig. 10A** gezeigt, konnte ein Primerextensionsprodukt von ca. 216 Nucleotiden beobachtet werden, was darauf hinweist, dass das 5'-Ende des Transkripts ca. 60 Nucleotide stromaufwärts vom Ende des cDNA-Klons 10-1 initiierte. Eine modifizierte Polymerasekettenreaktion (Fixier-PKR) wurde daraufhin zum Erleichtern des Klonierens der 5'-endigen Sequenz (**Fig. 10b**) verwendet. Zwei unabhängige 5'-Extensionsklone, einer aus der Pankreas und der andere aus T84 RNA, wurden durch DNA-Sequenzieren gekennzeichnet und es wurde festgestellt, dass sie sich bezüglich der Länge um nur eine Base unterscheiden, was darauf hinweist, dass es sich um die wahrscheinlichste Initiationsstelle für das Transkript, wie in **Fig. 1** gezeigt, handelt.

[0137] Da die meisten der ursprünglichen cDNA-Klone keinen auf das Ende einer mRNA hinweisende PolyA-Schwanz enthielten, wurde eine Fixier-PKR auch an das 3'-Ende des Transkripts angewendet (Frohman et al, 1988, supra). Es wurden drei 3'-Extensionsoligonucleotide an das Endteil des cDNA-Klons T16-4.5 angebracht. Wie in **Fig. 10C** gezeigt, wurden 3 PKR-Produkte verschiedener Größe erhalten. Alle stimmten mit der Interpretation überein, dass das Ende des Transkripts um ca. 1,2 kb stromabwärts von der Hind III-Stelle an der Nucleotidposition 5027 gelegen war (vergleiche **Fig. 1**). Die von repräsentativen Klonen abgeleitete DNA-Sequenz stimmte mit derjenigen des T84-cDNA-Klons T12a (vergleiche **Fig. 1** und 7) und der Sequenz des entsprechenden genomischen EcoRI-Fragments von 2,3 kb überein.

3.0 DIE MOLEKULARGENETIK VON ZF

3.1 EXPRESSIONSSTELLEN

[0138] Um das Transkript für das putative ZF-Gen darzustellen, wurden Experimente zur RNA-Gelblothybridisierung mit der 10-1 cDNA als Sonde durchgeführt. Die Ergebnisse der RNA-Hybridisierung sind in **Fig. 8** gezeigt.

[0139] Es wurden aus Gewebeproben, die aus der chirurgischen Pathologie oder einer Sektion schon beschriebenen Methoden entsprechend erhalten worden waren (A. M. Kimmel, S. L. Berger, Verfasser, Meth. Enzymol. 152, 1987) RNA-Proben zubereitet. Formaldehydgelte wurden auf Nylonmembranen übertragen (Zeta-probe®, BioRad Lab). Die Membranen wurden daraufhin mit DNA-Sonden, die auf eine hochspezifische Aktivität durch die Randomprimermethode markiert worden waren (A. P. Feinberg und B. Vogelstein, Anal. Biochem. 132, 6, 1983) schon veröffentlichten Verfahren (J. Rommens et al, Am. J. Hum. Genet. 43, 645–663, 1989) gemäß hybridisiert. **Fig. 8** zeigt die Hybridisierung durch den cDNA-Klon 10-1 an ein 6,5 kb-Transkript in den angezeigten Geweben. Die gesamte RNA (10 µg) eines jeden Gewebes und die Poly A + RNA (1 µg) der Kolonkrebszellenlinie T84 wurden auf einem 1%igen Formaldehydgel getrennt. Die Positionen der rRNA-Banden 28S und 18S sind angezeigt. Pfeile zeigen die Position von Transkripten an. Die Größe wurde durch Vergleich mit Standard-RNA-Markern (BRL Labs) bestimmt. HL60 ist eine menschliche Promyelozytenleukämie-Zelllinie und T84 ist eine menschliche Kolonkrebszellenlinie.

[0140] Die Analyse bringt eine auffallende Bande einer Größe von ca. 6,5 kb in T84-Zellen zum Vorschein. Ähnlicherweise wurden auch starke Hybridisierungssignale in Pankreas- und primären Kulturen von Zellen eines Nasenpolyps erfasst, was darauf hinweist, dass die reife mRNA des putativen ZF-Gens ca. 6,5 kb bemisst. Geringe Hybridisierungssignale, die wahrscheinlich Abbauprodukte darstellen, wurden in den niedrigeren Größenbereichen erfasst, sie waren jedoch von Versuch zu Versuch verschieden. Identische Ergebnisse wurden mit anderen, als Sonden verwendeten cDNA-Klonen erhalten. Auf der Basis der Intensität der Hybridisierungsbande und des Vergleichs mit denjenigen, die für andere Transkripte unter identischen Versuchsbedingungen erfasst worden sind, wurde schätzungsweise angenommen, dass die putativen ZF-Transkripte ca. 0,01% der gesamten mRNA in T84-Zellen darstellten.

[0141] Eine Anzahl anderer Gewebe wurde ebenfalls durch RNA-Gelblothybridisierungsanalyse beim Versuch untersucht, das Expressionsbild des 10-1-Gens und die Pathologie von ZF miteinander in Bezug zu setzen. Wie in **Fig. 8** gezeigt, wurden Transkripte identischer Größe in der Lunge, dem Dickdarm, in Schweißdrüsen (kultivierte Epithelzellen), der Plazenta, Leber und der Ohrspeicheldrüse aufgefunden, die Signalintensitäten in diesen Geweben waren jedoch bei verschiedenen Zubereitungen unterschiedlich und im Allgemeinen schwächer als diejenigen, die im Pankreas und in Nasenpolypen bestimmt worden ist. Die Intensität war bei verschiedenen Zubereitungen unterschiedlich; beispielsweise wurde bei der in **Fig. 8** gezeigten Zubereitung keine Hybridisierung in der Niere erfasst, kann jedoch in darauffolgenden wiederholten Untersuchungen erfasst werden. Im Gehirn oder in der Nebenniere (**Fig. 8**) konnten keine Hybridisierungssignale entdeckt werden, und in Hautfibroblast- und Lymphoblastzelllinien ebenfalls nicht.

[0142] Zusammenfassend scheint die Expression des ZF-Gens in vielen der untersuchten Gewebe vorzu-

kommen und in höherem Masse bei denjenigen Geweben, die stark von der ZF betroffen sind. Während diese für Epithelgewebe spezifische Expressionsbilder gut mit der Pathologie der Krankheit übereinstimmt, ist kein signifikanter Unterschied zwischen der Menge oder Größe von Transkripten von ZF und Kontrollgeweben festgestellt worden, was mit der Annahme übereinstimmt, dass ZF-Mutationen leichte Änderungen am Nucleotid-niveau darstellen.

3.2 DIE WICHTIGSTE ZF-MUTATION

[0143] **Fig. 17** zeigt die DNA-Sequenz an der F508-Deletion. Auf der linken Seite ist das Reverskomplement der Sequenz von der Basenposition 1649–1664 der normalen Sequenz (wie sie von dem cDNA-Klon T16 abgeleitet ist) gezeigt. Die Nucleotidsequenz ist als Output (in willkürlichen Fluoreszenzintensitätseinheiten, y-Achse) als Funktion der Zeit (x-Achse) für jede der 2 Photovervielfacherröhren (PMT#1 und #2) eines Genesis 2000TM DNA-Analysesystems von Dupont aufgezeigt. Die entsprechende Nucleotidsequenz ist darunter gezeigt. Auf der rechten Seite ist die gleiche Region aus einer mutanten Sequenz (wie sie aus dem cDNA-Klon C 16 abgeleitet ist) gezeigt. Doppelstrangige Plasmid-DNA-Matrizen wurden durch das alkalische Lyseverfahren zubereitet. Fünf µg Plasmid-DNA und 75 ng Oligonucleotidprimer wurden bei jeder Sequenzierreaktion dem von Dupont empfohlenen Protokoll entsprechend verwendet, mit Ausnahme der Tatsache, dass die Wärmebehandlung 30 Minuten bei 45°C und der Verlängerungs-/Terminationsschritt für 10 Minuten bei 42°C stattfanden. Die nicht eingebauten Fluoreszenznucleotide wurden durch Ausfällen des DNA-Sequenzierreaktionsprodukts mit Ethanol in Gegenwart von 2,5 M Ammoniumacetat bei einem pH-Wen von 7,0 entfernt und einmal mit 70% Ethanol gespült. Der für das T16-1-Sequenzieren verwendete Primer bestand aus einem spezifischen Oligonucleotid 5'GTTGGCATGCTTTGATGACGCTTC3', das die Basenposition 1708–1731 umfasste und der Primer für C16-1 bestand aus dem Universalprimer SX für den Bluescript-Vektor (Stratagene). **Fig. 18** zeigt auch die DNA-Sequenz um die F508-Deletion, wie sie durch manuelles Sequenzieren bestimmt worden ist. Die normale Sequenz von der Basenposition 1726–1651 (von cDNA C16–1) ist neben der ZF-Sequenz (aus cDNA C16-1) gezeigt. Der linke Abschnitt zeigt die Sequenzen aus den Kodierungssträngen, die mit dem B-Primer (5'GTTTTCTGGAT-TATGCCTGGGCAC3') erhalten worden ist und der rechte Abschnitt zeigt diejenigen aus dem gegenüberliegenden Strang mit dem D-Primer (5'GTTGGCATGCTTTGATGACGCTTC3'). Die Klammern zeigen die drei Nucleotide in der normalen, die bei ZF (Pfeilköpfe) fehlen. Das Sequenzieren wurde wie bei F. Sanger, S. Nicklen, A. R. Coulson, Proc. Nat. Acad. Sci. USA 74: 5463 (1977) beschrieben, ausgeführt.

[0144] Um den Anteil der ZF-Patienten, die Träger dieser Deletion (F508) sind, wurden genomische DNA-Proben aus Patienten und ihren Eltern jeweils mit die Mutation flankierenden Oligonucleotidprimern in einer Polymerasekettenreaktion amplifiziert und an mit ³²P markierten Oligonucleotide hybridisiert, die für die normale und die putativ mutante Sequenz (**Fig. 2**) spezifisch sind. Die Ergebnisse dieser Analyse sind in Tabelle 2 gezeigt.

TABELLE 2

VERTEILUNG VON ZF-UND NICHT-ZF-(N) CHROMOSOMEN MIT UND OHNE DIE 3 bp-DELETION

a.	<u>ZF-Chromosomen</u>	<u>N-Chromosomen</u>
ohne Deletion	69	198
mit Deletion	<u>145</u>	<u>0</u>
Insgesamt	214	198

ZF-Chromosomen

	<u>mit der 3 bp-Deletion</u>	<u>ohne Deletion</u>
ZF-PI	62	24
ZF-PS	5	9
nicht klassifiziert	<u>78</u>	<u>36</u>
Insgesamt	145 (68 %)	69 (32 %)

[0145] Die Daten für die ZF-PI- (Pankreasinsuffizienz) und die ZF-PS-(Parikreassuffizienz) Chromosomen wurden aus den bei unserer Verkoppelungsanalyse verwendeten ZF-Familien abgeleitet. Diese Familien wurden ursprünglich ohne Kenntnisse bezüglich PI oder PS ausgewählt; die daraufhin identifizierten 15 ZF-PS-Familien wurden nicht in diese Berechnung eingeschlossen. Die nicht klassifizierten ZF-Chromosomen wurden vom DNA-Diagnoselabor beim Hospital for Sick Children in Toronto erhalten, für die keine Pankreasfunktionsdaten zur Verfügung standen.

[0146] Es ist ersichtlich, dass 68% (145/214) der ZF-Chromosomen bei der Patientenpopulation im Allgemeinen die F508-Deletion aufwies (Tabelle 2). Im Gegensatz dazu, wiesen keine (0/198) der N-Chromosomen diese Deletion auf (Tabelle 2; χ^2 -207, $pk10^{-57,5}$), was darauf hinweist, dass diese Sequenzänderung für ZF spezifisch ist und dass es die die Krankheit verursachende Hauptmutation darstellt. Zwischen der F508-Deletion und ZF ist keine Rekombination festgestellt worden.

[0147] Andere Sequenzunterschiede wurden zwischen den normalen (T16-4.5) und ZF- (C1-1/5) cDNA-Klonen beobachtet. Bei der Basenposition 2629 wies T16-4.5 ein C und C1-1/5 wies ein T auf, was beim Aminosäureniveau zu einer Änderung von Leu zu Phe führte. Bei Position 4555 war die Base G in T16-4.5 aber A in C1-1/5 (Val zu Met). Es wird angenommen, dass diese Befunde den Sequenzpolymorphismus darstellen. Die spezifische Oligonucleotidhybridisierungsanalyse von DNA von Patienten/Familie identifiziert diese als andere mögliche Mutationen. Zusätzliche Nucleotidunterschiede wurden in den nichttranslatierten 3'-Regionen zwischen verschiedenen cDNA-Klonen und der genomischen DNA-Sequenz beobachtet. Derartige Unterschiede in den Sequenzen und, wie man sich im Klaren sein wird, andere Sequenzmodifikationen sind möglich; derartige Unterschiede sind beispielsweise auf normale Sequenzpolymorphismen und Klonierartefakte zurückzuführen, wobei alle derartigen Unterschiede im Wesentlichen der in **Fig. 1** beschriebenen Sequenz bezüglich ihrer Funktion und ihrer handelsmäßigen Anwendung äquivalent sind.

[0148] Aufgrund der umfangreichen genetischen und physikalischen Kartierungsdaten sind Molekularklonierungsstudien auf ein kleines Segment von DNA am Chromosom 7 ausgerichtet worden. Wegen des Mangels an Chromosomendeletionen und -umgruppierungen im ZF und des Mangels an einer gut entwickelten funktionellen Untersuchung bezüglich des ZF-Genprodukts hat die Identifizierung des ZF-Gens eine genaue Charakterisierung des Locus selbst und den Vergleich zwischen den ZF- und normalen (N) Allelen erforderlich gemacht. Phenotypisch normale, zufallsmäßig ausgewählte Individuen konnten beim Vergleich nicht als Kontrollen verwendet werden, aufgrund der hohen Frequenz symptomloser Träger in der Population. Deshalb waren nur Eltern von ZF-Patienten, die beide definitionsgemäß Träger eines N- und eines ZF-Chromosoms sind, für die Analyse geeignet. Auch war es aufgrund der starken Allel-Assoziation, die zwischen ZF und einigen der eng verkoppelten DNA-Markern beobachtet worden war, notwendig, die Möglichkeit auszuschließen, dass zwischen N und ZF beobachtete Sequenzunterschiede Polymorphismen darstellten, die mit dem Krankheitslocus assoziiert sind.

3.3 IDENTIFIZIERUNG VON RFLPs UND FAMILIENSTUDIEN

[0149] Um den Zusammenhang zwischen den DNA-Segmenten, die bei Chromosomenwanderungs- und -sprungexperimenten isoliert worden sind, und der ZF zu bestimmten, wurden Restriktionsfragmentlängen-Polymorphismen (RFLP) identifiziert und zum Untersuchen von Familien herangezogen, bei denen Crossover-Vorkommnisse vorher zwischen CR und anderen flankierenden DNA-Markern entdeckt worden sind. Wie in **Fig. 14** gezeigt, wurden insgesamt 18 RFLP in der 500 kb-Region entdeckt; 17 von diesen (von E6 bis CE1.0) sind in Tabelle 3 aufgelistet; einige von ihnen entsprechen schon vorher berichteten Markern.

[0150] Fünf der RFLPs, nämlich 10-1X.6, T6/20, H1.3 und CE1.0, wurden mit cDNA- und genomischen DNA-Sonden identifiziert, die von dem putativen ZF-Gen abgeleitet worden waren. Die RFLP-Daten sind in Tabelle 3 angegeben, wobei Marker in den MET- und D7S8-Regionen zum Vergleich ebenfalls angegeben sind. Die physikalischen Abstände zwischen diesen Markern sowie ihre Beziehung zu den MET- und D7S8-Regionen sind in **Fig. 14** aufgezeigt.

TABELLE 3. MIT DEM ZF-GEN ASSOZIIERTE RFLPS

<u>Sonden- Bezeichnung</u>	<u>Enzym</u>	<u>Fragment- länge</u>	<u>N^(a)</u>	<u>ZF-P₁^(a)</u>	<u>A^(b)</u>	<u>*^(c)</u>	<u>Literaturangabe</u>
metD	BanI	7,6(kb)	28	48	0,60	0,10	J.E. Spence et al <u>Am.J.Hum.Genet.</u> 39:729 (1986)
metD	TaqI	6,8 6,2	59 74	25 75	0,66	0,06	R. White et al, <u>Nature</u> 318:382 (1985)
meth	TaqI	4,8 7,5	19 45	4 49	0,35	0,05	White et al. <u>supra</u>
E6	TaqI	4,0 4,4	38 58	20 62	0,45	0,06	B. Keren et al, <u>Am.J.Hum.Genet</u> 44:827 (1989)
E7	TaqI	3,6 3,9 3+0,9	42 40 51	17 16 57	0,47	0,07	
pH131	HinfI	0,4	81	33	0,73	0,15	J.M. Rommens et al, <u>Am.J.Hum.Genet</u> 43:645 (1988)

TABELLE 3 (FORTSETZUNG)

W3D1,4	HindIII	0,3 20	18 82	47 33	0,68	0,13	B. Kerem et al, <u>supra</u>
H2,3A	TaqI	10 2,1	22 39	47 51	0,64	0,09	K. Estivill et al, <u>Nature</u> 326:840 (1987); X. Estivill et al <u>Genomics</u> 1:257 (1987)
(XV2C)		1,4	37	11			
EG1,4	HincII	3,8 2,8	31 56	69 7	0,89	0,17	
EG1,4	BglII	20 15	27 62	69 9	0,89	0,18	
JG2E1	PstI	7,8	69	10	0,88	0,18	X.Estivill et al <u>supra</u> And B. Karem et al <u>supra</u>
(KM19)		6,6	30	70			
E2,6/E.9	HspI	11 8,5	34 26	6 55	0,85	0,14	
H2,8A	NcoI	25 8	22 52	55 9	0,87	0,18	

TABELLE 3 (FORTSETZUNG)

J3,11	MspI	6	47	38	0,29	0,04	B.M. Wainright et al, <u>Nature</u> 318:384 (1985)
J29	PvuII	1,8 9	62 26	38 36	0,36	0,06	M.C. Iannuzzi et al, <u>Supra</u>
		6	55	36			

ANMERKUNGEN ZU TABELLE 3

(a) Die Anzahl von N- und ZF-PI- (ZF mit Pankreasinsuffizienz) Chromosomen wurden aus den Eltern in den denjenigen Familien abgeleitet, die bei Verkoppelungsanalysen eingesetzt worden sind (Tsui et al, Cold Springs Harbor Symp. Quant. Biol. 51: 325 (1986).

(b) Die standardisierte Assoziation (A), die durch Schwankungen in der Verteilung der DNA-Markierungsallele unter den N-Chromosomen weniger stark beeinflusst ist, wird hier für den Vergleichskoeffizienten der Yule-Assoziation $A = (ad - bc)/(ad + bc)$ verwendet, wobei a, b, c und d die Anzahl von N-Chromosomen beim DNA-Markierungsallel 1, ZF mit 1, N mit 2 bzw. ZF mit 2 sind. Das relative Risiko kann unter Zuhilfenahme des Verhältnisses von $RR = (1 + A)/(1 - A)$ oder dessen umgekehrte Beziehung berechnet werden.

(c) Die Allel-Assoziation (*), die A. Chakravarti et al, Am. J. Hum. Genet. 36: 1239, (1984) entsnrechend unter Annahme einer Frequenz des ZF-Chromosoms von 0,02 in der Population berechnet worden ist, ist zum Vergleich ebenfalls eingeschlossen.

[0151] Aufgrund der geringen Anzahl rekombinanter Familien, die für die Analyse zur Verfügung standen, wie aufgrund der geringen Entfernung zwischen den untersuchten Markern und ZF zu erwarten war, und der Möglichkeit einer Falschdiagnose waren alternative Ansätze zum weiteren Feinkartieren des ZF-Gens erforderlich.

3.4 ALLEL-ASSOZIATION

[0152] Eine Allel-Assoziation (Verkoppelungsungleichgewicht) ist bei vielen eng verkoppelten DNA-Markern entdeckt worden. Während die Nützlichkeit des Verwendens einer Allel-Assoziation für das Messen der genetischen Entfernung ungewiss ist, ist eine Korrelation insgesamt zwischen ZF und den flankierenden DNA-Markern beobachtet worden. Eine starke Assoziation mit ZF wurde für die nähergelegenen DNA-Macker D7S23 und D7S122 beobachtet, während nur eine geringe oder keine Assoziation bei den entfernter liegenden Markern MET, D7S8 oder D7S424 beobachtet worden ist (vergleiche **Fig. 1**).

[0153] Wie in Tabelle 3 gezeigt, stieg der Assoziationsgrad zwischen DNA-Mackern und ZF (wie durch den Yule-Assoziationskoeffizienten gemessen) von 0,35 bei methH und 0,17 bei J32 auf 0,91 bei 10-1X.6 an (es wurden nur ZF-PI-Patientenfamilien bei der Analyse herangezogen, da sie genetisch homogener als ZF-PS zu sein schienen). Die Assoziationskoeffizienten schienen über die 300 kb von EG1.4 bis N1.3 ziemlich konstant zu sein; die bei verschiedenen Position, insbesondere bei H2.3A, E4.1 und T6/20 entdeckten Schwankungen waren wahrscheinlich der Änderung der allelen Verteilung unter den N-Chromosomen zuzuschreiben (vergleiche Tabelle 2). Diese Daten stimmen daher mit den aus der Studie rekombinanter Familien erhaltenen Ergebnis überein (vergleiche **Fig. 14**). Eine ähnliche Schlussfolgerung konnte durch Überprüfen der extendierten DNA-Macker-Haplotypen, die mit den ZF-Chromosomen assoziiert sind, gezogen werden (vergleiche unten). Jedoch erlaubte die starke Allel-Assoziation, die über die große physische Entfernung zwischen EG1.4 und H1.3 bestimmt werden ist, kein weiteres Feinkartieren des ZF-Gens. Da J44 der letzte genomische DNA-Klon darstellte, der durch Chromosomenwanderung und Chromosomensprung isoliert worden ist, bevor ein cDNA-Klon identifiziert worden war, veranlasste uns die starke, für das JG2E1-J44-Intervall entdeckte Allel-Assoziation, über dieses gesamte Intervall nach Kandidatengensequenzen zu suchen. Dabei ist es interessant, zu bemerken, dass der höchste Grad an Allel-Assoziation in der Tat zwischen ZF und den 2 durch 10-1X.62 entdeckten RFLPs, einer Region in der Nähe der Haupt-ZF-Mutation, entdeckt worden ist.

[0154] Tabelle 4 zeigt eine paarweise Allel-Assoziation zwischen mit ZF eng verkoppelten DNA-Mackern. Die durchschnittliche Anzahl von Chromosomen, die bei diesen Berechnungen hinzugezogen worden ist, betrug 75–80 und es wurden nur Chromosomen aus ZF-PI-Familien beim Scoren von ZF-Chromosomen herangezogen. Ähnliche Ergebnisse wurden dann erhalten, wenn die standardisierte Yule-Assoziation (A) herangezogen wurde.

N-Chromosomen

Tabelle 3.

	meiD	meiH	E6	E7	PH131	W3	D1.4	H2.9A	EG1.4	JG2E1	E2.6	H2.8	E4.1	J44	10-1X.6	T6/20	H1.3	CE1.0	J32	J3.11	J29	
	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH	meiD	meiH
meiD Band	0.35	0.49	0.04	0.04	0.05	0.07	0.27	0.08	0.06	0.07	0.14	0.07	0.09	0.03	0.06	0.10	0.03	0.16	0.05	0.07	0.11	0.02
meiD Tagl	0.21	0.41	0.13	0.15	0.02	0.01	0.02	0.09	0.15	0.11	0.07	0.24	0.03	0.11	0.08	0.02	0.06	0.13	0.15	0.09	0.09	0.05
meiH Band	0.81	0.14	0.01	0.05	0.04	0.05	0.24	0.05	0.09	0.07	0.19	0.15	0.07	0.04	0.02	0.02	0.07	0.02	0.03	0.21	0.04	0.18
E6 Tagl	0.11	0.30	0.00	0.00	0.07	0.04	0.04	0.02	0.03	0.00	0.19	0.02	0.09	0.19	0.09	0.11	0.09	0.15	0.07	0.11	0.20	0.00
E7 Tagl	0.16	0.31	0.02	1.00	0.11	0.09	0.03	0.03	0.04	0.01	0.11	0.00	0.07	0.22	0.01	0.02	0.09	0.13	0.06	0.06	0.16	0.04
PH131 Hind	0.45	0.28	0.23	0.38	0.40	0.01	0.12	0.04	0.09	0.05	0.06	0.03	0.03	0.08	0.16	0.15	0.20	0.04	0.03	0.05	0.08	0.06
W3D1.4 HindIII	0.45	0.29	0.23	0.45	0.47	0.93	0.21	0.02	0.03	0.01	0.06	0.03	0.03	0.10	0.12	0.10	0.23	0.10	0.05	0.05	0.10	0.06
H2.9A Tagl	0.20	0.11	0.15	0.08	0.11	0.38	0.47	0.05	0.11	0.07	0.42	0.14	0.29	0.07	0.27	0.22	0.20	0.09	0.23	0.04	0.08	0.12
EG1.4 Hind	0.11	0.06	0.07	0.06	0.07	0.20	0.24	0.05	0.05	0.87	0.76	0.88	0.81	0.80	0.07	0.13	0.51	0.56	0.04	0.24	0.14	0.15
EG1.4 Bgl	0.09	0.08	0.07	0.08	0.07	0.27	0.40	1.00	0.00	0.92	0.77	0.93	0.71	0.55	0.09	0.07	0.58	0.55	0.12	0.28	0.24	0.20
JG2E1 PstI	0.07	0.08	0.03	0.09	0.08	0.30	0.43	0.93	0.94	0.00	0.84	1.00	0.78	0.64	0.11	0.11	0.61	0.57	0.13	0.31	0.26	0.22
E2.6/E.9 MspI	0.22	0.06	0.07	0.02	0.03	0.20	0.34	0.81	0.82	0.92	0.00	0.83	0.97	0.78	0.56	0.52	0.47	0.70	0.32	0.31	0.25	0.22
H2.8 NcoI	0.05	0.07	0.01	0.06	0.06	0.31	0.45	0.92	0.93	1.00	0.92	0.00	0.74	0.65	0.13	0.18	0.60	0.59	0.10	0.28	0.28	0.19
E4.1 MspI	0.12	0.06	0.07	0.05	0.03	0.25	0.48	0.82	0.86	0.94	1.00	0.93	0.71	0.49	0.49	0.49	0.68	0.35	0.27	0.25	0.21	0.21
J44 XbaI	0.18	0.05	0.06	0.01	0.01	0.28	0.28	0.45	0.71	0.80	0.90	0.80	0.85	0.00	0.50	0.40	0.65	0.64	0.32	0.24	0.22	0.23
10-1X.6 AccI	0.16	0.10	0.24	0.10	0.11	0.42	0.42	0.64	0.64	0.84	0.70	0.69	0.69	0.59	0.00	0.91	0.19	0.38	0.56	0.00	0.02	0.03
10-1X.6 HaeIII	0.16	0.10	0.25	0.08	0.11	0.41	0.41	0.65	0.64	0.64	0.70	0.69	0.69	0.69	1.00	0.00	0.18	0.43	0.62	0.02	0.02	0.08
T6/20 MspI	0.27	0.07	0.36	0.19	0.13	0.23	0.29	0.06	0.09	0.01	0.07	0.02	0.01	0.11	0.69	0.69	0.00	0.56	0.03	0.21	0.18	0.25
H1.3 NcoI	0.08	0.08	0.08	0.03	0.01	0.30	0.30	0.55	0.71	0.87	0.90	0.87	0.83	0.92	0.64	0.64	0.12	0.40	0.19	0.13	0.20	0.20
CE1.0 NcoI	0.00	0.04	0.02	0.11	0.11	0.25	0.25	0.08	0.09	0.55	0.43	0.55	0.37	0.44	0.24	0.24	0.07	0.40	0.19	0.20	0.14	0.14
J32 SacI	0.03	0.13	0.07	0.17	0.13	0.17	0.24	0.07	0.21	0.24	0.22	0.24	0.21	0.21	0.27	0.28	0.13	0.21	0.18	0.00	0.64	0.97
J3.11 MspI	0.14	0.11	0.15	0.07	0.06	0.05	0.12	0.11	0.10	0.13	0.18	0.19	0.15	0.20	0.28	0.28	0.24	0.14	0.07	0.81	0.00	0.71
J29 PvuII	0.11	0.12	0.09	0.10	0.10	0.00	0.09	0.10	0.10	0.14	0.17	0.20	0.16	0.16	0.29	0.29	0.23	0.18	0.06	0.85	0.97	0.00

ZF-Chromosomen

[0155] Eine starke Allel-Assoziation wurde auch unter Untergruppen von RFLPs an beiden der ZF- und N-Chromosomen entdeckt. Wie in Tabelle 4 gezeigt, scheinen DNA-Marker, die physisch eng beieinander liegen, im Allgemeinen eine starke Assoziation miteinander zu haben. Beispielsweise wurde eine starke (in vielen Fällen fast vollständige) Allel-Assoziation zwischen den nebeneinander liegenden Markern E6 und E7, zwischen PH13 und W3D1.4, zwischen den durch 10-1X.6 bestimmten polymorphen Accl- und HaeIII-Stellen und

unter EG1.4, JG2E1, E2.6(E.9), E2.8 und E4.1 entdeckt. Die beiden Gruppen von Distalmarkern in der MET- und D7S8-Region haben auch einen gewissen Grad an Verkoppelungsungleichgewicht untereinander gezeigt, jedoch zeigten sie kaum eine Assoziation mit Markern von E6 bis CE1.0, was mit den entfernten Positionen von MET und D7S8 vereinbar ist. Andererseits kann das Fehlen einer Assoziation zwischen DNA-Markern, die einander physisch nahestehen, das Vorliegen von Rekombinations-Hotspots anzeigen. Beispiele dieser potentiellen Hotspots sind die Region zwischen E7 und pH131, um N2.3A herum, zwischen J44 und den Regionen, über die sich die Sonden 10-1X.6 und T6/20 erstrecken (siehe **Fig. 14**). Diese Regionen, die häufig Rekombinationsbruchpunkte enthalten, sind bei der darauffolgenden Analyse extendierter Haplotypdaten für die ZF-Region nützlich gewesen.

3.5 HAPLOTYPANALYSE

[0156] Die extendierten Haplotypen auf der Basis von 23 DNA-Markern wurden für die ZF- und N-Chromosomen in der Sammlung von Familien gebildet, die vorher für die Verkoppelungsanalyse herangezogen worden waren. Aufgrund der Annahme einer Rekombination zwischen Chromosomen verschiedener Haplotypen war es möglich, mehrere Linien beobachteter ZF-Chromosomen zu konstruieren und die Position des Krankheitslocus vorauszusagen.

[0157] Um weitere, für das Verständnis der Natur verschiedener ZF-Mutationen nützte Informationen zu erhalten, wurden die F508-Deletionsdaten mit den extendierten DNA-Marker-Haplotypen korreliert. Wie in Tabelle 5 gezeigt, konnten fünf Hauptgruppen von N- und ZF-Haplotypen durch die RFLPs innerhalb oder direkt neben des putativen ZF-Gens (Regionen 6–8) definiert werden.

dien herangezogenen Familien (man vergleiche Anmerkung (a) zu der Tabelle 3), wobei zusätzliche ZF-PS-Familien hinterher zusammengebracht wurden, abgeleitet (Kerem et al. Am. J. Genet. 44: 827 (1989)). Um Platz zu sparen, sind die Daten in Gruppen (Regionen) aufgezeigt. Die Regionen sind hauptsächlich in Tabelle 4 gezeigten paarweisen Assoziationsdaten entsprechend zugeordnet, wobei die Regionen 6–8 die putativen ZF-Loci überspannen (die F508-Deletion befindet sich zwischen den Regionen 6 und 7). Ein Bindestrich (-) wird an der Region gezeigt, an der der Haplotyp aufgrund unvollständiger Daten oder der Unfähigkeit, eine Phase zu etablieren, nicht bestimmt worden ist. Alternative Haplotypzuordnungen sind auch dort angegeben, wo die Daten unvollständig sind. Nicht klassifiziert umschließt diejenigen Chromosomen mit mehr als 3 unbekanntem Zuordnungen. Die Haplotypdefinitionen für jede der 9 Regionen sind wie folgt:

Region I-	metD <u>BanI</u>	metD <u>TaqI</u>	metH <u>TaqI</u>	
A=	1	1	1	
B=	2	1	2	
C=	1	1	2	
D=	2	2	1	
E=	1	2	-	
F=	2	1	1	
G=	2	2	2	
Region 2 -	E6 <u>TaqI</u>	E7 <u>TaqI</u>	pH131 <u>HinfI</u>	W3D1,4 <u>HindIII</u>
A=	1	2	2	2
B=	2	1	1	1
C=	1	2	1	1
D=	2	1	2	2
E=	2	2	2	1
F=	2	2	1	1
G=	1	2	1	2
H=	1	1	2	2

Region 3 -	H2,3A <u>TaqI</u>		
A=	1		
B=	2		
Region 4-	Eg1,4 <u>HincII</u>	EG1,4 <u>BGI</u>	JG2E1 <u>PstI</u>
A=	1	1	2
B=	2	2	1
C=	2	2	2
D=	1	1	1
E=	1	2	1
Region 5 -	E2,6 <u>MspI</u>	E2,8 <u>NcoI</u>	E4,1 <u>MspI</u>
A=	2	1	2
B=	1	2	1
C=	2	2	2
Region 6-	J44 <u>XbaI</u>	10-1X. <u>AccI</u>	610-1X.6 <u>HaeIII</u>
A=	1	2	1
B=	2	1	2
C=	1	1	2
D=	1	2	2
E=	2	2	2
F=	2	2	1
Region 7 -	T6/20 <u>MspI</u>		
A=	1		
B=	2		
Region 8 -	H1,3 <u>NcoI</u>	CE 1,0 <u>NdeI</u>	
A=	2	1	
B=	1	2	
C=	1	1	
D=	2	2	

Region 9 -	J32 <u>SacI</u>	J3,11 <u>MspI</u>	J29 <u>PvuII</u>
A=	1	1	1
B=	2	2	2
C=	2	1	2
D=	2	2	1
E=	2	1	1

(b) Anzahl von Chromosomen, die in jeder Klasse angezeigt sind:

ZF-PI(F) = ZF-Chromosomen aus ZF-PI-Patienten mit der F508-Deletion;

ZF-PS(F) = ZF-Chromosomen aus ZF-PS-Patienten mit der F508-Deletion;

ZF-PI = andere ZF-Chromosomen aus ZF-PI-Patienten;

ZF-PS = andere ZF-Chromosomen aus ZF-PS-Patienten;

N = normale Chromosomen, die aus Trägereltern abgeleitet sind.

[0158] Es war offensichtlich, dass die meisten Rekombinationen zwischen Haplotypen zwischen den Regionen 1 und 2 und zwischen den Regionen 8 und 9 vorkamen, was wiederum in gutem Einklang mit der relativ langen physischen Entfernung zwischen diesen Regionen steht. Andere weniger häufige Bruchpunkte wurden zwischen kurzen Entfernungsintervallen beobachtet und sie entsprachen im Allgemeinen den Hotspots, die durch paarweise Allel-Assoziationsstudien, wie oben gezeigt, identifiziert worden sind. Auffallend daran war, dass die F508-Deletion fast ausschließlich mit der Gruppe I, dem häufigsten ZF-Haplotyp, assoziiert ist, was die Auffassung unterstützt, dass diese Deletion die Hauptmutation bei ZF darstellt. Noch wichtiger ist, dass die F508-Deletion zwar in 89% (62/70) der ZF-Chromosomen beim AA-Haplotyp (der den beiden Regionen 6 und 7 entspricht), der die Deletion flankiert, beobachtet worden ist, jedoch keine bei den 14 N-Chromosomen innerhalb der gleichen Gruppe ($\chi^2 = 47.3$, $p < 10^{-4}$) vorgefunden worden ist. Die F508-Deletion war deshalb kein allgemeiner Sequenzpolymorphismus, der mit dem Kern des Gruppe I-Haplotyps assoziiert ist (vergleiche Tabelle 5).

[0159] Eines der ZF-Chromosomen, die durch die spezifische Oligonucleotidsonde für die F508-Deletion bestimmt worden ist, erwies sich als zu einer anderen Haplotypgruppe gehörend (Gruppe III). Weder irgendeines der 9 anderen ZF-Chromosomen noch die 17 N-Chromosomen mit der gleichen Gruppe hybridisierten zu der Sonde. Dieses spezifische Hybridisierungsergebnis deutet darauf hin, dass die an diesem Chromosom untergebrachte Mutation F508 ähnlich ist. Obwohl eine Rekombination oder Genumwandlung eventuell Mechanismen sind, die zum Erklären des Vorliegens dieser Deletion an einem Nicht-Gruppen I-Haplotyp dienen können, ist es wahrscheinlicher, dass diese 2 Gruppen III-Chromosomen einen sich wiederholenden Mutationsfall darstellen, d. h. eine Situation, die den β^S und β^E -Mutationen am β -Globinlocus ähnlich sind.

[0160] Zusammengenommen untermauern die Ergebnisse der Oligonucleotidhybridisierungsstudie und der Haplotypanalyse die Tatsache, dass der hier beschriebene GenLocus das ZF-Gen ist und dass die 3 bp-(F508)Deletion die häufigste Mutation bei ZF ist.

3.6 ANDERE ZF-MUTATIONEN

[0161] Die Assoziation der F508-Deletion mit einem allgemeinen und einem seltenen ZF-Haplotyp hat weitere Einsicht geboten bezüglich der Anzahl von Mutationsvorkommnissen, die zur gegenwärtigen Patientenspopulation beitragen könnten. Auf der Basis der extendierten Haplotypdaten tragen die 2 ursprünglichen Chromosomen, bei denen die F508-Deletion vorgekommen ist, wahrscheinlichweise den Haplotyp – AAAAAA – (Gruppe Ia) und -CBAACBA-(Gruppe IIIa), wie in Tabelle 5 definiert. Die anderen, die Deletion tragenden Gruppe I-ZF-Chromosomen sind wahrscheinlich Rekombinationsprodukte, die vom ursprünglichen Chromosom abgeleitet sind. Betrachtet man die ZF-Chromosomen in jeder Haplotypgruppe als vom gleichen Ursprung abgeleitet, so würden nur 3–4 zusätzliche Mutationsvorkommen vorausgesagt werden (Tabelle 5). Da jedoch viele der ZF-Chromosomen in der gleichen Gruppen deutlich voneinander verschieden sind, sind weitere Unterteilungen innerhalb jeder Gruppe möglich. Dadurch könnte eine größere Anzahl unabhängiger Mutationsvorkommnisse in Betracht gezogen werden, und die Daten weisen darauf hin, dass mindestens 7 zusätzliche putative Mutationen ebenfalls zum ZF-PI-Phenotyp beitragen (vergleiche Tabelle 4). Die zu den ZF-PS-Untergruppen führenden Mutationen sind wahrscheinlichweise heterogener.

[0162] Die 7 zusätzlichen ZF-PI-Mutationen sind durch folgende Haplotypen dargestellt: -CAAAAA- (Gruppe Ib), -CABCAAD- (Gruppe Ic), ---BBBAC- (Gruppe IIa), -CABBBAB- (Gruppe Va). Obwohl der Molekulardefekt bei jeder dieser Mutationen noch zu definieren ist, ist es klar, dass keine dieser Mutationen die Region stark beeinflusst, die den Oligonucleotidbindungsstellen entspricht, die bei dem PKR-Hybridisierungsexperiment

verwendet worden sind.

3.7 PANKREASSUFFIZIENZ

[0163] CF-PS wird klinisch als ausreichende Pankreasexokrinfunktion für das Verdauen von Nahrung definiert; jedoch ist der Grad residueller Pankreasenzymeaktivität im Verdauungssystem von Patient zu Patient verschieden. Frühere Haplotypdaten haben darauf hingedeutet, dass die ZF-PI- und ZF-PS-Patienten auf verschiedene mutante Allele zurückzuführen waren. Obwohl der grundlegende biochemische Defekt bei ZF noch definiert werden muss, ist es möglich, dass die residuelle Pankreasenzymaktivität bei ZF-PS-Patienten die Aktivität des mutanten ZF-Genprodukts direkt widerspiegelt. So würde die residuelle Exokrinfunktion, die durch ein mildes (ZF-PS-) Allel übertragen wird – obgleich sie viel geringer als diejenige des normalen Genprodukts ist – einen dominanten Phänotyp im Vergleich mit demjenigen stärkerer (ZF-PI) Mutationen mit kaum einer oder keiner Funktionsfähigkeit darstellen. Daraus kann der Schluss gezogen werden, dass nur Patienten, die Träger von 2 Kopien starker Allele sind, ZF-PI-Patienten sein würden und dass Patienten, die 1 oder 2 milde Allele tragen, ZF-PS-Patienten wären.

[0164] Um die obigen Thesen zu überprüfen, könnten Informationen bezüglich des Anteils von ZF-Patienten, die Träger der F508-Deletion sind, verwendet werden. Unter der Annahme, dass eine starke Mutation im Vergleich mit einer milden Mutation rezessiv ist und dass eine Verteilung von ZF-Allelen unter der Patientenpopulation dem Hardy-Weinberg-Gesetz entspricht, könnte die Häufigkeit schwerer Allele auf 0,92 geschätzt werden und diejenige der milden Allele (M), auf 0,08 (vergleiche Tabelle 6).

TABELLE 6
POPULATIONSANALYSE VON ZF-PI UND ZF-PS

	<u>Ange-</u> <u>nommener</u> <u>Genotyp (a)</u>	<u>vorher-</u> <u>gesagte</u> <u>Häufig-</u> <u>keit (b)</u>	<u>beobachtet (c)</u>	<u>erwartet (d)</u>
Pankreas- insuffizienz (PI)	FF	0,459	21	21,2
	FS	0,331	14	15,2
	SS	0,060	4	2,7
	Gesamt	0,850	39	-
Pankreas- suffizienz (PS)	FN	0,106	15 (a)	14,8
	SN	0,038	6	6,2
	MM	0,006		
	Gesamt	0,150	21	

(a) Allelbezeichnungen: F = die 3 bp-Deletion (Deletion von Phenylalanin an der Aminosäureposition 508); S = uncharakterisierte stark mutante Allele; M = uncharakterisierte mild mutante Allele.

(b) Angenommen, dass der mutante ZF-PI-Phänotyp bezüglich des mutanten ZF-PS-Phänotyps rezessiv ist, so könnte die Häufigkeit der mutanten ZF-PI-Allele, einschließlich der 3 bp-Deletion, aufgrund des beobachteten Anteils der ZF-PI-Patienten in der ZF-Klinik geschätzt werden (Corey et al J. Pediatr. 115: 274 (1989)), d. h. $(0,85) = 0,92$. Die beobachtete Allelhäufigkeit für F in der gesamten ZF-Population beträgt 0,68 (Tabelle 3); die Häufigkeit für S = $0,92 - 0,68 = 0,24$; die Häufigkeit für M = $1 - 0,92 = 0,08$. Die Häufigkeit für jeden Genotyp wurde daraufhin unter Zuhilfenahme des Hardy-Weinberg-Gesetzes berechnet.

(c) Die Anzahl von ZF-PI- und ZF-PS-Patienten in jeder Kategorie wurde durch Oligonucleotid-Hybridisierungsanalyse, wie in Fig. 15 veranschaulicht, erhalten. Die Patienten stammten aus den bei unserer Verkopplungsanalyse herangezogenen ZF-Familien, wobei 14 zusätzliche ZF-PS-Patienten/Familien aus einer darauffolgenden Studie stammten. Da SM und MM nicht genotypisch oder phenotypisch unterschieden werden konnte, wurden sie bei der Analyse kombiniert.

(d) Die erwarteten Zahlen wurden für ZF-PI und ZF-PS nach der Standardisierung innerhalb jeder Gruppe berechnet. Das χ^2 der Passung beträgt 0,86, d. f. = 3, $0,74 < p < 0,90$.

(e) Diese Zahl ist höher als zu erwarten gewesen wäre (15 beobachtete gegen 9,6 zu erwartende), wenn

die F508-Deletion unter allen ZF-Chromosomen im Hardy-Weinberg-Gleichgewicht steht ($\chi^2 = 6,48$, d. f. = 1, $p < 0,011$).

[0165] Da es sich herausstellte, dass die Mehrzahl der ZF-PI-Patienten für die F508-Mutation (F) homozygot war, war es vernünftig anzunehmen, dass diese Mutation einem der starken Allele entsprach. Angesichts der bei der untersuchten ZF-Population beobachteten Häufigkeit von F (0,68) konnte die Häufigkeit der verbleibenden starken Allele (8) abgeleitet werden. Der Anteil von FF-, SS-, MM-, FS-, FM- und SM-Patienten wurde berechnet. Da Individuen mit SM und MM phenotypisch oder genotypisch nicht unterschieden werden konnten, wurden sie bei der Analyse kombiniert. Wie in Tabelle 6 gezeigt, waren die für alle 5 Patientengruppen beobachteten Häufigkeiten wie es aufgrund dieser Hypothese zu erwarten gewesen war.

[0166] Die obige Analyse bietet daher einen starken Rückhalt für unsere Annahme, dass ZF-PI der Gegenwart von zwei starken Allelen zuzuschreiben ist und dass ein ZF-PS-Patient Träger entweder eines starken Allels oder von zwei milden Allelen ist. Dieses Modell erklärt auch die geringere Häufigkeit der F508-Deletion bei der ZF-PS- als bei der ZF-PI-Population und die übermäßige Anzahl von ZF-PS-Patienten mit einer Kopie der Deletion (vergleiche Anmerkung in Tabelle 6).

[0167] Angesichts des vorausgesagten dominanten Phenotyps, der durch die M-Allele übertragen wird, war es notwendig, die ZF-Chromosomen bei ZF-PS-Patienten einzeln zu untersuchen, um diejenigen, die die M-Allele tragen, zu identifizieren. Wie in Tabelle 7 gezeigt, sind fünf der sieben repräsentativen ZF-PS-Patienten Träger einer Kopie der F508-Deletion; mindestens 5 verschiedenen Haplotypen konnten den anderen ZF-Chromosomen zugeordnet werden.

TABELLE 7. Haplotyp der ZF-Chromosomen in ZF-PS-Individuen und Familien mit MI

Familie#	1	2	3	4	5	6	7	8	9	CF-Allele
(a) CF-PS-Individuen										
3	A D	A C	A B	A A	A A	A C	- B	- A	A A	F (Gruppe Ia) M (vorausgesagt, Gruppe IIIb)
14	B B	A C	A B	A B	- -	A B	A A	A C	A C	F (Gruppe Ia) M (vorausgesagt, Gruppe IIa)
27	A A	B C	- -	A A	A A	A A	A A	A A	E A	F (Gruppe Ia) M (vorausgesagt, Gruppe Ib)
29	A B	C A	- -	C B	- -	A B	A A	A A/C	B B	F (Gruppe Ia) M (vorausgesagt, Gruppe IIa)
40	D F	A C	A B	A A	A A	A C	A B	A C	B A	P (Gruppe Ia) N (vorausgesagt, Gruppe IV)
51	C	C	A	B	B	B/C	A	CA		M (vorausgesagt, Gruppe IIa)
	F	D	A	B	B	B/C	A	CC		M (vorausgesagt, Gruppe IIa)
54	B	C	A	B	C	C	A	C	A	M oder S vorausgesagt, Gruppe Vb)

B B A A A C B A A M (vorausgesagt,
Gruppe IIIb)

(b) Familien mit MI

4	B	A	A	A	A	A	A	A	F (Gruppe Ia)
	B	A	A	A	A	A	A	A	F (Gruppe Ia)
10	D	B	A	A	-	A	-	A	F (Gruppe Ia)
	A	D	A	A	-	A	A	A	F (Gruppe Ia)
23	A	E	B	A	A	A	A	A	F (Gruppe Ia)
	B	C	A	A	A	A	A	A	S (vorausgesagt, Gruppe Ib)
28	A	A	-	A	A	A	A	A	F (Gruppe Ia)
	A	A	-	A	A	A	A	A	F (Gruppe Ia)
33	B	B	-	A	A	A	-	A	F (Gruppe Ia)
	B	A	-	A	A	A	A	A	F (Gruppe Ia)
49	A	A	A	A	A	A	A	A	F (Gruppe Ia)
	A	A	A	A	A	A	A	A	F (Gruppe Ia)

(a) Die Haplotypdefinitionen sind die gleichen wie in Tabelle 5.

(b) Die Allelbezeichnungen sind die gleichen wie in Tabelle 6: F = die F508-Deletion; S = uncharakterisiertes stark mutantes Allel; M = uncharakterisiertes mild mutantes Allel.

[0168] Diese letzteren Beobachtungen bieten eine weitere Untermauerung der Annahme, dass die Mehrzahl von ZF-PS-Patienten verbindungsheterozygot sind.

4.0 ZFTR-PROTEIN

[0169] Wie bezüglich der DNA-Sequenz von **Fig. 1** besprochen, wurde bei der Analyse der Sequenz der sich überlappenden cDNA-Klone ein unverarbeitetes Polypeptid von 1480 Aminosäuren mit einer Molmasse von 168.138 Dalton vorausgesagt. Wie später beschrieben kann aufgrund von Polymorphismen im Protein die Molmasse des Proteins durch mögliche Substitutionen oder die Deletion von gewissen Aminosäuren variabel sein. Die Molmasse ändert sich auch aufgrund des Anlagerns von Kohlenhydrateinheiten unter Bildung eines Glykoproteins. Es ist auch klar, dass das funktionelle Protein in der Zelle dem unverarbeiteten Polypeptid ähnlich sein wird, jedoch aufgrund eines Zellmetabolismus modifiziert werden kann.

[0170] Dementsprechend bietet die Erfindung ein gereinigtes normales ZFTR-Polypeptid, das durch eine Molmasse von 170.000 Dalton gekennzeichnet ist und eine Epithelzellen-Transmembranionen-Konduktanzaktivität aufweist. Das normale ZFTR-Polypeptid, das im Wesentlichen von anderen menschlichen Proteinen frei ist, ist durch die oben erwähnten DNA-Sequenzen kodiert und einer Ausführungsform gemäß derjenigen von **Fig. 1**. Ein derartiges Polypeptid weist die immunologische oder biologische Aktivität von normalem ZFTR-Polypeptid auf. Wie weiter unten noch besprochen werden wird, kann das ZFTR-Polypeptid und Fragmente desselben durch chemische oder enzymatische Peptidsynthese hergestellt oder in einem geeigneten gezüchteten Zellsystem exprimiert werden. Die Erfindung bietet auch ein gereinigtes mutantes ZFTR-Polypeptid, das durch eine mit der zystischen Fibrose assoziierte Aktivität in menschlichen Epithelzellen gekennzeichnet ist. Ein derartiges Mutantes ZFTR-Polypeptid, kann, da es im wesentlichen von anderen menschlichen Proteinen frei ist, durch die mutante DNA-Sequenz kodiert werden.

4.1 STRUKTUR VON ZFTR

[0171] Das charakteristischste Merkmal des vorhergesagten Proteins ist das Vorliegen von zwei wiederholten Motiven, von denen jedes aus einem Satz von Aminosäureresten besteht, die in der Lage sind, die Membran

mehrere Male zu überspannen, gefolgt von sequenzähnlichen Konsens-Nucleotid(ATB)-Bindungsfalten (NBFs) (**Fig. 11, 12 und 16**). Diese charakteristischen Merkmale sind denjenigen des gegen eine Reihe von Arzneimitteln widerstandsfähigen P-Glykoproteins von Säugern und einer Reihe anderer membranassoziiierter Proteine bemerkenswert ähnlich, was darauf hinweist, dass das vorhergesagte ZF-Genprodukt wahrscheinlich beim Transport von Substanzen (Ionen) über die Membran eine Rolle spielt und wahrscheinlich ein Mitglied einer Membranprotein-Superfamilie ist.

[0172] **Fig. 13** zeigt ein schematisches Modell des vorhergesagten ZFTR-Proteins. In **Fig. 13** zeigen Zylinder Membranen überspannende Wendel, die schraffierten Kreise zeigen NBFs an. Der punktierte Kreis ist die polare R-Domäne. Die sechs Membranen überspannenden Wendel in jeder Hälfte des Moleküls sind als Zylinder gezeigt. Die inneren zytoplasmisch orientierten NBFs sind als schattierte Kreise mit Schlitzen aufgezeigt, um die Eintrittsmöglichkeit durch das Nucleotid anzuzeigen. Die große polare R-Domäne, die die beiden Hälften verbindet, ist durch einen punktierten Kreis dargestellt. Geladene einzelne Aminosäuren innerhalb der Transmembransegmente und an der Oberfläche der R-Domäne sind als kleine das Ladungszeichen enthaltende Kreise dargestellt. Nettoladungen an den internen und externen Schleifen, die die Membranzylinder verbinden, und an den Regionen der NBFs sind durch offene Quadrate umschlossen. Stellen für die Phosphorylierung durch die Proteinkinasen A oder C sind durch ausgefüllte bzw. offene Dreiecke dargestellt. K, R, H, D und E sind die Standardnomenklatur für die Aminosäuren Lysin, Arginin, Histidin, Asparaginsäure bzw. Glutaminsäure.

[0173] Jede der vorhergesagten membranassoziierten Regionen des ZFTR-Proteins besteht aus 6 stark hydrophoben Segmenten, die in der Lage sind, eine Lipiddoppelschicht den Algorithmen von Kyte und Doolittle und von Garnier et al (J. Mol. Biol. 12D, 97 (1978) (**Fig. 13**) zu überspannen. Auf die membranassoziierten Regionen folgt jeweils eine große hydrophile, die NBFs enthaltende Region. Auf der Basis der Sequenzausrichtung mit anderen bekannten Nucleotid bindenden Proteinen umfasst eine jede der putativen NBFs im ZFTR mindestens 150 Reste (**Fig. 13**). Die bei den meisten der ZF-Patienten entdeckte 3 bp-Deletion ist zwischen den 2 am besten erhaltenen Segmenten der ersten NBF in ZFTR positioniert. Die Aminosäuresequenzidentität zwischen der Region, die die Phenylalanin-Deletion umgibt, und den entsprechenden Regionen einer Reihe anderer Proteine legt nahe, dass diese Region von funktioneller Wichtigkeit ist (**Fig. 16**). Eine hydrophobe Aminosäure, gewöhnlich eine mit einer aromatischen Seitenkette, liegt in den meisten dieser Proteine an der Position vor, die F508 des ZFTR-Proteins entspricht. Man muss sich im Klaren darüber sein, dass Aminosäurepolymorphismen als Ergebnis von DNA-Polymorphismen existieren können.

[0174] **Fig. 16** zeigt die Ausrichtung der 3 am besten konservierten Segmente der erweiterten NBFs von ZFTR auf vergleichbare Regionen anderer Proteine. Diese 3 Segmente bestehen aus den Resten 433–473, 488–513 und 542–584 der N-terminalen Hälfte und 1219–1259, 1277–1302 und 1340–1382 der C-terminalen Hälfte von ZFTR. Die dicke Kopflinie weist auf die Regionen der größten Ähnlichkeit hin. Eine zusätzliche allgemeine Homologie ist selbst ohne Einführung von Leerstellen ersichtlich.

[0175] Trotz der Symmetrie der Struktur des Proteins insgesamt und der Sequenzkonservation der NBFs ist die Sequenzhomologie zwischen den beiden Hälften des vorausgesagten ZFTR-Proteins bescheiden. Dies wird in **Fig. 12** bewiesen, wo die Aminosäure 1–1480 auf jeder Achse dargestellt sind. Linien auf jeder Seite der Identitätsdiagonale zeigen die Positionen interner Ähnlichkeiten an. Aus diesem Grund sind zwar vier Sätze interner Sequenzidentitäten, wie in **Fig. 12** gezeigt, mit Hilfe der Dayhoff Score-Matrix, wie von Lawrence et al. [C. B. Lawrence, D. A. Goldman und R. T. Hood, Bull Math Biol. 48, 569 (1986)] angewendet, erfassbar, drei von diesen sind jedoch nur bei den niedrigen Schwellenwerteinstellung für die Standardabweichung ersichtlich. Die starke Identität liegt zwischen Sequenzen an den Carboxylenden der NBFs vor. Von den 66 ausgerichteten Resten sind 27% identisch und weitere 11% funktionell ähnlich. Die insgesamt schwache interne Homologie steht im Gegensatz zu dem viel höheren Grad (> 70%) im P-Glykoprotein, wofür eine Genduplikationshypothese vorgeschlagen worden ist (Gros et al, Cell 47, 371, 1986; C. Chen et al, Cell 47, 381, 1986, Gerlach et al, Nature, 324, 485, 1986, Gros et al, Mol. Cell. Biol. 8, 2770, 1988). Der Mangel an Konservierung in den relativen Positionen der Exon-Intron-Grenzen kann eventuell gegen ein derartiges Modell für ZFTR sprechen (**Fig. 2**).

[0176] Da keine Signal-Peptidsequenz am Aminoterminus von ZFTR vorzuliegen scheint, ist das hochgeladene hydrophile Segment, das der ersten Transmembransequenz vorausgeht, wahrscheinlich im Zytoplasma orientiert. Es ist zu erwarten, dass jeder der zwei Sätze hydrophober Wendel 3 die Membran überquerende Schleifen bildet und es ist zu erwarten, dass wenig der Sequenz des gesamten Proteins der äußeren Oberfläche gegenüber ausgesetzt ist, mit Ausnahme der Region zwischen den Transmembransegmenten 7 und 8. Es ist interessant, festzustellen, dass die letztere Region zwei potentielle Stellen für die N-verkoppelte Glykosylierung enthält.

[0177] Auf jede der membranassoziierten Regionen folgt eine NBF, wie oben angezeigt. Außerdem kann eine hochbeladene zytoplasmische Domäne in der Mitte des vorausgesagten ZFTR-Polypeptids identifiziert werden, die die beiden Hälften des Proteins verbindet. Diese Domäne, nämlich die R-Domäne, ist operationell durch ein einziges großes Exon definiert, in dem 69 der 241 Aminosäuren polare Reste sind, die in sich in ab-

wechselnden Gruppen positiver und negativer Ladung angeordnet sind.

[0178] Außerdem befinden sich 9 der 10 Konsenssequenzen, die für die Phosphorylierung durch die Proteinkinase A (PKA) erforderlich sind, und 7 der potentiellen Substratstellen für die Proteinkinase C (PKC), die in ZFTR aufzufinden sind, in diesem Exon.

4.2 DIE FUNKTION VON ZFTR

[0179] Die Eigenschaften von ZFTR können aus einem Vergleich mit anderen membranassoziierten Proteinen abgeleitet werden (**Fig. 16**). Außen der allgemeinen strukturellen Ähnlichkeit mit dem P-Glykoprotein von Säugern, weist eine jeder der beiden vorausgesagten Domänen im ZFTR auch eine bemerkenswerte Ähnlichkeit mit der Eindomänenstruktur von Hämolyysin B von *E. coli* und dem Produkt des White-Gens von *Drosophila* auf. Diese letzteren Proteine spielen beim Transport des lytischen Peptids des Hämolyysin-Systems bzw. von Augenpigmentmolekülen eine Rolle. Das Vitamin B₁₂-Transportsystem von *E. coli*, BtuD und MbpX, bei dem es sich um ein Leberblümchen-Chloroblastgen handelt, dessen Funktion unbekannt ist, weist ebenfalls ein ähnliches Strukturmotiv auf. Des Weiteren hat das ZFTR-Protein eine strukturelle Ähnlichkeit mit mehreren der periplasmischen Gelöststoff-Transportsysteme Gram-negativer Bakterien gemeinsam, bei denen die Transmembranregion und die ATP-Bindungsfalten in getrennten Proteinen enthalten sind, die im Zusammenspiel mit einem dritten substratbindenden Polypeptid funktionieren.

[0180] Die gesamte strukturelle Anordnung der Transmembrandomänen in ZFTR ist derjenigen mehrerer Kationenkanalproteine und einiger Kationentranslokations-ATPasen sowie der vor kurzem beschriebenen Adenylatcyclase von Rinderhirn ähnlich. Die funktionelle Signifikanz dieser topologischen Klassifikation, die aus 6 Transmembrandomänen besteht, bleibt auch weiter hin rein theoretisch.

[0181] Kurze Regionen einer Sequenzidentität sind auch zwischen den putativen Transmembranregionen von ZFTR und anderen, Membranen überspannenden Proteinen entdeckt worden.

[0182] Interessanterweise gibt es auch Sequenzen einer Länge von 18 Aminosäure, die ca. 50 Reste vom Carboxyterminus von ZFTR und dem raf-Serin/Threoninkinaseprotoonkogen auf *Xenopus laevis* entfernt gelegen sind, die bei 12 dieser Positionen identisch sind.

[0183] Schließlich ist eine Aminosäuresequenzidentität (10–13 konservierte Reste) zwischen einem hydrophilen Segment (Position 701–713) innerhalb der hochbeladene n R-Domäne von ZFTR und einer Region beobachtet worden, die direkt vor der ersten Transmembranschleife der Natriumkanäle sowohl im Rattenhirn als auch beim Aal liegt. Das topologisch eng verwandte P-Glykoprotein hat die geladene R-Domäne von ZFTR nicht gemeinsam: das die 241 Aminosäure verkoppelnde Peptid scheint der Hauptunterschied zwischen den beiden Proteinen zu sein.

[0184] Zusammenfassend zeigen die Merkmale der primären Struktur des ZFTR-Proteins an, dass es Eigenschaften besitzt, auf Grund derer es geeignet ist, an der Regulierung und Kontrolle des Innentransports in den Epithelzellen von mit ZF befallenen Geweben teilzunehmen. Die feste Anbindung an die Membran in zwei Regionen dient dazu, ihre drei intrazellulären Hauptdomänen (die Nucleotid bindenden Falten 1 und 2 und die R-Domäne) in die Nähe der zytoplasmischen Oberfläche der Zellmembran zu positionieren, wo sie die Innenbewegung durch Kanäle modulieren können, die entweder durch ZFTR-Transmembransegmente selbst oder durch andere Membranproteine gebildet sind.

[0185] Angesichts der genetischen Daten, der Gewebespezifität und der vorausgesagten Eigenschaften des ZFTR-Proteins kann vernünftigerweise der Schluss gezogen werden, dass das ZFTR direkt für die ZF verantwortlich ist. Es bleibt jedoch immer noch unklar, auf welche Weise das ZFTR bei der Regulierung der Ionenkonduktanz über die apikale Membran von Epithelzellen hinweg eine Rolle spielt.

[0186] Es ist möglich, dass das ZFTR als Innenkanal selbst dient. Wie in **Fig. 13** gezeigt, enthalten 10 der 12 Transmembranregionen eine oder mehrere Aminosäuren mit geladenen Seitenketten, eine Eigenschaft, die dem Natriumkanal des Gehirns und den GABA-Rezeptorchloridkanal-Untereinheiten ähnlich ist, wo geladene Rückstände in vier der sechs bzw. drei der vier membranassoziierten Domänen pro Untereinheit oder Wiederholeinheit vorliegen. Man glaubt, dass die amphipathische Natur dieser Transmembransegmente zur Kanalbildungsfähigkeit dieser Moleküle beiträgt. Als Alternative ist das ZFTR eventuell kein Innenkanal sondern dient stattdessen zum Regulieren der Ionenkanalaktivitäten. Als Untermauerung letzterer Annahme scheinen keine der gereinigten Polypeptide aus der Luftröhre und den Nieren, die in der Lage sind, Chloridkanäle in Lipidmembranen zu rekonstituieren [Landry, et al, *Science* 224: 1469 (1989)] ZFTR zu sein, wenn sie auf der Basis der Molekularmasse beurteilt werden.

[0187] In jedem der beiden Fälle weist das Vorliegen von ATP-bindenden Domänen im ZFTR daraufhin, dass die ATP-Hydrolyse eine direkte Rolle spielt und für die Transportfunktion erforderlich ist. Die hohe Dichte von Phosphorylierungsstellen für PKA und PKC und die Gruppen geladener Reste in der R-Domäne können beide zum Regulieren dieser Aktivität dienen. Die Deletion eines Phenylalaninrest in der NBF kann die richtige Bindung von ATP oder die Konformationsänderung verhindern, die diese normalerweise hervorruft, und deshalb zur beobachteten Unempfindlichkeit gegen Aktivierung durch PKA- oder PKC-vermittelte Phosphorylierung

des apikalen ZF-Chloridkonduktanzpfads führen. Da das vorausgesagte Protein mehrere Domänen enthält und zu einer Familie von Proteinen gehört, die häufig als Teil von Multikomponenten-Molekularsystemen funktionieren, kann das ZFTR auch an Epithelgewebefunktionen bezüglich der Aktivität oder Regulierung, die nicht mit dem Ionentransport in Verbindung stehen, teilnehmen.

[0188] Da das isolierte ZF-Gen (cDNA) nun im Griff ist, ist es möglich, den grundlegenden biochemischen Defekt bei ZF zu definieren und die Kontrolle der Innentransportwege in Epithelzellen im Allgemeinen noch weiter aufzuklären. Am wichtigsten ist dabei, dass das bisher aus der vorhergesagten Struktur von ZFTR zusammen mit den zusätzlichen Informationen aus Studien des Proteins selbst erhaltene Wissen die Grundlage bietet für die Entwicklung verbesserter Mittel für die Behandlung dieser Krankheit. Bei derartigen Studien sind Antikörper zum ZFTR-Protein gebildet worden, wie später noch beschrieben wird.

4.3 PROTEINREINIGUNG

[0189] Das ZFTR-Protein kann durch Methoden gereinigt werden, die auf der Basis von Eigenschaften, wie sie durch dessen Sequenz enthüllt worden sind; ausgewählt werden. Da es beispielsweise unverkennbare Eigenschaften eines integralen Membranproteins aufweist, wird eine Membranfraktion der Epithelzellen, in denen es stark exprimiert ist (z. B. der kultivierten Kolonkarzinomzelllinie T84) zuerst unter Zuhilfenahme bewährter Methoden isoliert [J. E. Langridge, et al, *Biochem. Biophys. Acts.* 751: 318 (1983)]. Die peripheren Proteine dieser Membranen sind diejenigen, die durch Extraktion mit hohen Salzkonzentrationen, Mitteln mit hohem pH-Wert oder chaotropischen Mitteln wie Lithiumdiiodosalicylat entfernt werden. Daraufhin unter Zuhilfenahme eines Detergents wie Octylglukosid (Landry, et al, Alle der verbleiben integralen Proteine, einschließlich des ZFTR-Proteins, werden daraufhin unter Zuhilfenahme eines Detergents wie Octylglukoaid)Landry, et al, supra) CHAPS [D. J. Bereo et al, *J. Biol. Chem.* 262: 10613 (1987)] oder anderen Verbindungen ähnlicher Wirkung löslich gemacht. Unter Zuhilfenahme der Nucleotidbindungsdomänen von ZFTR wird daraufhin die Cibacron-blau-[S. T. Thompson et al. *Proc. Nat. Acad. Sci. USA*, 72: 669 (1975)] Affinitätschromatographie zum Binden des ZFTR-Proteins und Entfernen desselben von anderen integralen Proteinen der mit Detergens stabilisierten Mischung verwendet. Da ZFTR ein Glykoprotein ist, kann die Differentiallektin-Chromatographie eine weitere Reinigung herbeiführen [Riordan et al. *J. Biol. Chem.* 254: 1270 (1979)]. Die endgültige Reinigung bis zur Homogenität wird daraufhin unter Zuhilfenahme anderer Standardproteinreinigungsverfahren, z. B. Ionenaustauschchromatographie, Gelpermeationschromatographie, Adsorptionschromatographie oder des isoelektrischen Fokussierens, der Notwendigkeit entsprechend erreicht. Als Alternative werden Einzelschrittreinigungsverfahren wie beispielsweise Immunaaffinitätschromatographie unter Zuhilfenahme immobilisierter Antikörper zum ZFTR-Protein (oder Fragmenten desselben) oder die präparative Polyacrylamidgel-Elektrophorese unter Verwendung fortschrittlicher Instrumentierung wie des "230A HPEC-Systems" von Applied Biosystems verwendet. Auf der Basis der Erfahrung bei der Reinigung von P-Glykoprotein (Riordan et al, supra), einem anderen Mitglied der allgemeinen Kategorie von Nucleotid-bindenden Transport-assoziierten Membranproteinen, wird die Reinigung des ZFTR-Proteins erleichtert.

[0190] Außer der Entfernung von Geweben und Zellen, in denen das ZFTR-Protein stark exprimiert ist, werden ähnliche Verfahren zum Entfernen von ZFTR aus Zellen verwendet, die mit das ZF-Gen (cDNA), wie oben beschrieben, enthaltenden Vektoren transfiziert sind. Proteinprodukte, die aus der Expression modifizierter Versionen der cDNA-Sequenz herrühren, werden auf ähnliche Weise gereinigt. Die Homogenitätskriterien von derart bereitgestelltem Protein umfassen diejenigen, die zum Standard des Gebiets der Proteinchemie gehören, einschließlich der ein- und zweidimensionalen Gelelektrophorese und der N-terminalen Aminosäurebestimmung. Das gereinigte Protein wird bei der weiteren physikalischen biochemischen Analyse zum Bestimmen von Merkmalen seiner sekundären und tertiären Struktur verwendet, um zum Konzipieren von Arzneimitteln beizutragen, um das richtige Funktionieren der mutanten ZF-Formen zu unterstützen. Bei der Zubereitung für die Verwendung bei der Proteintherapie wird die Abwesenheit potentiell toxischer kontaminierender Substanzen in Betracht gezogen. Es wird dabei anerkannt, dass die hydrophobe Natur des Proteins den Einschluss von amphiphilen Verbindungen, wie Detergentien und anderen (J. V. Ambud Kar und P. C. Maloney *J. Biol. Chem.* 261: 10079 (1986)) zu allen Handhabungsstufe notwendig macht.

5.0 ZF-SCREENEN

5.1 DIAGNOSE AUF DER BASIS DER DNA

[0191] Angesichts der Kenntnisse bezüglich der Hauptmutation, wie sie hier offenbart sind, können das Screenen des Trägers und die pränatale Diagnose wie folgt durchgeführt werden.

[0192] Die Population mit erhöhtem Risiko bezüglich der zystischen Fibrose ist kaukasisch. Beispielsweise würden alle kaukasischen Frauen und/oder Männer im gebärfähigen Alter einem Screenen unterzogen, um zu bestimmen, ob sie oder er ein Träger ist (ca. 5%ige Wahrscheinlichkeit bei jedem Individuum). Sind beide Trä-

ger, so besteht bei ihnen als Elternpaar das Risiko ein Kind mit zystischer Fibrose zu bekommen. Bei jedem Kind eines Elternpaars von erhöhtem Risiko liegt die Wahrscheinlichkeit, dass es von der zystischen Fibrose befallen wird, bei 25%. Das Verfahren zum Bestimmen des Trägerstatus unter Zuhilfenahme der hier offenbarten Sonden ist wie folgt.

[0193] Eine wichtige Anwendung der DNA-Sequenzinformationen bezüglich der normalen und mutanten ZF-Gene liegt auf dem Gebiet des genetischen Testens, der Trägerbestimmung und der pränatalen Diagnose. Individuen, die Träger von Mutationen im ZF-Gen (Krankheitsträger oder Patienten) sind, können bezüglich der DNA unter Zuhilfenahme verschiedener Techniken entfernt werden. Die für die Diagnose verwendete genomische DNA kann aus Körperzellen wie denjenigen, die in peripherem Blut, Urin, Speichel, der Gewebebiopsie, chirurgischen Proben und Sezionsmaterial vorliegen, gewonnen werden. Die DNA kann direkt für die Bestimmung einer spezifischen Sequenz verwendet oder enzymatisch *in vitro* mit Hilfe von PKR (Saiki et al. *Science* 230: 1350–1353, (1985), Saiki et al. *Nature* 324: 163–166 (1986) vor der Analyse amplifiziert werden. RNA oder ihre cDNA-Form kann für den gleichen Zweck ebenfalls verwendet werden. Kürzliche Besprechungen dieses Themas sind von Caskey (*Science* 236: 1223–8 (1989) und von Landegren et al. (*Science* 242: 229–237 (1989) vorgelegt worden.

[0194] Die Erfassung der spezifischen DNA-Sequenz kann durch Methoden wie die Hybridisierung unter Zuhilfenahme spezifischer Oligonucleotide (Wallace et al. *Cold Spring Harbour Symp. Quant. Biol.* 51: 257–261 (1986)), direktes DNA-Sequenzieren (Church and Gilbert, *Proc. Nat. Acad. Sci. USA* 81: 1991–1995 (1988)), die Verwendung von Restriktionsenzymen (Flavell et al. *Cell* 15: 25 (1978), Geever et al. *Proc. Nat. Acad. Sci. USA* 78: 5081 (1981)), die Diskriminierung auf der Basis der elektrophoretischen Mobilität in Gelen mit denaturiertem Reagens (Myers und Maniatis, *Cold Spring Harbour Sym. Quant. Biol.* 51: 275–284 (1986) die RNase-Protektion (Myers, R. M., Larin, J., und T. Maniatis *Science* 230: 1242 (1985)), die chemische Spaltung (Cotton et al. *Proc. Nat. Acad. Sci. USA* 85: 4397–4401 (1985) und das durch Ligase vermittelte Erfassungsverfahren (Landegren et al. *Science* 241: 1077 (1988)) erzielt werden.

[0195] Für normale oder mutante Sequenzen spezifische Oligonucleotide werden chemisch unter Zuhilfenahme im Handel erhältlicher Maschinen synthetisiert, radioaktiv mit Isotopen (wie ^{32}P) oder nichtradioaktiv (mit Markern wie Biotin (Ward and Langer et al. *Proc. Nat. Acad. Sci. USA* 78: 6633–6657 (1981)) markiert und mit einzelnen DNA-Proben hybridisiert, die auf Membranen oder anderen festen Trägern durch Dot-blot oder Übertragung aus Gelen auf die Elektrophorese hin immobilisiert. Das Vorliegen oder die Abwesenheit dieser spezifischer Sequenzen werden durch Methoden wie die Autoradiographie oder fluorimetrische (Landegren et al. 1989, *supra*) oder colorimetrischen Reaktion (Gebeyehu et al. *Nucleic Acids Research* 15: 4513–4534 (1987)) sichtbar gemacht. Eine Ausführungsform dieser Oligonucleotid-Screenmethode ist bei der Erfassung der F508-Deletion, wie hier beschrieben, angewendet worden.

[0196] Unterschiede zwischen normalen und mutanten Sequenzen können durch die direkte DNA-Sequenziermethode von Church und Gilbert (*supra*) erfasst werden. Klonierte DNA-Segmente können als Sonden zum Erfassen spezifischer DNA-Segmente verwendet werden. Die Empfindlichkeit dieser Methode wird dann wesentlich verbessert, wenn sie mit der PKR kombiniert wird (Wrichnik et al, *Nucleic Acids Res.* 15: 529–542 (1987); Wong et al. *Nature* 330: 384–386 (1987); Stofflet et al. *Science* 239: 491–494 (1988)). Bei letzterem Verfahren wird ein Sequenzierprimer, der innerhalb der amplifizierten Sequenz liegt, mit dem Doppelstrang-PKR-Produkt oder einer einsträngigen, durch eine modifizierte PKR gebildeten Matrize verwendet. Die Bestimmung der Sequenz findet durch herkömmliche Verfahren mit radiomarkierten Nucleotiden oder durch automatische Sequenzierverfahren mit Fluoreszenzmarkern statt.

[0197] Änderungen der Sequenz können ab und zu zur Bildung zufälliger Restriktionsenzym-Erkennungsstellen führen, die durch Anwendung einer geeigneten Enzymdigestion, gefolgt von herkömmlicher Gelblothybridisierung (Southern, *J. Mol. Biol.* 98: 503 (1975)) erfasst werden. Die jeweilige Stelle tragende DNA-Fragmente (entweder normale oder mutante) werden durch Reduzierung ihrer Größe oder Erhöhung der entsprechenden Anzahl von Restriktionsfragmenten erfasst. Genomische DNA-Proben können auch durch PKR vor der Behandlung mit dem entsprechenden Restriktionsenzym amplifiziert werden; Fragmente verschiedener Größe werden dann unter UV-Licht in Gegenwart von Ethidiumbromid auf die Gelelektrophorese hin sichtbar gemacht.

[0198] Die genetische Prüfung auf der Basis von DNA-Sequenz-Unterschieden kann durch Erfassen der Änderung der elektrophoretischen Mobilität von DNA-Fragmenten in Gelen mit oder ohne Denaturierungsreagens erreicht werden. Kleine Sequenzdeletionen und -Insertionen können durch Hochauflösungsgelelektrophorese sichtbar gemacht werden. Beispielsweise lässt sich das PKR-Produkt mit der 3 bp-Deletion deutlich von der normalen Sequenz an einem 8%igen nichtdenaturierten Polyacrylamidgel unterscheiden. DNA-Fragmente verschiedener Sequenzzusammensetzungen können an denaturiertem Formamidgradientengel unterschieden werden, in dem die Mobilitäten verschiedener DNA-Fragmente in dem Gel an verschiedenen Positionen, je nach ihren spezifischen "Partialschmelz"-Temperaturen verzögert sind (Myers, *supra*). Außerdem können Sequenzänderungen, insbesondere geringe Deletionen, als Änderungen im Migrationsmuster von DNA Heteroduplexen durch der Elektrophorese mit nichtdenaturiertem Gel erfasst werden, wie sie bezüglich

der 3 bp-(F508) Mutation und in anderen Versuchssystemen erfasst worden sind (Nagamine et al. *Am J. Hum. Genet.* 45: 337–339 (1989)). Als Alternative könnte eine Methode zum Erfassen einer Mutation, die eine Einbasensubstitution oder andere geringe Veränderung umfasst, auf der Differentialprimerlänge in einer PCR basieren. Beispielsweise könnte ein nichtvarianter Primer zusätzlich zu einem für eine Mutation spezifischen Primer verwendet werden. Die PCR-Produkte der normalen und mutanten Gene können daraufhin durch Differential in Acrylamidgelen erfasst werden.

[0199] Sequenzänderungen an spezifischen Positionen können auch durch Nucleaseschutzassays, wie RNase- (Myers, *supra*) und S1-Schutz (Berk, A. J., und P. A. Sharpe *Proc. Nat. Acad. Sci. USA* 75: 1274 (1978)), die chemische Spaltmethode (Cotton, *supra*) oder das durch Ligase vermittelte Erfassungsverfahren (Landegren, *supra*) enthüllt werden.

[0200] Außer den herkömmlichen Gelelektrophorese- und Blothybridisierungsmethoden ist es auch möglich, DNA-Fragmente durch Methoden sichtbar zu machen, bei denen die einzelnen DNA-Proben nicht auf Membranen immobilisiert werden. Die Sonde und die Zielsequenzen können beide in Lösung vorliegen oder die Sondensequenz kann immobilisiert sein (Saiki et al. *Proc. Natl. Acad. Sci. USA* 86: 6230–6234 (1989)). Eine Reihe von Erfassungsmethoden wie die Autoradiographie unter Zuhilfenahme von Radioisotopen, die direkte Erfassung des radioaktiven Zerfalls (in Gegenwart oder Abwesenheit von Szintillant), die Spektrophotometrie unter Zuhilfenahme kolorigener Reaktionen und die Fluorimetrie unter Zuhilfenahme fluorogener Reaktionen kann zum Identifizieren spezifischer individueller Genotypen angewendet werden.

[0201] Da mehr als eine Mutation im ZF-Gen zu erwarten ist, ist ein Mehrfachsystem ein ideales Protokoll für das Screenen von ZF-Trägern und die Erfassung spezifischer Mutationen. Beispielsweise kann eine PCR mit mehrfachen spezifischen Oligonucleotidprimern und Hybridisierungssonden zum gleichzeitigen Identifizieren aller möglichen Mutationen verwendet werden (Chamberlain et al. *Nucleic Acids Research* 16: 1141–1155 (1988)). Das Verfahren kann immobilisierte Sequenz-spezifische Oligonucleotidsonden umfassen (Saiki et al. *supra*).

5.2 DIE ERFASSUNG DER HAUPTMUTATION

[0202] Diese Erfassungsmethoden können bei der pränatalen Diagnose unter Verwendung von Zellen aus dem Fruchtwasser, der chorionischen Villi-Biopsie oder dem Sortieren fötaler Zellen aus dem mütterlichen Kreislauf angewendet werden. Die Untersuchung auf ZF-Träger in der Population kann als wesentlichen Teil in ein breit angelegtes genetisches Testprogramm auf weitverbreitete Krankheiten eingebaut werden.

[0203] Einer Ausführungsform der Erfindung gemäß kann der Teil des DNA-Segments, der der Informationen liefert bezüglich einer Mutation, wie beispielsweise der Mutation gemäß dieser Ausführungsform, d. h. des Teils, der die F508-Deletion direkt umgibt, durch Standard-PCR-Techniken amplifiziert werden [wie sie von Landegren, Ulf, Robert Kaiser, C. Thomas Caskey und Leroy Hood, *DNA Diganostics- Molecular Techniques und Automation*, in *Science* 242: 229–237 (1988), besprochen worden sind]. Es wird die Möglichkeit in Betracht gezogen, dass es sich bei dem Teil des DNA-Segments, der verwendet wird, eventuell um ein Einzel-DNA-Segment oder eine Mischung von verschiedenen DNA-Segmenten handeln kann. Es folgt nun eine genaue Beschreibung dieser Technik.

[0204] Es soll eine spezifische Region der genomischen DNA der Person oder des Fötus einem Screenen unterzogen werden. Eine derart spezifische Region wird durch die Oligonucleotidprimer C16B (5'GTTTTCTGGATTATGCCTGGGCAC3') und C16D (5'GTTGGCATGCTTTGATGACGCTTC3') definiert. Die spezifischen Regionen wurden durch die Polymerasekettenreaktion (PKR) amplifiziert. 200–400 ng genomischer DNA entweder aus kultivierten Lymphoblasten oder peripheren Blutproben aus an ZF leidenden Individuen und ihren Eltern wurden bei jeder PCR mit den oben angegebenen Oligonucleotidprimern verwendet. Die Oligonucleotide wurden mit Oligonucleotid-Reinigungskartuschen[®] (Applied Biosystems) oder NENSORB[®] PREP-Säulen (Dupont) mit Hilfe der von den Lieferanten empfohlenen Verfahren gereinigt. Die Primer wurden 45 Sekunden bei 62°C reassoziert, 120 Sekunden bei 72°C (mit 2 Einheiten Taq-DNA-Polymerase) extendiert und 60 Sekunden bei 94°C denaturiert, und zwar über 28 Zyklen mit einem Endzyklus von 7 Minuten für die Extension in einem automatischen Perkin-Elme/Cetus-Thermozykler mit einem schrittweisen Zyklusprogramm (Übergangseinstellung auf 1,5 Minuten). Anteile der PCR-Produkte wurden durch Elektrophorese auf 1,4% Agarosegel getrennt, auf Zetabind[®] (Biorad)-Membranen den Standardverfahren gemäß übertragen. Die beiden Oligonucleotidsonden aus **Fig. 15** (jeweils 10 ng) wurden getrennt mit 10 Einheiten T4-Polynucleotidkinase (Pharmacia) in 10 µl Reaktionsmedium, das 50 ml Tris-HCl (pH-Wert 7,6), 10 mM MgCl₂, 0,5 mM Dithiothreitol, 10 mM Spermidin, 1 mM EDTA und 30–40 µCi γ³²P] = ATP enthielt, für 20–30 min. bei 37°C markiert. Die freibleibenden Radionucleotide wurden vor der Verwendung mit einer Sephadex G-25-Säule entfernt. Die Hybridisierungsbedingungen waren die gleichen wie schon beschrieben (J. M. Rommens et al. *Am. J. Hum. Genet.* 43, 645 (1988)), mit Ausnahme der Tatsache, dass die Temperatur 37°C betrug. Die Membranen wurden zweimal bei Raumtemperatur mit 5 × SSC und zweimal bei 39°C mit 2 × SSC (1 × SSC = 150 mM NaCl und 13 mM Na-Citrat) gewaschen. Über Nacht wurde eine Autoradiographie bei Raumtemperatur durchgeführt. Die Auto-

radiographen zeigen die Hybridisierungsergebnisse genomischer DNA mit den beiden spezifischen Oligonucleotidsonden, wie in **Fig. 15** angegeben. Die Sonde C erfasst die normale DNA-Sequenz und die Sonde F erfasst die mutante Sequenz. Eine Probe genomischer DNA von jedem Familienmitglied wurde durch die Polymerasekettenreaktion amplifiziert und die Produkte durch Elektrophorese auf einem 1,4%igen Agarosegel getrennt und daraufhin Standardverfahren entsprechend auf eine Zetabind-(Biorad) Membrane übertragen. Wasserblindproben und Plasmid-DNA, T16 und C16, die der normalen Sequenz (N) bzw. der F508-Deletion (ZF) entsprachen, wurden als Kontrollen eingeschlossen.

[0205] Die 3 bp-Deletion wurde auch durch Polyacrylamidgelelektrophorese enthüllt. Wenn die durch die oben erwähnten C16B- und C16D-Primer hervorgerufenen PKR auf ein 8%iges Polyacrylamidgel angewendet und 2 Stunden bei 20 V/cm in einem 90 mM Tris-Boratpuffer (pH-Wert 8,3) eine Elektrophorese durchgeführt wurden, so waren DNA-Fragmente anderer Mobilität bei Individuen ohne die 3 bp-Deletion klar erfassbar, die für die Deletion heterozygot oder homozygot sind. Außerdem wurde eine extra DNA-Bande, bei der es sich um den Heteroduplex zwischen normalen und mutanten DNA-Strängen handelt dürfte, bei Heterozygoten bemerkt. Eine ähnliche Änderung der Gelmobilität bei während der PKR gebildeten Heteroduplexen ist auch für experimentelle Systeme berichtet worden, bei denen kleine Deletionen eine Rolle spielen (Nagamine et al supra). Diese Mobilitätsveränderungen können als Basis für die Tests für das nichtradioaktive genetische Screening verwendet werden.

5.3 ZF-SCREENINGPROGRAMM

[0206] Es ist klar, dass nur 70% der Träger mit Hilfe der spezifischen F508-Sonden dieser spezifischen Ausführungsform der Erfindung erfasst werden können. Erweist sich deshalb ein untersuchtes Individuum bei Anwendung der F508-Sonden nicht als Träger, so kann dessen Trägerstatus nicht ausgeschlossen werden, denn es kann Träger anderer Mutationen, wie oben bemerkt, sein. Sind jedoch sowohl das Individuum als auch der Gatte/die Gattin des untersuchten Individuums Träger der F508-Mutation, so kann mit Sicherheit angegeben werden, dass es sich dabei um ein Paar mit erhöhtem Risiko handelt. Die Sequenz des Gens, wie sie hier offenbart wird, ist eine wesentliche Voraussetzung für die Bestimmung der anderen Mutationen.

[0207] Die pränatale Diagnose ist eine logische Erweiterung des Screenens von Trägern. Ein Paar kann als ein solches mit erhöhtem Risiko, ein Kind mit zystischer Fibrose zu bekommen, durch eine von zwei Möglichkeiten identifiziert werden: falls sie schon ein an zystischer Fibrose leidendes Kind haben, so sind sie beide definitionsgemäß notwendigerweise Träger der Krankheit und bei jedem darauffolgenden Kind liegt die Wahrscheinlichkeit, dass es von zystischer Fibrose befallen ist, bei 25%. Ein Hauptvorteil der vorliegenden Erfindung liegt darin, dass sie die Notwendigkeit der Analyse des Stammbaums erübrigt, während ein Genmutationsscreeningprogramm, wie es oben skizziert ist, oder eine andere ähnliche Methode erfindungsgemäß verwendet werden kann, um eine genetische Mutation zu identifizieren, die zu einem Protein mit geänderter Funktion führt. Das hängt nicht von einer vorherigen Feststellung der Familie durch ein betroffenes Kind ab. Fötale DNA-Proben können beispielsweise, wie oben erwähnt, aus den Zellen des Fruchtwasser und chorionischen Villi-Proben erhalten werden. Eine Amplifikation kann daraufhin durch Standard-PKR-Techniken an dieser Matrizen-DNA durchgeführt werden.

[0208] Erweisen sich beide Eltern als Träger mit der F508-Deletion, so wäre die Interpretation der Ergebnisse wie folgt. Findet eine Hybridisierung der fötalen DNA mit der normalen Sonde (keine Deletion, wie in **Fig. 15** gezeigt) statt, so wird der Fötus nicht durch zystische Fibrose angegriffen, obwohl er ein ZF-Träger sein kann (50%ige Wahrscheinlichkeit bei jedem Fötus eines Paares mit erhöhtem Risiko). Hybridisiert die fötale DNA nur mit der F508-Deletionssonde und nicht der normalen Sonde (wie in **Fig. 15** gezeigt), so wird der Fötus durch zystische Fibrose angegriffen.

[0209] Es ist klar, dass bezüglich dieser und anderer Mutationen im ZF-Gen, eine Reihe verschiedener spezifischer Verfahren angewendet werden können, um eine vollständige Diagnose aller potentiellen ZF-Träger oder -patienten zu liefern. Eine vollständige Beschreibung dieser Verfahren erfolgt später.

[0210] Die Erfindung bietet daher eine Methode und ein Kit für das Feststellen, ob eine Versuchsperson ein ZF-Träger oder ZF-Patient ist. Zusammenfassend umfasst die Methode für das Screening folgende Schritte: Zurverfügungstellen einer biologischen Probe der Versuchsperson, die dem Screenen unterzogen werden soll und Durchführen einer Untersuchung für die Bestimmung, in der biologischen Probe, des Vorliegens von mindestens einem Mitglied der Gruppe, die aus Folgendem besteht: dem normalen ZF-Gen, normalen ZF-Genprodukten, einem mutanten ZF-Gen, mutanten ZF-Genprodukten und Mischungen derselben.

[0211] Die Methode kann des Weiteren dadurch gekennzeichnet sein, dass mindestens eine weitere Nucleotidsonde eingeschlossen ist, bei der es sich um ein anderes DNA-Sequenzfragment von beispielsweise der DNA von **Fig. 1** oder ein anderes DNA-Sequenzfragment des menschlichen Chromosoms 7 handelt, und die zu beiden Zeiten der DNA-Sequenz in **Fig. 1** positioniert ist.

[0212] Ein Kit nach einer Ausführungsform der Erfindung gemäß, der für die Verwendung bei Screeningverfahren und Untersuchungen auf das Vorliegen des ZF-Gens durch Immunassay geeignet, ist, umfasst:

- (a) einen Antikörper der spezifisch an ein Genprodukt des ZF-Gens bindet;
- (b) ein Reagensmittel für das Feststellen der Bindung des Antikörpers an das Genprodukt; und
- (c) wobei der Antikörper und das Reagensmittel in Mengen vorliegen, die zur Durchführung des Immunoassays effektiv sind.

[0213] Der Kit für das Untersuchen auf das Vorliegen des ZF-Gens kann auch durch Hybridisierungstechniken bereitgestellt werden. Der Kit umfasst:

- (a) eine Oligonucleotidsonde, die spezifische an das ZF-Gen bindet;
- (b) ein Reagensmittel für das Feststellen der Hybridisierung der Oligonucleotidsonde an dem ZF-Gen und;
- (c) wobei die Sonde und das Reagensmittel in Mengen vorliegen, die zur Durchführung des Hybridisierungssassays effektiv sind.

5.4 ANTIKÖRPER ZUM ERFASSEN DES ZFTR

[0214] Wie oben erwähnt, werden Antikörper zu Epitopen innerhalb des ZFTR-Proteins gezüchtet, um umfangreiche Informationen bezüglich der charakteristischen Eigenschaften des Proteins und andere wertvolle Informationen zu bieten, einschließlich Folgender:

1. Um ein Sichtbarmachen des Proteins in Zellen und Geweben zu ermöglichen, in denen es durch Immunblotting ("Western Blots") gefolgt von Polyacrylamidgelelektrophorese exprimiert wird. Das erlaubt eine schätzungsweise Bestimmung der Molekülgröße des reifen Proteins, einschließlich des Beitrags aus den Zellen von auf die Translation hin angelagerten Anteilen, einschließlich Oligosaccharidketten und Phosphatgruppen, zum Beispiel. Immunzytochemische Techniken, einschließlich der Immunfluoreszenz und Immunelektronmikroskopie, können verwendet werden, um die subzelluläre Position des Proteins in Zellmembranen zu bestimmen. Die Antikörper können auch dazu verwendet werden, eine andere Technik zum Erfassen irgendwelcher der anderen ZF-Mutationen zu bieten, die zur Synthese eines Proteins geänderter Größe führen.
2. Antikörper zu bestimmten Domänen des Proteins können zum Bestimmen der topologischen Anordnung des Proteins in der Zellmembran verwendet werden. Dies bietet Informationen über Segmente des Proteins, die für extern zugesetzte Modulierer für Zwecke der Arzneimitteltherapie zugänglich sind.
3. Der Zusammenhang zwischen Struktur und Funktion von Anteilen des Proteins kann mit Hilfe spezifischer Antikörper untersucht werden. Beispielsweise ist es möglich, Antikörper in die Zellen einzuführen, die jede der beladenen zytoplasmischen Schleifen, die die Transmembransequenzen verbinden, sowie Anteile der Nucleotidbindungsfalten und der R-Domäne erkennen. Der Einfluss dieser Antikörper auf funktionelle Parameter des Proteins bietet Einsicht in die Zellregulierungsmechanismen und bietet potentielle Vorschläge bezüglich der Mittel zum Modulieren der Aktivität des defekten Proteins bei einem ZF-Patienten.
4. Antikörper mit der entsprechenden Avidität ermöglichen auch die Immunpräzipitation und Immunaffinitätsreinigung des Proteins. Die Immunpräzipitation erleichtert die Charakterisierung der Synthese und die Posttranslationsmodifikation, einschließlich der ATP-Bindung und Phosphorylierung. Die Reinigung ist für die Untersuchung der Proteinstruktur und für die Rekonstitution ihrer Funktion sowie die Therapie auf Proteinbasis erforderlich.

[0215] Um die Antikörper zuzubereiten, sind Fusionsproteine, die definierte Anteile von ZFTR-Polypeptiden enthalten, in Bakterien durch die Expression von entsprechenden DNA-Sequenzen in einem geeigneten Kloniervehikel synthetisiert worden, während kleinere Peptide chemisch wie in Tabelle 8 beschrieben synthetisiert worden sind. Die Fusionsproteine wurden beispielsweise durch Affinitätschromatographie auf Glutathion-Agarose gereinigt und die Peptide wurden an ein Trägerprotein (Hämocyanin) angekoppelt, mit Freund-Adjuvans gemischt und in Kaninchen eingespritzt. Auf Auffrischungsimpfungen in zweiwöchentlichen Abständen hin wurden die Kaninchen geblutet und Sera isoliert. Die gefärbten Fusionsproteine sind in den **Fig. 19a** gezeigt. Die Bahn 1 stellt das nichtinduzierte Kontrollplasmid, die Bahn 2 das IPTG-induzierte Kontrollplasmid, das ausschließlich Glutathion-S-Transferase (GST) exprimiert, die Bahn 3 stellt die affinitätsgereinigte GST-Bande bei 27 Kilodalton (kD) dar, die Bahn 4 ist nicht induziert, die Bahn 5 ist induziert und die Bahn 6 stellt das gereinigte Fusionsprotein #1 aus Tabelle 8 dar. In **Fig. 19b** ist die Gelelektrophorese von Lysaten von Bakterien dargestellt, die mit pGEX-Plasmiden transformiert worden sind, die die Fusionsproteine #5 aus Tabelle 8 für die Bahnen 1 und 2 und die Fusionsproteine #2 aus Tabelle 8 für die Bahnen 3 und 4 enthalten. Die Bahn 1 aus **Fig. 19b** betrifft das nichtinduzierte Plasmid, während die Bahn 2 das induzierte Plasmid, zum Exprimieren des Fusionsproteins #5 betrifft. Die Bahn 3 der **Fig. 19b** steht für das nichtinduzierte Plasmid, während die Bahn 4 für das induzierte Plasmid zum Exprimieren des Fusionsproteins #2 steht. Immunblots des Fusionsproteins #1 das mit Antisera sondiert worden ist, die aus den zweiten Blutungen von zwei verschiedenen Kaninchen erhalten worden sind, sind in **Fig. 20** gezeigt. Die Färbung erfolgte mit einem durch Alkaliphosphatase konjugierten zweiten Antikörper [Blake et al, Anal. Biochem. 136: 175, (1984)]. Beide dieser Immunsera färben das 32

kD-Fusionsprotein, während die Vorimmunsera es nicht tun. **Fig. 21** zeigt die Reaktivität einer dieser Immunsera mit einer Bande einer Größe von ca. 200 kD in Membranen, die aus 784-Kolonkrebiszellen isoliert worden sind, die das ZFTR-Transkript in hohem Maße exprimieren. Diese Bande liegt im Größenbereich, der für das ZFTR-Protein zu erwarten ist, das vor Posttranslationsmodifikationen ein vorausgesagtes Molekulargewicht von 169 kD aufweist.

[0216] Sera aus Kaninchen, die mit dem LKH-Konjugat des Peptids #2 immunisiert worden sind, wurden einem Screenen sowohl auf reines Peptid als auch KLH, wie in **Fig. 22** gezeigt, unterzogen. In dieser Figur steht H für Hämocyanin, P 1 für Peptid #1, P2 für Peptid #2. Die Mengen von Protein oder Peptid, in ng angegeben, sind aufgezeigt. Dieses Antiserum erfasst Peptid in Mengen bis hinunter zu 1 ng und reagiert in keiner Weise mit dem Kontrollpeptid #1.

[0217] So ist es möglich, polyklonale Antikörper zu züchten, die sowohl für Teile des ZFTR-Proteins enthaltende Fusionsproteine als auch Peptide spezifisch sind, die kurzen Segmenten ihrer Sequenz entsprechen. Desgleichen können Mäuse mit KLH-Konjugaten der Peptide 1, 2 und 7 aus Tabelle 8 injiziert werden, um die Bildung monoklonaler Antikörper zu diesen Segmenten des ZFTR-Proteins einzuleiten. Monoklonale Antikörper können desgleichen zu anderen Domänen des ZFTR-Proteins gezüchtet werden.

[0218] Was die Bildung polyklonaler Antikörper anbetrifft, so sind Immunogene für das Züchten monoklonaler Antikörper (mAbs) zum ZFTR-Protein bakterielle Fusionsproteine [Smith et al, *Gene* 67: 31 (1988)], die Anteile des ZFTR-Polypeptids enthalten, oder synthetische Peptide, die kurzen (12 bis 25 Aminosäurenlängen) Segmenten der Sequenz entsprechen. Die wesentliche Methodologie ist diejenige von Kohler und Milstein [*Nature* 256: 495 (1975)].

[0219] Balb/c-Mäuse werden durch intraperitoneales Einspritzen von 500 µg reinem Fusionsprotein oder synthetischem Peptid in unvollständigem Freund-Adjuvans immunisiert. Eine zweite Einspritzung erfolgt nach 14 Tagen, eine dritte nach 21 Tagen und eine vierte nach 28 Tagen. Einzelne so immunisierte Tiere werden eine, zwei bzw. vier Wochen nach der endgültigen Einspritzung getötet. Die Milz wird entfernt, ihre Zellen dissoziiert, eingesammelt und mit Sp2/O-Ag14-Myelomzellen Geffer et al, *Somatic Cell Genetics* 3: 231 (1977) entsprechend fusioniert. Die Fusionsmischung wird im Kulturmedium verteilt, das für die Propagierung von fusionierten Zellen selektiv ist, die gezüchtet werden, bis sie ca. 25% konfluent sind. Zu diesem Zeitpunkt werden Züchtungsüberstände auf die Anwesenheit von Antikörpern hin getestet, die mit einem spezifischen ZFTR-Antigen reagieren. Ein zweiter mit Alkaliphosphatase markierter Antikörper wird daraufhin für die Erfassung von Positiven verwendet. Zellen aus positiven Züchtungsvertiefungen werden dann in Kulturmedium expandiert, ihre überstehenden Flüssigkeiten für weitere Tests eingesammelt und die Zellen im tiefgefrorenen Zustand in Cryoschutzmittel enthaltendem Medium gelagert. Um große Mengen mAb zu erhalten, werden Bildungszellen in das Bauchfell in einer Menge von 5×10^6 Zellen pro Tier eingespritzt und Ascites-Fluid erhalten. Die Reinigung erfolgt durch Chromatographie auf Protein G- oder Protein A-Agarose Ey et al, *Immunochemistry* 15: 429 (1977) entsprechend.

[0220] Die Reaktivität dieser mAbs mit dem ZFTR-Protein wird durch Polyacrylamidgel-Elektrophorese von Membranen bestätigt, die von Epithelzellen isoliert worden sind, in denen es exprimiert wird, sowie durch Immunblotting [Towbin et al, *Proc. Natl. Acad. Sci. USA* 76: 4350 (1979)].

[0221] Zusätzlich zur Verwendung monoklonaler Antikörper, die für jede der verschiedenen Domänen des ZFTR-Proteins spezifisch sind, zum Sondieren ihrer individuellen Funktionen, werden andere mAbs, die zwischen den normalen und den mutanten Formen des ZFTR-Proteins unterscheiden können, zum Erfassen des mutanten Proteins in Proben von Epithelzellen verwendet, die von Patienten erhalten worden sind, wie beispielsweise durch Nasenschleimhautbiopsie erhaltene "Pinselabstriche", [R. De-Lough und J. Rutland, *J. Clin. Pathol.* 42, 613 (1989)] oder Schweißdrüsen enthaltenden Hautbiopsieproben.

[0222] Antikörper, die dieser Unterscheidung fähig sind, werden durch differentielles Screenen von Hybridomen aus Gruppen von Mauspaaren erhalten, die mit einem Peptid, das das Phenylalanin an der Aminosäureposition 508 enthält (z. B. GTIKENIIFGVSY) oder einem Peptid, das mit Ausnahme der Abwesenheit von F508 identisch ist, immunisiert worden sind. mAbs, die in der Lage sind, die anderen mutanten Formen von in Patienten vorliegendem ZFTR-Protein zusätzlich oder anstatt der F508-Deletion zu erkennen, werden mit Hilfe ähnlicher Strategien zur Bildung monoklonaler Antikörper erhalten.

[0223] Antikörper zum normalen und zu ZF-Versionen des ZFTR-Proteins und Segmenten desselben werden bei der diagnostischen Immunzytochemischen und Immunfluoreszenz-Lichtmikroskopie und Immunelektronmikroskopie zum Aufzeigen der Gewebe-, Zell- und Subzellverteilung von ZFTR innerhalb der Organe von ZF-Patienten, Trägern und Nicht-ZF-Individuen verwendet.

[0224] Antikörper werden zum therapeutischen Modulieren durch Fördern der Aktivität des ZFTR-Proteins bei ZF-Patienten und in Zellen von ZF-Patienten verwendet. Mögliche Verfahrensweisen für eine derartige Modulation können unter Umständen die Stimulierung aufgrund von Vernetzung von ZFTR-Proteinmolekülen mit mehrwertigen Antikörpern in Analogie zum Stimulieren einiger Zelloberflächenmembranrezeptoren, wie dem Insulinrezeptor [O'Brien et al, *Euro. Mol. Biol. Organ. J.* 6: 4003 (1987)], dem Epidermiswachstums-Faktorezeptor [Schreiber et al, *J. Biol. Chem.* 258: 846 (1983)] und mit T-Zellrezeptoren assoziierten Molekülen wie

CD4 [Veillette et al Nature, 338: 357 (1989)] involvieren.

[0225] Antikörper werden zum Ausrichten der Abgabe therapeutischer Mittel an die Zellen, die das defekte ZFTR-Protein bei ZF exprimieren, verwendet. Zu diesem Zweck werden die Antikörper in ein Vehikel, wie beispielsweise ein Liposom, eingearbeitet [Matthay et al, Cancer Res. 46: 4904 (1986)], das das therapeutische Mittel, wie beispielsweise ein Arzneimittel oder das normale Gen, trägt.

TABELLE 8
ZUR BILDUNG VON ANTIKÖRPERN VERWENDETE ZFTR-FRAGMENTE

GST ^a -Fusionsproteine die ZFTR-Rückstände enthalten		ZFTR-Domäne von Fig. 13
1.	204-249	TM3, Ext. 2, TMA
2.	347-698	NFB-1, N-term., 1/2 R-Domäne
3.	710-757	Negativ geladener mittlerer Teil der R-Domäne
4.	758-796	Positiv geladenes Segment der R-Domäne
5.	1188-1480	C-term. zyto. Domäne mit NBF-2
KLH ^b -Konjugate, die ZFTR-Peptide enthalten		
1.	28-45	N-term. zytoplasmisch
2.	58-75	N-term. zytoplasmisch
3.	104-117	1. extrazellulär
4.	139-153	2. zytoplasmisch
5.	279-294	N-term. vom 3. zytoplasmisch
6.	500-512	NBF-1, um die F508-Deletion herum
7.	725-739	geladener Mittelteil der R-Domäne
8.	933-946	5. zytoplasmisch
9.	1066-1084	6. zytoplasmisch

^a Restriktionsfragmente, die für diese Fragmente kodieren, die an das 3'-Ende der Glutathion-S-Transferase (GsT) von Schistosoma Japonicum im pGEX-Plasmidexpressionsvektor, wie bei Smith et al, Gene 67: 31 (1988) identifiziert, ligiert sind.

^b Peptide, die durch ein N-terminales Cystein an das Trägerprotein-Schlüsselloch-Limpethämocyanin (KLH) Green et al, Cell 28: 477 (1982) entsprechend gekoppelt sind. TM bedeutet Transmembransequenzen.

5.5 DIE RFLP-ANALYSE

[0226] Diese Erfindung bietet eine Reihe von Vorteilen, die direkt aus der Entdeckung und der Charakterisierung des ZF-Gens abgeleitet sind und eine direkte praktische Anwendung besitzen. Die Aminosäuresequenz von ZFTR bietet Einsicht in die Struktur und Funktion des Proteins sowie des Molekularmechanismus, an denen ZFTR teilnimmt, und die bei der zystischen Fibrose defekt sind. Diese Information ermöglicht die Bildung weiterer Werkzeuge und Konzepte zur Forschung und Therapie bezüglich dieser Krankheit.

[0227] Die Erfassung der Träger, die DNA-Diagnose und Beratung der Familien sind einige der Applikationen dieser Erfindung. Bisher ist die genetische Untersuchung auf der Basis der DNA bezüglich der ZF vor allem Familien mit betroffenen Kindern und ihren engen Verwandten zur Verfügung gestanden. Das Wissen bezüglich der ZF-Mutationen auf dem Niveau der DNA-Sequenz erlaubt die Stichprobenuntersuchung von Individuen. Unserer Schätzung nach können 46% der ZF-Patienten ohne vorherige Familienkrankengeschichte durch DNA-Analyse akkurat diagnostiziert werden und 68% der ZF-Träger in der Bevölkerung können über die

F508-Deletion identifiziert werden.

[0228] Angesichts der Tatsache, dass die Trägerhäufigkeit in der nordamerikanischen Bevölkerung ca. 1 unter 20 beträgt, ist es realisierbar, beispielsweise alle Männer bzw. Frauen von gebärfähigem Alter bezüglich ihres Trägerstatus zu screenen. Durch die Erfassung der Träger unter Zuhilfenahme von Sonden, die für die F508-Deletion spezifisch sind, wird bis zu 70% der Träger identifiziert. Die übrigen Träger werden durch eine Reihe von Sonden, die für die verschiedenen, oben identifizierten Haplotypgruppen spezifisch sind, erfasst.

[0229] Da die F508-Deletion ca. 70% aller ZF-Mutationen darstellt, kann die RFLP-Analyse zusätzlich zum Testen auf direkte Deletion bei Familienmitgliedern oder nahen Verwandten von ZF-Patienten angewendet werden. Von ca. 55% der ZF-Eltern, die keine Träger der F508-Mutation sind, ist zu erwarten, dass sie auf der Basis der rückblickenden Analyse unserer ZF-Verkoppelungsfamilien bezüglich des DNA-Markers JG2E1 (KM19) informativ sind [Kerem et al, Am. J. Hum. Genet. 44: 827–834 (1989); Estivill et al, Genomics 1: 257 (1987)]; weitere 39% wären informativ, wenn E6 (Taq I) [Kerem et al supra] und J3.11 (Map I) [Wainwright et al Nature (1995)] ebenfalls getestet würden; praktisch alle Eltern wären informativ, wenn H2.3 (XV2C-Taq I) [Kerem et al, supra; Estivill et al, Nature (1987)], E2.6 (E.9) (Msp I) [auf Anfrage erhältliche Sonde], E4.1 (Mp6d.9) (Msp I) [auf Anfrage erhältliche Sonde; Estivill et al, Am. J. Hum. Genet. (1989)], J44 (E3.1) (Xba I) [auf Anfrage erhältliche Sonde] und metD (Ban I) [Spence et al. Am. J. Hum. Genet (1986), [ATCC #40219] in den Test eingeschlossen würden.

[0230] Die Nützlichkeit dieser Sonden beruht auf der Tatsache, dass sie polymorphe Restriktionsstellen erkennen. So sind die Sonden typischerweise nicht durch ihre Sequenz über die spezifische polymorphe Stelle definiert, sondern können aufgrund des Wissens bezüglich der flankierenden Sequenzen angewendet werden unter Berücksichtigung der Bildung der Polymerasekettenreaktion (PKR) der fraglichen Region, wie es einem Fachmann bekannt wäre.

[0231] Beispielsweise wird die Sonde E2.6 (Msp I) vollständig durch zwei flankierende Oligomere definiert: 5'GTGATCCAGTTTGCTCTCCA3' und 'GGAATCACTCTTCCTGATAT3'. Die Verwendung dieser durch E2.6 PRC gebildeten Sonde zum Erfassen eines Msp I-Polymerphismus erfasst zwei verschiedenen Allele: entweder ein Fragment von 850 by oder ein Fragment von 490 by und 360 bp, je nachdem, ob die Msp I-Stelle vorliegt oder nicht. Desgleichen wird die Sonde J44 (E3.1) (Xba I) vollständig durch zwei flankierende Oligomere definiert:

5'CAATGTGATTGGTGAACAACTA3' und 5'CTTCTCCTCCTAGACACCTGC AT3'. Die Verwendung dieser durch J44 (E3.1) PKR gebildeten Sonde zum Erfassen eines Xba I-Polymerphismus führt zum Erfassen zwei verschiedener Allele: entweder eines Fragments von 860 by oder eines Fragments von 610 by und eines Fragments von 250 pb, je nachdem, ob die Xba I-Stelle vorliegt oder nicht.

[0232] Die verkoppelten RFLPs können auch bei der Berechnung des Risikos bei Einzelpersonen, die nicht Träger der F508-Deletion sind, verwendet werden. Ein allgemeines Risikoabschätzungsverfahren ist von Baudet et al Am. J. Hum. Genet. 44: 319–326 besprochen worden.

[0233] Zur pränatalen Diagnose kann eine Mikrovillus-Darmenzymanalyse (Brock, Lancet 2: 941 (1983)) zur Erhöhung der Verlässlichkeit der Diagnose in den Fällen, in denen eine DNA-Diagnose nicht schlüssig ist, durchgeführt werden.

[0234] Die DNA-Diagnose wird gegenwärtig zum Beurteilen angewendet, ob ein Fötus mit zystischer Fibrose geboren werden wird, in der Vergangenheit ist dies jedoch nur dann ausgeführt worden, wenn ein bestimmtes Elternpaar schon ein an zystischer Fibrose leidendes Kind bekommen hat, durch das es notwendigerweise als Träger identifiziert worden ist. In Kombination mit der Trägererfassung, wie sie oben skizziert ist, wird jedoch eine DNA-Diagnose für alle Schwangerschaften bei Trägerpaaren möglich. Wenn die Eltern schon ein an zystischer Fibrose leidendes Kind bekommen haben, kann eine extendierte Haplotypanalyse am Fötus ausgeführt und so der Prozentsatz an falschen positiven oder falschen negativen Ergebnissen stark reduziert werden. Haben die Eltern noch kein davon befallenes Kind bekommen und die DNA-Diagnose am Fötus wird auf der Basis der Trägererfassung durchgeführt, so kann die Haplotypanalyse trotzdem durchgeführt werden.

[0235] Obwohl schon seit vielen Jahren der Gedanke vorherrscht, dass bei der zystischen Fibrosekrankheit ein hohes Niveau klinischer Heterogenität vorliegt, stellt es sich nur heraus, dass zwei allgemeine Kategorien, die so genannte Pankreassuffizienz (ZF-PS) und die Pankreasinsuffizienz (ZF-PI) bestehen. Sind die mit diesem Krankheitskategorien verbundenen Mutationen gut charakterisiert, so kann man eine spezifische Mutation mit einem klinischen Phenotyp der Krankheit in Verbindung bringen. Dadurch sind Änderungen der Behandlung jedes Patienten möglich. Durch die Natur der Mutation lässt sich auf diese Weise die Prognose des Patienten bis zu einem gewissen Grad voraussagen und eine spezifische Behandlung indizieren.

6.0 MOLEKULARBIOLOGY DER ZYSTISCHEN FIBROSE

[0236] Das Postulat, dass das ZFTR eventuell die Aktivität von Innenkanälen, insbesondere des nach außen rektifizierenden C1-Kanals, der als funktioneller Defekt bei der ZF eine Rolle spielt, regulieren kann, kann durch die Einspritzung und Translation von in vitro transkribierter ZFTR-mRNA voller Länge in Xenopus-Oozyten ge-

testet werden. Die dadurch hervorgerufenen Änderungen in den Innenströmen über die Oozytenmembran hinweg können gemessen werden, da das Potential auf einen festgelegten Wert eingestellt wird. Das ZFTR kann endogene Oozytenkanäle regulieren oder es kann eventuell notwendig sein, Epithelzellen-RNA ebenfalls einzuführen, um die Translation von Kanalproteinen zu lenken. Die Anwendung des mRNA-Kodierens für normales und für mutantes ZFTR, wie es durch diese Erfindung geboten wird, macht diese Versuche möglich.

[0237] Andere Expressionsarten im heterologen Zellsystem ermöglichen auch die genaue Analyse von Struktur-Funktion-Zusammenhängen. Die vollständige in den Plasmidexpressionsvektor ligierte ZFTR-DNA-Sequenz wird zum Transfizieren von Zellen derart verwendet, dass ihr Einfluss auf den Innentransport beurteilt werden kann. Plasmidexpressionsvektoren, die einen Teil der normalen ZFTR-Sequenz zusammen mit Anteilen modifizierter Sequenz an ausgewählten Stellen enthalten, können bei Mutageneseversuchert *in vitro* verwendet werden, die mit dem Ziel durchgeführt werden, diejenigen Anteile des ZFTR-Proteins zu identifizieren, die für die Regulierfunktion ausschlaggebend sind.

6.1 DIE EXPRESSION DER DNA-SEQUENZ

[0238] Die DNA-Sequenz kann bei Studien zur Verbesserung des Verständnisses der Expression des Gens und seiner Produkte und zur Erzielung der Produktion großer Mengen des Proteins für die funktionelle Analyse, die Herstellung von Antikörpern und die Patiententherapie manipuliert werden. Die Änderungen in der Sequenz können das Expressionsbild mit Bezug auf die relativen Mengen, die Gewebespezifität und die funktionellen Eigenschaften ändern, oder auch nicht. Die Teil-cDNA-Sequenz oder die cDNA-Sequenz voller Länge, die für das Protein der Untersuchungsperson kodieren und unmodifiziert oder modifiziert sind, können an Bakterienexpressionsvektoren wie pRIT-(Nilsson et al. ENBO J. 4: 1075–1080 (1985)), pGEX-(Smith und Johnson, Gene 67: 31–40 (1988)) oder pATH-(Spindler et al. J. Virol. 49: 132–141 (1984)) Plasmide ligiert werden, die in *E. coli*-Zellen eingeführt werden können für die Bildung der entsprechenden Proteine, die den oben besprochenen Proteinreinigungsverfahren entsprechend isoliert werden können. Die DNA-Sequenz kann auch von ihrem bestehenden Kontext auf andere Kloniervehikel, wie beispielsweise andere Plasmide, Bakteriophagen, Cosmide, tierische Viren, künstliche Hefechromosomen (YAC) (Burke et al. Science 236: 806–812 (1987)), somatische Zellen und andere einfache oder komplexe Organismen wie Bakterien, Pilze (Timberlake und Marshall, Science 244: 1313–1317 (1989), Invertebrate, Pflanzen (Gasser und Fraley, Science 244: 1293 (1989)) und Schweine (Pwsel et al. Science 244: 1281–1288 (1989)) übertragen werden.

[0239] Für die Expression in Säugerzellen kann die cDNA-Sequenz an heterologe Promotoren wie beispielsweise das Simian-Virus (SV) 40, einen Promotor im pSV2-Vektor [Mulligan und Berg, Proc. Natl. Acad. Sci. USA, 78: 2072–2076 (1981)] ligiert und in Zellen wie Affen-COS-1-Zellen [Guzman, Cell. 23: 175–182 (1981)] zur Erzielung einer vorübergehenden oder Langzeit-Expression eingeführt werden. Die beständige Integration des chimären Genkonstrukts kann in Säugerzellen durch biochemische Selektion, wie beispielsweise Neomycin [Southern und Berg, J. Mol. Appln. Genet. 1: 327–341 (1982)] und Mycophenolsäure [Mulligan und Berg, supra] aufrechterhalten werden.

[0240] DNA-Sequenzen können durch Standardverfahren wie die Restriktionsenzymdigestion, Lückenfüllung mit DNA-Polymerase, Deletion durch ExoNuclease, Extension durch terminelle Desoxynucleotidtransferase, Ligierung von synthetischen oder klonierten DNA-Sequenzen, positionsgeregelte Sequenzänderung über Einstrang-Bakteriophagenzwischenprodukt oder unter Verwendung spezifischer Die cDNA-Sequenz (oder von ihr abgeleitet Teile) oder ein Minigen (eine Oligonukleotide in Kombinationen mit der PKR manipuliert werden.

[0241] cDNA mit einem Intron und ihrem eigenen Promotor) wird in eukaryotische Expressionsvektoren durch herkömmliche Verfahren eingeführt. Diese Vektoren sind so konzipiert, dass sie die Transkription der cDNA in eukaryotische Zellen durch Bereitstellen von Reguliersequenzen erlauben, die die Transkription der cDNA einleiten und verbessern und deren richtige Spleißung und Polyadenylierung sicherstellen. Vektoren, die den Promotor und Enhancerbereiche des Simianvirus (SV) 40 oder lange Terminalwiederholungen (LTVn des Rous-Sarkom-Virus und ein Polyadenylations- und Spleißsignal von SV 40 enthalten, stehen ohne weiteres zur Verfügung [Mulligan et al. Proc. Natl. Acad. Sci. USA 78: 1078–2076 (1981), Gorman et al. Proc. Natl. Acad. Sci. USA 79: 6777–6781 (1982)].

[0242] Als Alternative kann der endogene ZFTR-Promotor verwendet werden. Das Expressionsniveau der cDNA kann bei diesem Typ Vektor entweder durch Verwendung von Promotoren verschiedener Aktivitäten (beispielsweise kann das Baculovirus pAC373 cDNAs in hohem Niveaus in *S. frugiperda*-Zellen exprimieren [M. D. Sommers und G. E. Smith in Genetically Altered Viruses and the Environment (B. Fields, et al. Verfasser), Band 22, Seite 319–328, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, 1985] oder durch Verwendung von Vektoren, die der Modulation zugängliche Promotoren, beispielsweise den auf Glukocorticoid ansprechenden Promotor aus dem Maussäugertumorstumorvirus [Lee et al. Nature 294: 228 (1982)] enthalten, manipuliert werden. Die Expression der cDNA kann in den Empfängerzellen für 24 bis 72 Stunden nach Einführung (vorübergehende Expression) überwacht werden.

[0243] Außerdem enthalten einige Vektoren selektierbare Marker, [wie gut [Mulligan und Berg et al., supra]

oder neobakterielle Gene (Southern und Berg J. Mol. Appl. Genet 1: 327–341 (1982)), die das Isolieren, durch chemische Selektion, von Zellen gestatten die eine stabile Langzeitexpression der Vektoren (und deshalb der cDNA) in der Empfängerzelle aufweisen. Die Vektoren können in den Zellen als episomale, sich frei wiederholende Einheiten aufrechterhalten werden, und zwar durch Anwendung von Regulierelementen von Viren wie Papillom [Sarver et al. Mol. Cell Biol. 1: 486 (1981) oder Epstein-Barr (Sugden et al. Mol. Cell Biol. 5: 410 (1985)).

[0244] Als Alternative können auch Zelllinien hergestellt werden, bei denen der Vektor in die genomische DNA integriert ist. Beide dieser Typen von Zelllinien bilden das Genprodukt auf kontinuierlicher Basis. Es ist auch möglich, Zelllinien herzustellen, die die Anzahl von Kopien des Vektors (und daher auch der cDNA) zur Bildung von Zelllinien, die Genprodukt in hohem Niveau bilden können [Alt et al. J. Biol. Chem. 253: 1357 (1978)] amplifiziert haben.

[0245] Die Übertragung von DNA in eukaryotische, insbesondere menschliche oder andere Säugerzellen, ist inzwischen eine herkömmliche Technik geworden. Die Vektoren werden als reine DNA in die Empfängerzellen (Transfektion) beispielsweise durch Ausfällen mit Calciumphosphat [(Graham und vander Eb. Virology 52: 466 (1973) oder Strontiumphosphat [Brash et al. Mol. Cell Biol. 7: 2013 (1987)], Elektroporation [Neumann et al. EMBO J 1: 841 (1982)], Lipofektion [(Felgner et al. Proc. Natl. Acad. Sci USA 84: 7413 (1987)], DEAE-Dextran [McCuthan et al J. Natl. Cancer Inst. 41: 51 1968)], Mikroinjektion [Mueller et al Cell 15: 579 (1978)], Protoplastfusion [Schafner, Proc Natl. Aca. Sci USA 72: 2163] oder Granulatpistole [Klein et al, Nature 327: 70 (1987)] eingeführt. Als Alternative kann die cDNA durch Infektion mit Virusvektoren eingeführt werden. Es sind Systeme entwickelt worden, bei denen beispielsweise Retroviren [Bernstein et al. Genetic Engineering 7: 235, (1985)], Adenoviren [Ahmad et al. J. Virol 57: 267 (1986)] oder Herpesviren [Spaete et al Cell 30: 295 (1982)] verwendet werden.

[0246] Diese eukaryotischen Expressionssysteme können für viele Studien des ZF-Gens und des ZFTR-Produkts verwendet werden. Diese umfassen beispielsweise: (1) Feststellen, dass das Gen richtig exprimiert ist und dass alle für die volle biologische Aktivität notwendigen Posttranslationsmodifikationen richtig abgeschlossen sind, (2) Identifizieren von Regulierelementen, die in der 5'-Region des ZF-Gens positioniert sind und ihrer Rolle bei der Gewebe- oder vorübergehenden Regulierung der Expression des ZF-Gens, (3) Bildung großer Mengen des normalen Proteins für das Isolieren und Reinigen, (4) Verwendung von Zellen, die das ZFTR-Protein exprimieren, als Assaysystem für Antikörper, die gegen das ZFTR-Protein gebildet worden sind, oder als Assaysystem zum Prüfen der Wirksamkeit von Arzneimitteln, (5) das Studium der Funktion des normalen vollständigen Proteins, spezifischer Anteile des Proteins oder natürlich vorkommender oder künstlich hergestellt mutanter Proteine. Natürliche vorkommende mutante Proteine liegen in Patienten mit ZF vor, während künstlich hergestelltes mutantes Protein durch positionsspezifisch ausgerichtete Sequenzänderungen konstruiert werden kann. Diese letzteren Studien können die Funktion irgendeines erwünschten Aminosäurerests im Protein durch Mutation der Nucleotide, die für diese Aminosäure kodieren, sondieren.

[0247] Mit Hilfe der obigen Techniken, können die Expressionsvektoren, die die ZF-Gensequenz oder Fragmente derselben enthalten, in menschliche Zellen, Säugerzellen aus anderen Spezies oder Nichtsäugerzellen, wie erwünscht, eingeführt werden. Die Wahl der Zellen hängt vom Zweck der Behandlung ab. Beispielsweise ist es möglich, Affen-COS-Zellen (Gluzan, Cell 23: 175 (1981)) zu verwenden, die das SV40 T-Antigen in hohem Maße bilden und die Replikation von Vektoren gestatten, die den SV40-Replikationsursprung enthalten, um zu zeigen, dass der Vektor das Proteinprodukt exprimieren kann, da eine Funktion nicht erforderlich ist. Eine ähnliche Behandlung könnte mit chinesischen Hamsterovarium-(CHO) oder Maus-NIH 3T3-Fibroblasten oder mit menschlichen Fibroblasten oder Lymphoblasten durchgeführt werden.

[0248] Der erfindungsgemäße rekombinante Kloniervektor umfasst dann die ausgewählte DNA der erfindungsgemäßen DNA-Sequenz für die Expression in einem geeigneten Wirt. Die DNA ist im Vektor funktionsfähig an eine Expressionskontrollsequenz im Molekül der rekombinanten DNA derart verkoppelt, dass das normale ZFTR-Polypeptid exprimiert werden kann. Die Expressionskontroll-Sequenz kann aus der Gruppe ausgewählt werden, die aus Sequenzen besteht, die die Expression von Genen von prokaryotischen oder eukaryotischen Zellen und ihrer Viren und Kombinationen derselben steuern. Die Expressionskontrollsequenz kann spezifisch aus der Gruppe ausgewählt werden, die aus dem lac-System, dem trp-System, dem tac-System, dem trc-System, den Hauptoperator- und Promotorregionen der Phage lambda, der Kontrollregion von fd-Beschichtungsprotein, den frühen und späten Promotoren von SV40, den von Polyom abgeleiteten Promotoren, Adenoviren, Retroviren, Baculoviren und dem Simianvirus, dem Promotor für 3-Phosphoglyceratkinase, dem Promotor von Hefensäurephosphatase, dem Promotor von Hefe-alpha-Kopplungsfaktoren und Kombinationen derselben besteht.

[0249] Die Wirtszellen, die mit dem erfindungsgemäßen Vektor transfiziert werden kann, kann aus der Gruppe ausgewählt werden, die aus E. coli, Pseudomonas, Bacillus subtilis, Bacillus stearothermophilus oder anderen Bazillen, anderen Bakterien, Hefe, Pilzen, Insekten, der Maus oder anderen Tieren oder pflanzlichen Wirten oder menschlichen Gewebezellen besteht.

[0250] Man wird sich im Klaren darüber sein, dass für die mutante DNA-Sequenz zum Exprimieren und Bilden

des mutanten Produkts ähnliche Systeme verwendet werden.

6.2 BETRACHTUNGEN BEZÜGLICH DER PROTEINFUNKTION

[0251] Um die Funktion des ZFTR-Proteins zu untersuchen, wird es bevorzugt, Epithelzellen als Empfänger zu verwenden, da die richtige funktionelle Expression das Vorliegen anderer Wege oder Genprodukte, die nur in derartigen Zellen exprimiert werden, erforderlich machen kann. Zellen, die verwendet werden können, umfassen beispielsweise menschliche Epithelzelllinien wie T84 (ATCC#CRL 248) oder PANC-1 (ATCC # CLL 1469) oder die immortalisierte T43-ZF-Nasenepithelzelllinie [Jettan et al, Science (1989)] und primäre [Yanhoskes et al. Ann. Rev. Resp. Dis. 132: 1281 (1985)] oder transformierte [Scholte et al. Exp. Cell Res. 132: 559 (1989)] menschliche Nasenpolyp- oder Luftwegzellen, Pankreaszellen [Harns und Coleman J. Cell. Sci. 87: 695 (1987) oder Schweißdrüsenzellen [Collie et al. In Vitro 21: 597 (1985)], die aus normalen oder an ZF leidenden Personen abgeleitet sind. Die ZF-Zellen können zum Untersuchen auf funktionelle Aktivität der mutanten ZF-Gene verwendet werden. Die zur Zeit zur Verfügung stehenden funktionellen Assays umfassen die Studie der Bewegung von Anionen (Cl oder I) über die Zellmembranen hindurch als Funktion der Stimulierung von Zellen durch Mittel, die die intrazellulären AMP-Spiegel erhöhen und Chloridkanäle aktivieren [Strutto et al. Proc. Nat. Acad. Sci. USA. 82: 6677 (1985)]. Andere Assays umfassen das Messen von Änderungen der Zellpotentiale durch Patch-Klemmen ganzer Zellen oder isolierter Membranen [Frizzell et al. Science 233 : 558 (1986), Welsch und Liedtke Nature 322 : 467 (1986)] oder die Studie von Innenströmen in Epithelschichten konfluenter Zellen [Widdicombe et al. Proc. Nat. Acad. Sci. 82: 6167 (1985)]. Als Alternative könnte aus dem ZF-Gen hergestellte RNA in Xenopus-Oozyten eingespritzt werden. Der Oozyt translatiert RNA in Protein und erlaubt dessen Studium. Im Zuge der Entwicklung noch spezifischerer Assays können diese ebenfalls zum Studium der Funktion des transfizierten ZFTR-Proteins verwendet werden.

[0252] "Domänenumschaltungs"-Experimente zwischen dem ZFTR und dem menschlichen, gegen Arzneimittel multiresistenten P-Glykoprotein können ebenfalls ausgeführt werden, um das Studium des ZFTR-Proteins zu fördern. Bei diesen Experimenten werden Plasmidexpressionsvektoren durch routinemäßige Verfahren aus Fragmenten der ZFTR-Sequenz und Fragmenten der Sequenz von P-Glykoprotein, die durch DNA-Ligase miteinander ligiert sind, derart konstruiert, dass ein die entsprechenden Anteile dieser beiden Proteine enthaltendes Protein durch eine mit dem Plasmid transfizierte Wirtszelle synthetisiert wird. Dieser letztere Ansatz hat den Vorteil, dass viele experimentelle Parameter, die mit der Multiarzneimittel-Resistenz assoziiert sind, gemessen werden können. Es ist daher heute möglich, die Fähigkeit von Segmenten von ZFTR, diese Parameter zu beeinflussen, zu beurteilen.

[0253] Diese Studien des Einflusses vom ZFTR auf den Innentransport dient dazu, das Gebiet des Epitheltransports in den molekularen Bereich einzuführen. Es handelt sich dabei um das erste mit dem Transport verbundene Molekül aus Epithelzellen, für das die vollständige Primärstruktur aufgezeigt ist.

[0254] Die Kenntnis bezüglich des ZFTR kann dazu benutzt werden, zu einem besseren Verständnis der Charakteristik der Epithelzelnhnembran in diesem Bereich auf molekularer Ebene zu gelangen. Beispielsweise können die Moleküle in nächster Nähe zum ZFTR durch Vernetzungsexperimente bestimmt werden. Die Hypothese, dass die Rolle von ZFTR darin liegt, Innenkanäle zu regulieren, würde voraussagen, dass diese Kanäle notwendigerweise unter diese Kategorie fallen. Die großen cDNA-Bibliotheken hoher Qualität, die für das Klonieren von ZFTR-cDNA konstruiert worden sind, werden auch für das molekulare Klonieren von cDNAs für Polypeptide nützlich sein, die andere Epithelionentransportsysteme darstellen, einschließlich andere Kanäle sowie Ko-, Gegen- und Aktivtransportsysteme.

6.3 THERAPIEN

[0255] Es versteht sich von selbst, dass das Hauptziel der verschiedenen biochemischen Studien unter Zuhilfenahme der erfindungsgemäßen Zusammensetzungen aus der Entwicklung von Therapien zum Umgehen oder Bezwingen des ZF-Defekts unter Zuhilfenahme sowohl pharmakologischer als auch von "Gentherapie"-Ansätzen besteht.

[0256] Beim pharmakologischen Ansatz werden Arzneimittel, die den ZF-Defekt umgehen oder überwinden, gesucht. Zu Beginn können Verbindungen im Wesentlichen dem Zufall entsprechend getestet werden und die Screensysteme müssen zwischen vielen Kandidatenverbindungen unterscheiden. Diese Erfindung bietet Wirtszellensysteme, die verschiedene der mutanten ZF-Gene exprimieren und zur Verwendung als Systeme für das Screenen erster Ebene besonders gut geeignet sind. Bevorzugt wird ein Zellkultursystem unter Zuhilfenahme von Säugerzellen (am bevorzugtesten menschlichen Zellen), die mit einem Expressionsvektor transfiziert sind, der eine DNA-Sequenz umfasst, die für ein eine ZF-bildende Mutation, beispielsweise die F508-Deletion, enthaltendes ZFTR-Protein kodiert, für den Screeningvorgang verwendet. Kandidatenarzneimittel werden durch Inkubieren der Zellen in Gegenwart des Kandidatenarzneimittels und Messen derjenigen Zellfunktionen, die vom ZFTR abhängen, insbesondere Messen der Innenströme, in welchem Fall das Transmembran-

potential auf einen festgelegten Wert eingestellt ist, geprüft. Um jedoch Assays in großer Zahl unterzubringen, basieren bequemere Assays beispielsweise auf der Verwendung von gegen Ionen sensiblen fluoreszierenden Farbstoffen. Um Änderungen in der Cl^- -Innenkonzentration zu erfassen, sind SPQ oder dessen Analoge nützlich.

[0257] Als Alternative könnte ein von Zellen freies System verwendet werden. Gereinigtes ZFTR könnte in künstliche Membranen rekonstituiert und Arzneimittel in einem zeltfreien Assay einem Screening unterzogen werden [A1-Agwatt, Science, (1989)].

[0258] Auf der zweiten Stufe sind Tierversuche erforderlich. Es ist möglich, ein Modell der ZF durch Beeinflussen der normalen Expression des Gegenstücks des ZF-Gens in einem Tier wie der Maus zu entwickeln. Das "Knock-out" dieses Gens durch Einführen einer mutanten Form desselben in die Keimlinie von Tieren führt zu Bildung eines Stamms von Tieren mit ZF-ähnlichen Syndromen. Das ermöglicht das Testen von Arzneimitteln, die beim Screening auf Zeltbasis der ersten Stufe erfolgversprechend gewesen sind.

[0259] Im Zuge des Erzielens weiterer Kenntnisse bezüglich der Natur des Proteines und seiner Funktion wird es möglich, Strukturen von Proteinen oder anderen Verbindungen, die mit dem ZFTR-Protein in Wechselwirkung treten, voraussagen. Das erlaubt wiederum gewisse Vorhersagen bezüglich potentieller Arzneimittel, die mit diesem Protein in Wechselwirkung treten und eine Wirkung auf die Behandlung der Patienten ausüben. Schließlich können derartige Arzneimittel konzipiert und chemisch auf der Basis von Strukturen synthetisiert werden, von denen vorausgesagt wird, dass sie für die Wechselwirkung mit Domänen von ZFTR erforderlich sind. Dieser Ansatz wird von Capsey und Delvate, Genetically Engineered Human Therapeutic Drugs Stockton Press, New York, 1988 besprochen. Diese potentiellen Arzneimittel müssen im Screeningsystem ebenfalls geprüft werden.

6.3.1 PROTEINERSATZTHERAPIE

[0260] Die Behandlung von ZF kann durch Ersetzen des defekten Proteins durch normales Protein, durch Modulieren der Funktion des defekten Proteins oder durch Modifizieren eines anderen Schritts auf dem Weg, auf dem ZFTR teilnimmt, um die physiologische Abnormalität zu berichtigen, durchgeführt werden.

[0261] Zum Ersetzen des defekten Proteins durch das normale Protein müssen angemessene Mengen des reinen ZFTR-Proteins zur Verfügung stehen. Das reine Protein kann wie früher schon beschrieben aus gezüchteten Zeltsystemen erhalten werden. Die Abgabe des Proteins an das betroffene Lungengewebe erfordert, dass es in lipidhaltigen Vesikeln verpackt ist, die die Eingabe des Proteins in die Zellmembran erleichtern. Unter Umständen kann es auch realisierbar sein, Vehikel zu verwenden, die Proteine wie Tensidprotein, beispielsweise SAP(Val) oder SAF(Phs) enthalten, die diese Funktion auf natürliche Weise, zumindest bei den Lungenalveolarzellen erfüllen (PCT-Patentanmeldung WO/8803170, Whitsett et al, 7. Mai 1988 und PCT-Patentanmeldung WO89/04327, Benson et al, 18. Mai 1989). Die ZFTR-haltigen Vesikel werden in die Luftwege durch Inhalieren oder Spülung, d. h. Techniken, die gegenwärtig bei der ZF-Behandlung angewendet werden, eingeführt (Bat et al, supra).

6.3.2 ARZNEIMITTEL THERAPIE

[0262] Die Modulation der ZFTR-Funktion kann durch Verwendung therapeutischer Mittel (von Arzneimitteln) erzielt werden. Diese können durch Zufallsansätze unter Verwendung eines Screeningprogramms identifiziert werden, bei dem ihre Wirksamkeit bezüglich des Modulierens des defekten ZFTR-Proteins in vitro überwacht wird. Bei den Screeningprogrammen können kultivierte Zeltsysteme verwendet werden, in denen das defekte ZFTR-Protein exprimiert wird. Als Alternative können Arzneimittel zum Modulieren der ZFTR-Aktivität aufgrund der Kenntnisse bezüglich der Struktur- und Funktionskorrelationen von ZFTR-Protein und der Kenntnis bezüglich des spezifischen Defekts bei den verschiedenen mutanten ZFTR-Proteinen konzipiert werden (Capsey und Delvate, supra). Es ist möglich, dass jedes mutante ZFTR-Protein ein anderes Arzneimittel für eine spezifische Modulation erfordert. Es wird dann notwendig, die spezifische(n) Mutation(en) bei jedem ZF-Patienten vor Beginn der Arzneimitteltherapie zu identifizieren.

[0263] Es können Arzneimittel so konzipiert werden, dass sie auf verschiedene Aspekte der ZFTR-Proteinstruktur oder -funktion einwirken. Beispielsweise kann ein Arzneimittel (oder Antikörper) an eine strukturelle Falte des Proteins zum Korrigieren einer defekten Struktur binden. Als Alternative kann ein Arzneimittel eventuell an einen spezifischen funktionellen Rest binden und dessen Affinität für ein Substrat oder einen Kofaktor erhöhen. Da es bekannt ist, dass Mitglieder der Klasse von Proteinen, für die ZFTR eine strukturelle Homologie aufweist, mit einer Reihe verschiedener Arzneimittel in Wechselwirkung treten, diese binden und transportieren können, ist es vernünftig zu erwarten, dass mit Arzneimitteln verbundene Therapien bei der Behandlung von ZF wirksam sein können.

[0264] Ein dritter Mechanismus für das Verbessern der Aktivität eines wirksamen Arzneimittels wäre es, die Bildung oder die Stabilität von ZFTR innerhalb der Zelle zu modulieren. Dieser Erhöhung der Menge von ZFTR

könnte für ihre defekte Funktion kompensieren.

[0265] Die Arzneimitteltherapie könnte auch zum Kompensieren der defekten ZFTR-Funktion durch Wechselwirkung mit anderen Komponenten der physiologischen oder biochemischen Wege verwendet werden, die für die Expression der ZFTR-Funktion erforderlich sind. Diese Wechselwirkungen können zum Erhöhen oder Reduzieren der Aktivität dieser Nebenproteine führen. Die Methoden für das Identifizieren dieser Arzneimittel wären denjenigen ähnlich, die oben schon für die ZFTR-bezogenen Arzneimittel beschrieben worden sind.

[0266] Bei anderen genetischen Krankheiten ist es möglich gewesen, die Konsequenzen geänderter oder fehlender normaler Funktionen durch Einführen von Änderungen in der Ernährung zu korrigieren. Dies ist in Form der Entfernung von Metaboliten, wie im Fall der Phenylketonurie, erfolgt, wo Phenylalanin in den ersten fünf Lebensjahren aus der Ernährung entfernt worden ist, um die Geistesschwäche zu verhindern, oder durch Zusatz großer Mengen von Metaboliten zur Nahrung, wie im Falle des Adenosimdesaminasemangels, wo die funktionelle Korrektur der Aktivität des Enzyms durch Zusatz des Enzyms zur Nahrung erzielt werden kann. So kann, nachdem die Einzelheiten der ZFTR-Funktion geklärt und der grundlegende Defekt bei der ZF definiert worden sind, die Therapie durch Manipulationen der Ernährung erzielt werden.

[0267] Der zweite potentielle therapeutische Ansatz ist die sogenannte "Gentherapie", bei der normale Kopien des ZF-Gens in den Patienten eingeführt werden, um erfolgreich für das normale Protein in den Schlüssel-epithelzellen der betroffenen Gewebe zu kodieren. Es ist äußerst wichtig, zu versuchen, dies bei den Luftwegepithelzellen der Atemwege zu erzielen. Das ZF-Gen wird an diese Zellen in der Form geliefert, in der es aufgenommen und für ausreichend Protein kodieren kann, um die Regulierfunktion zu bieten. Dadurch wird die Lebensqualität und -länge des Patienten wesentlich verbessert. Schließlich besteht das Ziel natürlich darin, das Gen an alle betroffenen Gewebe zu liefern.

6.3.3 GENTHERAPIE

[0268] Ein Ansatz bezüglich der ZF-Therapie besteht darin, eine normale Version des ZF-Gens in das Luftwegeepithel betroffener Patienten zu inserieren. Dabei ist es wichtig, zu beachten, dass das Atmungssystem die Hauptursache der Morbidität und Mortalität bei ZF ist; die Erkrankung des Pankreas ist zwar ein wichtiges Merkmal, sie ist jedoch durch Zugabe von Enzym zur Nahrung heute relativ einfach zu behandeln. So würde die somatische Zellgentherapie [zur Rezension vergleiche T. Friedmann, *Science* 244: 1275 (1989)], mit gezielter Ansteuerung der Luftwege selbst die schwersten, mit der ZF assoziierten Probleme mildern.

A. Retrovirusvektoren. Retroviren sind als bevorzugter Vektor für Experimente bei der somatischen Gentherapie betrachtet worden, denn sie weisen eine hohe Infektionsineffizienz und eine stabile Integration und Expression auf [Orkin et al *Prog. Med. Genet* 7: 130 (1988)]. Ein möglicher Nachteil besteht darin, dass für die Retrovirusintegration eine Zellteilung erforderlich ist, so dass die Zielzellen im Luftweg eventuell vor der Retrovirusinfektion vielleicht durch chemische Mittel in den Zellzyklus gestoppt werden müssen. Die cDNA des ZF-Gens voller Länge kann in einen Retrovirusvektor kloniert und entweder von dessen endogenem Promotor oder vom Retrovirus-LRT (Langterminalwiederholung) her getrieben werden. Es wäre zu erwarten, dass die Expression des normalen Proteins in Mengen, die eventuell nur 10% des endogenen mutanten Proteins betragen können, bei ZF-Patienten heilsam sind, denn es handelt sich dabei um eine rezessive Krankheit. Das Abgeben des Virus könnte durch Aerosol oder Instillation in die Luftröhre erfolgen.

B. Andere Virusvektoren. Andere Abgabesysteme, die verwendet werden können, umfassen das Adeno-assoziierte Virus [AAV, McLaughlin et al, *J. Virol* 62: 1963 (1988)], das Vaccinia-Virus [Moss et al *Annu. Rev. Immunol*, 5: 305, 1987], das Rinderpapillomvirus [Rasmussen et al, *Methods Enzymol* 139: 642 (1987)] oder Mitglieder der Herpesvirusgruppe wie das Epstein-Barr-Virus [Margolis et al, *Mol. Cell. Biol* 8: 2937 (1988)]. Obwohl noch viel bezüglich ihrer grundlegenden Biologie zu lernen ist, ist der Gedanke der Verwendung eines Virusvektors mit natürlichem Tropismus für die Atemwege (z. B. das Atmungssynzytiumvirus, Echovirus, Coxsackievirus usw.) eine Möglichkeit.

C. Nichtvirus-Genübertragung. Andere Methoden zum Inserieren des ZF-Gens in das Atemwegeepithel können ebenfalls produktiv sein; viele dieser Methoden weisen eine geringere Effizienz auf und würden potentiell eine Infektion *in vitro*, die Selektion von Transfektanten und eine Reimplantation erforderlich machen. Dies würde Calciumphosphat, DEAE-Dextran, die Elektroporation und Protoplastfusion umfassen. Eine besonders ansprechende Idee ist die Verwendung von Liposom, die eventuell *in vivo* ausgeführt werden kann [Ostro, *Liposomes*, Marcel-Dekker, 1987]. Synthetische kationische Lipide wie DOTMA [Felger et al *Proc. Natl. Acad. Sci. USA* 84: 7413 (1987)] erhöhen eventuell die Effizienz und erleichtern die Durchführung dieses Ansatzes.

6.4 TIERISCHE ZF-MODELLE

[0269] Die Bildung eines Maus- oder anderen tierischen Modells für ZF wird für das Verständnis der Krankheit und das Testen bezüglich möglicher Therapien äußerst wichtig sein (bezüglich einer allgemeinen Besprechung

der Herstellung von Tiermodellen vergleiche man Erickson, Am. J. Hum. Genet 43: 582 (1988). Zur Zeit steht kein tierisches Modell der ZF zur Verfügung. Die Evolutionskonservierung des ZF-Gens (wie sie durch Kreuzspezies-Hybridisierungsblots für E4.3 und H1.6 aufgezeigt worden ist) wie in **Fig. 4** gezeigt, weist darauf hin, dass ein orthologes Gen in der Maus (im folgenden mit mZF und dessen entsprechendes Protein als mZFTR bezeichnet) existiert und es wird möglich sein, dieses in genomischen Maus- und cDNA-Bibliotheken unter Zuhilfenahme der menschlichen ZF-Gensonden zu klonieren. Es ist zu erwarten, dass die Bildung einer spezifischen Mutation im Mausgen, die der F508-Mutation analog ist, die optimalste sein wird, um den Phenotyp zu reproduzieren, obwohl eine vollständige Inaktivierung des mZFTR-Gens ebenfalls ein nützlicher zu bildender Mutant sein wird.

A. Mutagenese. Die Inaktivierung des mZF-Gens kann durch chemische [z. B. Johnson et al, Proc. Natl. Acad. Sci. USA 78: 3138 (1981)] oder Röntgertmutagenese [Popp et al, J. Mol. Biol. 127: 141 (1979)] von Mausgameten, gefolgt von der Fertilisation erzielt werden. Nachkommen, die bezüglich der Inaktivierung von mZFTR heterozygot sind, können daraufhin durch Southern Blotting identifiziert werden, um den Verlust eines Allels durch Dosierung oder die Unfähigkeit, ein elterliches Allel zu erben, wenn ein RFLP-Marker beurteilt wird, aufzuzeigen. Dieser Ansatz ist früher schon erfolgreich zum Identifizieren von Mausmutanten für α -Globin [Whitney et al, Proc. Natl. Acad. Sci. USA 77: 1087 (1980)], Phenylalaninhydroxylase [McDonald et al Pediatr. Res 23: 63 (1988)] und Kohlen säureanhydrase II [Lewis et al Proc. Natl. Acad. Sci. USA 85: 1962 (1988)] angewendet worden.

B. Transgene Aspekte. Eine normale oder mutante Version von ZFTR oder mZFTR kann in die Mauskeimlinie unter Zuhilfenahme heutiger Standardverfahren der Oozyteninjektion [Camper, Trends in Genetics (1988)] inseriert werden; als Alternative kann, falls es wünschenswert ist, das endogene mZF zu inaktivieren oder zu ersetzen, das homologe Rekombinationssystem unter Zuhilfenahme von Embryostammzellen (ES-Zellen) [Capecchi, Science 244: 1288 (1989)] angewendet werden.

1. Oozyteninjektion.

[0270] Das Einführen einer oder mehrerer Kopien des normalen oder mutanten mZF-Gens an eine willkürliche Stelle in der Mauskeimlinie kann durch Mikroinjizierung des ProNucleus eines gerade befruchteten Mausoozyten, gefolgt von der Reimplantation in eine pseudoschwangere Pflegemutter erzielt werden. Die lebend geborenen Mäuse können daraufhin einem Screenen bezüglich Integrierten unter Zuhilfenahme der Analyse der Schwanz-DNA bezüglich des Vorliegens menschlicher ZF-Gensequenzen unterzogen werden. Das gleiche Protokoll lässt sich zum Inserieren eines mutanten mZF-Gens verwendet. Um ein Mausmodell zu bilden, würde man dieses Transgen in einen Maushintergrund einsetzen, wo das endogene mZF-Gen inaktiviert worden ist, entweder durch Mutagenese (vergleiche oben) oder durch homologe Rekombination (vergleiche unten). Bei dem Transgen kann es sich entweder um: a) eine vollständige genomische Sequenz, obgleich die Größe derselben (ca. 250 kb) es erforderlich machen würde, dass sie als künstliches Hefechromosom oder Chromosomfragment injiziert würde, oder b) eine cDNA mit entweder dem natürlichen Promotor oder einem heterologen Promotor oder c) ein "Minigen" handeln, das die gesamte Kodierregion und verschiedene andere Elemente wie Introne, Promotoren und 3'-flankierende Elemente enthält, von denen es sich gezeigt hat, dass sie für eine optimale Expression notwendig sind.

2. Retrovirusinfektion von Embryos im Frühstadium.

[0271] Diese Alternative involviert das Inserieren des ZFTR- oder mZF-Gens in einen Retrovirusvektor und das direkte Infizieren von Mausembryos im Frühstadium der Entwicklung unter Bildung einer Chimäre [Soriano et al Cell 46: 19 (1986)]. Mindestens einige derselben führen zur Keimlinientransmission.

3. ES-Zellen und homologe Rekombination.

[0272] Der Ansatz mit Embryostammzellen (Capecchi, supra und Capecchi, Trends Genet 5: 70 (1989)) bietet die Möglichkeit des Durchführens der Genübertragung und des darauffolgenden Screenens der so erhaltenen totipotenten Zellen zum Identifizieren der seltenen homologen Rekombinationsfälle. Auf das Identifizieren hin können diese zur Bildung von Chimären durch Einspritzen von Mausblastozyten verwendet werden und ein Teil der so gebildeten Mäuse werden die Keimlinientransmission aus der rekombinanten Linie aufweisen. Es gibt verschiedene Wege, auf die dies bei der Bildung eines Mausmodells für ZF nützlich sein könnte:

a) Die Inaktivierung des mZF-Gens kann auf bequeme Weise durch Konstruieren eines DNA-Fragments erzielt werden, das Sequenzen aus einem mZFTR-Exon enthält, das einen auswählbaren Marker, wie neo flankiert. Die homologe Rekombination führt zum Inserieren der neo-Sequenzen in die Mitte eines Exons unter Inaktivierung von mZFTR. Die homologen Rekombinationsereignisse (gewöhnlich ca. 1 unter 1000) können von den heterologen durch DNA-Analyse einzelner Klone (gewöhnlich unter Anwendung von PCR,

Kim et al Nucleic Acids Res. 16: 8887 (1988), Joyner et al. Nature 338: 153 (1989), Zimmer et al supra, Seite 150) oder durch Zuhilfenahme einer negativen Selektion gegen die heterologen Vorfälle (wie beispielsweise die Anwendung des HSV TK-Gens am Ende des Konstrukts, gefolgt von der Gancyclovir-Wahl, Mansow et al, Nature 336: 348 (1988) unterschieden werden. Diese inaktivierte mZFTR-Maus kann daraufhin verwendet werden, um ein mutantes ZF-Gen oder ein die F508-Abnormität oder irgendeine andere erwünschte Mutation enthaltendes mZF-Gen einzuführen.

b) Es ist möglich, dass spezifische Mutanten vom mZFTR in einem Schritt gebildet werden können. Beispielsweise kann ein Konstrukt, das mZF-Intron-9-Sequenzen am 5'-Ende, ein selektierbares neo-Gen in der Mitte und ein Intro 9 + Exon 10 (das die Mausversion der F508-Mutation enthält) am 3'-Ende enthält, hergestellt werden. Ein homologer Rekombinationsvorfall würde zum Inserieren des neo-Gens in Intron 9 und das Ersetzen von Exon 10 durch die mutante Version führen.

c) Falls das Vorliegen des auswählbaren neo-Markers im Intron die Expression des mZF-Gens ändert, wäre es möglich, es in einem zweiten homologen Rekombinationsschritt zu exzidieren.

d) Es ist auch möglich, Mutationen in der Mauskeimlinie durch Injizieren von Oligonucleotiden, die die interessierende Mutation enthalten, und Screenen der dabei gebildeten Zellen durch PCR zu bilden.

[0273] Bei dieser Ausführungsform der Erfindung ist hauptsächlich ein Mausmodell für die zystische Fibrose in Betracht gezogen worden. **Fig. 4** zeigt die Kreuzspezieshybridisierung nicht nur zur Maus-DNA sondern auch zur Rinder-, Hamster- und Hühner-DNA. So wird in Betracht gezogen, dass ein orthologes Gen auch in vielen anderen Spezies vorhanden ist. Es wird daher in Betracht gezogen, dass es möglich sein wird, unter Zuhilfenahme einer ähnlichen Technologie andere Tiermodelle zu bilden.

Patentansprüche

1. Gereinigtes DNA-Molekül, das ein Gen der zystischen Fibrose umfasst, wobei das Gen eine DNA-Sequenz umfasst, die aus der Gruppe ausgewählt wird, die aus Folgenden besteht:

(a) DNA-Sequenzen, die der DNA-Sequenz in **Fig. 1** von der Aminosäurerestposition 1 zur -position 1480 oder einer polymorphen Form der DNA-Sequenz entsprechen, in der die Nucleotidsubstitution in der DNA-Sequenz die wesentliche Funktion des dadurch codierten Polypeptids nicht beeinflusst;

(b) DNA-Sequenzen, die ein normales Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR), das die Sequenz der **Fig. 1** entsprechend von der Aminosäurerestposition 1 bis zur - position 1480 aufweist, oder eine polymorphe Form des Polypeptids codiert, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydrophatisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst;

(c) DNA-Sequenzen, die einem Fragment der DNA-Sequenz in **Fig. 1** entsprechen, das mindestens 16 sequentielle Nucleotide zwischen den Nucleotidsequenzpositionen 134 und 4573 umfasst, und die in der Lage sind, nach einem normalem oder einem defekten Gen oder Genprodukt der zystischen Fibrose zu suchen oder das Vorliegen derselben zu bestimmen, und

(d) DNA-Sequenzen, die mindestens 16 sequentielle Nucleotide umfassen und ein Fragment der Aminosäuresequenz der **Fig. 1** codieren, die in der Lage sind, nach einem normalem oder einem defekten Gen oder Genprodukt der zystischen Fibrose zu suchen oder das Vorliegen derselben zu bestimmen, und

(e) DNA-Sequenzen, die ein Epitop codieren, das durch mindestens 18 sequentielle Nucleotide in der DNA-Sequenz der **Fig. 1** zwischen den Aminosäurerestpositionen 1 und 1480 codiert ist.

2. Gereinigtes DNA-Modul aus (a), (b) oder (c) nach Anspruch 1, das des Weiteren mindestens eine Mutation in der DNA-Sequenz enthält, die, falls sie in den Zellen des menschlichen Körpers exprimiert ist, mit einer geänderten Zellfunktion assoziiert ist, die mit der zystischen Fibrose in Beziehung steht, wobei die Mutation eine Deletion von 3 Nucleotiden einschließt, die Phenylalanin in der Aminosäurerestposition 508 codieren.

3. Sonde aus gereinigter Nucleinsäure, die eine DNA- oder RNA-Nucleotidsequenz umfasst, die der Sequenz den Abschnitten (c), (d) oder (e) nach Anspruch 1 oder 2 entspricht.

4. Nucleinsäuresonde nach Anspruch 3, wobei die Sequenz AAA GAA AAT ATC ATC TTT GGT GTT und ihr Komplement umfasst.

5. Rekombinanter Kloniervektor, der das DNA-Molekül nach Anspruch 1 oder 2 umfasst.

6. Nichtmenschlicher, mit dem Vektor nach Anspruch 5 transformierter Wirt.

7. Methode für die Herstellung eines normalen oder mutanten Transmembrankonduktanzregulator-Poly-

peptids der zystischen Fibrose (ZFTR), die folgende Schritte umfasst:

- (a) Kultivieren einer Wirtszelle, die mit dem Vektor nach Anspruch 5 in einem Medium und unter Bedingungen transfiziert ist, die für die Expression eines normalen oder mutanten ZFTR-Polypeptids günstig sind, und
- (b) Isolieren des exprimierten normalen oder mutanten ZFTR-Polypeptids.

8. Gereinigtes DNA-Molekül, das eine DNA-Sequenz umfasst, die ein mutantes Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR), das eine Sequenz nach **Fig. 1** für Aminosäurerestpositionen 1 bis 1480 aufweist, oder eine polymorphe Form des Polypeptids codiert, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydrophathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, des Weiteren durch eine Deletion der drei Nucleotide gekennzeichnet, die Phenylalanin an der Aminosäurerestposition 508 codieren.

9. Gereinigtes Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR), dadurch gekennzeichnet, dass es eine Peptidmolmasse von 170.000 Dalton aufweist und aus der Gruppe ausgewählt wird, die aus Folgenden besteht:

- (a) einem normalen ZFTR-Polypeptid mit einer die Zellmembranionenkonduktanz beeinflussenden Aktivität, das die Sequenz nach **Fig. 1** für die Aminosäurepositionen 1 bis 1480 aufweist, oder eine polymorphe Form des normalen ZFTR-Polypeptids, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydrophathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, und
- (b) einem mutanten ZFTR-Polypeptid mit einer zystischen Fibroseaktivität in menschlichen Zellen, das die Sequenz nach **Fig. 1** für die Aminosäurepositionen 1 bis 1480 aufweist, mit Ausnahme der Deletion von Phenylalanin an der Aminosäurerestposition 508.

10. Polypeptid, das durch die Expression einer DNA-Sequenz nach Anspruch 1 oder 2 codiert ist, wobei das Polypeptid die immunologische Aktivität eines normalen oder mutanten Transmembrankonduktanzregulator-Polypeptids für zystische Fibrose (ZFTR) aufweist.

11. Verfahren für das Isolieren des normalen oder mutanten ZFTR-Proteins nach Anspruch 10 aus das Protein enthaltenden Zellen, welches Verfahren folgende Schritte umfasst:

- (a) Löslichmachen des Proteins einer ausgewählten Zellmembran, in der entweder das normale oder das mutante ZFTR-Protein exprimiert ist, unter Bildung einer Lösung des ZFTR-Proteins,
- (b) Trennen des ZFTR-Proteins von der Lösung durch Kontaktieren der Lösung mit Antikörpern gegen das normale oder mutante ZFTR-Protein, wobei die Antikörper auf einem Substrat immobilisiert werden,
- (c) Spülen des Substrats zur Entfernung des Proteins, das nicht an den Antikörpern anhaftet,
- (d) Freisetzen des ZFTR-Proteins von den Antikörpern zur Isolierung des ZFTR-Proteins und
- (e) Reinigen des ZFTR-Proteins zur Entfernung von jeglichem anderem Mammalienprotein.

12. Methode für das Screening einer Versuchsperson, um zu bestimmen, ob diese Versuchsperson ein Träger der zystischen Fibrose (ZF) oder ZF-Patient ist, welche Methode folgende Schritte umfasst:

- (a) zur Verfügungstellen einer biologischen Probe der Versuchsperson, die dem Screening unterzogen werden soll, und
- (b) Durchführen einer Untersuchung für die Bestimmung, in der biologischen Probe, des Vorliegens von mindestens einem Mitglied der Gruppe, die aus Folgendem besteht: einem normalen ZF-Gen, normalen ZF-Genprodukten, einem mutanten ZF-Gen, mutanten ZF-Genprodukten und Mischungen derselben, einem normalen ZFTR-Polypeptid mit einer die Zelltransmembran-Ionenkonduktanz beeinflussenden Aktivität und der Sequenz nach **Fig. 1** für Aminosäurepositionen 1 bis 1480 oder einer polymorphen Form des normalen ZFTR-Polypeptids, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben, oder ihr hydrophathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, einem mutanten ZFTR-Polypeptid mit einer zystischen Fibroseaktivität in menschlichen Zellen und der Sequenz nach **Fig. 1** für Aminosäurepositionen 1 bis 1480, mit der Ausnahme der Deletion von Phenylalanin in der Aminosäurerestposition 508, wobei es sich bei der Untersuchung um
 - (i) eine Hybridisierung involvierende Untersuchung oder
 - (ii) eine immunologische Untersuchung handeltund/oder die unter Zuhilfenahme eines oder mehrerer der Folgenden erfolgt: Hybridisierung unter Zuhilfenahme spezifischer Oligonucleotide, des direkten DNA-Sequenzierens, von Restriktionsenzymen, der Diskriminierung auf der Basis der elektrophoretischen Mobilität in Gelen mit denaturierendem Mittel, RNase-Protektion, chemischer Spaltung oder einem durch Ligase-mediatiertes Bestimmungsverfahren.

13. Methode nach Anspruch 12, wobei es sich bei der Untersuchung um eine die Hybridisierung involvierende Untersuchung handelt und die biologische Probe mindestens einen Teil des Genoms der Versuchsperson umfasst.

14. Methode nach Anspruch 13, wobei die Untersuchung des Weiteren eine gelabelte Nucleotidsonde nach Anspruch 3 umfasst.

15. Methode nach Anspruch 14, wobei die Sonde die Nucleotidsequenz von Anspruch 4 umfasst.

16. Methode nach Anspruch 12, wobei es sich bei der Untersuchung eine immunologische Untersuchung handelt und die biologische Probe ein normales Transmembrankonduktanzregulator-Polypeptid für zystische Fibrose (ZFTR) der Versuchsperson und/oder ein mutantes ZFTR-Polypeptid der Versuchsperson enthält: wobei das normale ZFTR-Polypeptid die Sequenz nach **Fig. 1** für Aminosäurepositionen 1 bis 1480 oder eine polymorphe Form des Polypeptids aufweist, in der die Aminosäuresubstitution in den verschiedenen Regionen des Polypeptids mit der Sequenz nach Anspruch das wesentliche Funktionieren derselben oder ihr hydropathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, und das mutante ZFTR-Polypeptid die Sequenz nach **Fig. 1** für Aminosäurepositionen 1 bis 1480, mit Ausnahme der Deletion von Phenylalanin an der Aminosäurerestposition 508 aufweist.

17. Methode nach Anspruch 16, wobei die Untersuchung des Weiteren einen für ein normales ZFTR-Polypeptid spezifischen Antikörper umfasst.

18. Methode nach Anspruch 16, wobei die Untersuchung des Weiteren einen für ein mutantes ZFTR-Polypeptid spezifischen Antikörper umfasst.

19. Methode nach Anspruch 17 oder 18, wobei der Antikörper mindestens ein Monoclonalantikörper ist.

20. Methode nach Anspruch 12, wobei die Versuchsperson ein menschlicher Fötus in utero ist.

21. Methode nach Anspruch 14 oder 15, wobei die Untersuchung des Weiteren mindestens eine zusätzlich Nucleotidsonde nach Anspruch 3 umfasst, vorausgesetzt, jede Nucleotidsonde in dieser Untersuchung ist verschieden.

22. Verfahren für das Screening eines potentiellen Trägers der zystischen Fibrose (ZF) oder eines an zystischer Fibrose leidenden Patienten zur Identifizierung des Vorliegens einer identifizierten zystischen Fibrose-mutation im ZF-Gen, das ein Transmembrankonduktanzregulator-Polypeptid für zystische Fibrose (ZFTR) mit der Sequenz nach **Fig. 1** für Aminosäureposition 1 bis Position 1480 oder eine polymorphe Form des Polypeptids aufweist, in der eine Aminosäuresubstitution in verschiedenen Regionen des Polypeptids mit der Sequenz nach **Fig. 1** das wesentliche Funktionieren derselben oder ihr hydropathisches Profil oder die sekundäre oder tertiäre Struktur nicht beeinflusst, wobei das Verfahren folgende Schritte umfasst:

(a) Isolieren von genomischer DNA des potentiellen ZF-Trägers oder des potentiellen Patienten;

(b) Hybridisieren einer DNA-Sonde nach Anspruch 3 auf die isolierte genomische DNA, wobei die DNA-Sonde die zystische Fibrose-mutation in dem ZF-Gen umfasst, und die DNA-Sonde in der Lage ist, die zystische Fibrose-mutation zu bestimmen und;

(c) Behandeln des genomischen DNA zur Bestimmung des Vorliegens der DNA-Sonde, wobei das Vorliegen oder die Abwesenheit der zystischen Fibrose-mutation einer vorbestimmten Art der Hybridisierung entsprechend angezeigt wird.

23. Verfahren für das Feststellen von Träger der zystischen Fibrose oder an zystischer Fibrose leidenden Patienten, wobei das Verfahren aus Folgendem besteht: Kontaktieren eines ZF-Genes des Patienten oder Trägers mit einer Restriktionsendonuclease, Bestimmen des Vorliegens oder der Abwesenheit einer Restriktionsendonucleasestelle in dem ZF-Gen und Vergleichen des Musters der Restriktionsendonucleasestellen in dem ZF-Gen mit einem normalen ZF-Gen, das der Nucleotidsequenz entspricht, die das Polypeptid zwischen dem Aminosäurerest in Position 1 und 1480 der **Fig. 1** codiert, wobei das Vorliegen zusätzlicher Stellen am normalen ZF-Gen oder die Abwesenheit von Stellen im normalen ZF-Gen auf ein mutantes ZF-Gen hinweist.

24. Verfahren für das Feststellen von Trägern der zystischen Fibrose, wobei das Verfahren aus Folgendem besteht: Bestimmen der Differentialmobilität von Heteroduplex-PCR-Produkten in Polyacrylamidgelen durch Insertionen oder Deletionen in dem mutanten zystischen Fibrose-(ZF)-Gen im Vergleich mit der normalen ZF-Gen, das der Nucleotidsequenz entspricht, das das Polypeptid zwischen der Aminosäurerestposition 1 und

1480 der **Fig. 1** codiert.

25. Versuchskit für das Untersuchen auf das Vorliegen eines normalen Gens der zystischen Fibrose (ZF), das der Nucleotidsequenz entspricht, die das Polypeptid zwischen der Aminosäurerestposition 1 und 1480 der **Fig. 1** codiert, oder eines mutanten ZF-Gens, das mindestens eine drei Basenpaare umfassende Deletion im normalen ZF-Gen aufweist, welche Deletion zur Deletion eines Phenylalanins an der Aminosäurerestposition 508 der **Fig. 1** führt, durch einen Immunoassay, der Folgendes umfasst:

- (a) einen Antikörper, der spezifisch an ein Genprodukt des normalen oder mutanten ZF-Gens nach Anspruch 9 bindet,
- (b) ein Reagensmittel für das Feststellen der Bindung des Antikörpers an das Genprodukt und
- (c) wobei der Antikörper und das Reagensmittel in Mengen vorliegen, die zur Durchführung des Immunoassays effektiv sind.

26. Versuchskit für das Untersuchen auf das Vorliegen eines normalen Gens der zystischen Fibrose (ZF), das der Nucleotidsequenz entspricht, die das Polypeptid zwischen der Aminosäurerestposition 1 und 1480 der **Fig. 1** codiert, oder eines mutanten ZF-Gens, das mindestens eine drei Basenpaare umfassende Deletion im normalen ZF-Gen aufweist, welche Deletion zur Deletion eines Phenylalanins an der Aminosäurerestposition 508 der **Fig. 1** führt, durch eine Hybridisierung, die Folgendes umfasst:

- (a) eine Oligonucleotidsonde nach Anspruch 3, die spezifisch an das normale oder das mutante ZF-Gen bindet,
- (b) ein Reagensmittel für das Feststellen der Hybridisierung der Oligonucleotidsonde an dem normalen oder dem mutanten ZF-Gen und
- (c) wobei die Sonde und das Reagensmittel in Mengen vorliegen, die zur Durchführung des Hybridisierungsassays effektiv sind.

27. Immunologisch aktiver gegen die zystische Fibrose wirkender polyclonaler oder monoclonaler Transmembrankonduktanzregulator (ZFTR)-Antikörper, der für ein ZFTR-Polypeptid nach Anspruch 9 spezifisch ist.

28. Hybridom, das einen monoclonalen Antikörper bildet, der für ein Transmembrankonduktanzregulator-Polypeptid der zystischen Fibrose (ZFTR) nach Anspruch 9 spezifisch ist.

29. Zusammensetzung für die Behandlung der zystischen Fibrose bei einem Patienten, die eine therapeutisch wirksame Menge des Proteins nach Abschnitte (a) des Anspruchs 9 in Kombination mit einem pharmazeutisch akzeptablen Träger umfasst.

30. Zusammensetzung nach Anspruch 29, wobei der Träger ein Lungentensidprotein umfasst, um die Anwendung der Zusammensetzung auf die Epitelzellen der Atmungswege zu erleichtern.

31. Zusammensetzung für die Gentherapie gegen zystische Fibrose, die ein DNA-Molekül nach Anspruch 1 und ein Vehikel für die Verabreichung des DNA-Moleküls an eine Zelle bei einem an zystischer Fibrose leidenden Patienten umfasst.

32. Zusammensetzung nach Anspruch 31, wobei es sich bei dem Vehikel um einen rekombinanten Vektor handelt.

33. Nichtmenschliches Tier, das ein heterologes Zellsystem umfasst, das einen rekombinanten Kloniervektor nach Anspruch 5 umfasst, welcher Vektor die Symptome der zystischen Fibrose in dem Tier induziert.

34. Transgene Maus, die die Symptome der zystischen Fibrose aufweist, die durch die Schritte erzielt werden können, die Folgendes umfassen: Einführen eines zystischen Fibrosegens nach Anspruch 2, das ein mutantes Transmembrankonduktanzregulatorprotein der zystischen Fibrose codiert, welches mindestens eine Deletion von Phenylalanin in der 508 in **Fig. 1** entsprechenden Position enthält, in einen Mausoozyten oder einen Mausembryo und gegebenenfalls Inaktivieren des endogenen Mausgens für zystische Fibrose.

Es folgen 32 Blatt Zeichnungen

Anhängende Zeichnungen

FIG.1.

1 AATTGGAAGCAAATGACATCACAGCAGGTCAGAGAAAAAGGGTTGAGCGGCAGGCACCCA
 61 GAGTAGTAGGTCTTTGGCATTAGGAGCTTGAGCCCAGACGGCCCTAGCAGGGACCCAGC
 M Q R S P L E K A S V V S K L F 16
 121 GCCCGAGAGACCATGCAGAGGTCGCTCTGGAAAAGGCCAGCGTTGTCTCCAACTTTTT
 F S W T R P I L R K G Y R Q R L E L S D 36
 181 TTCAGCTGGACCAGACCAATTTGAGGAAAGGATACAGACAGCGCCTGGAATTGTCAGAC
 I Y Q I P S V D S A D N L S E K L E H E 56
 241 ATATACCAAATCCCTTCTGTTGATTCTGCTGACAATCTATCTGAAAAATTGGAAAGAGAA
 W D R E L A S K K N P K L I N A L R R C 76
 301 TGGGATAGAGAGCTGGCTTCAAAGAAAAATCCTAAATCATTAAATGCCCTTCGGCGATGT
 F F W R F M F Y G I F L Y L G E V T K A 96
 361 TTTTCTGGAGATTTATGTTCTATGGAATCTTTTTATATTAGGGGAAGTCACCAAAGCA
 V O P L L L G R I I A S Y D P D N K E E 116
 421 GTACAGCCTCTCTTACTGGGAAGAATCATAGCTTCTATGACCCGGATAACAAGGAGGAA
 R S I A I Y L G I G L C L L F I V R T L 136
 481 CGCTCTATCGCGATTTATCTAGGCATAGGCTTATGCCCTCTCTTATTGTGAGGACTGT
 L L H P A I F G L H H I G M Q M R I A M 156
 541 CTCCTACACCCAGCCATTTTGGCCCTCATCACATTGGAATGCAGATGAGAATAGCTATG
 F S L I Y K K T L K L S S R V L D K I S 176
 601 TTTAGTTGATTATAAGAAGACTTTAAAGCTGTCAAGCCGTGTTCTAGATAAAATAAGT
 I G Q L V S L L S N N L N K F D E G L A 196
 661 ATTGGACAACCTGTTAGTCTCCTTCCAAACAACCTGAACAAATTTGATGAAGGACTTGCA
 L A H F V W I A P L O V A L L M G L I W 216
 721 TTGGCACATTTGTTGATCGCTCCTTGGCAAGTGGCACTCCTCATGGGGCTAATCTGG
 E L L O A S A F C G L G F L I V L A L F 236
 781 GAGTTGTTACAGGCTCTGCCTTCTGTGGACTTGGTTTCTGATAGTCTTGCCTTTTT
 O A G L G R M M H K Y R D Q R A G K I S 256
 841 CAGGCTGGGCTAGGGAGAATGATGATGAAGTACAGAGATCAGAGACTGGGAAGATCAGT
 E R L V I T S E M I E N I O S V K A Y C 276
 901 GAAAGACTTGTGATTACCTCAGAAATGATTGAAAATATCCAATCTGTTAAGGCATACTGC
 W E E A M E K H I E N L R Q T E L K L T 296
 961 TGGGAAGAAGCAATGGAAAAAATGATTGAAAATTAAGACAACAGAACTGAAACTGACT
 R K A A Y V R Y F N S S A F F F S G F F 316
 1021 CGGAAGGCAGCCTATGTGAGATACTTCAATAGCTCAGCCTTCTTCTCAGGGTTCTTT
 V V F L S V L P Y A L I M G I I L R K I 336
 1081 GTGGTGTTTTTATCTGTGCTTCCCTATGCACTAATCAAAGGAATCATCTCCGAAAATA
 F T T I S F C I V L R M A V T R Q F P W 356
 1141 TTCACCACCATCTCATTCTGCATTGTTCTGCGCATGGCGGTCACTCGGCAATTTCCCTGG
 A V Q T W Y D S L G A I N K I O D F L O 376
 1201 GCTGTACAAACATGGTATGACTCTCTTGGAGCAATAAACAATAACAGGATTTCTTACAA
 K Q E Y K T L E Y N L T T T E V V M E N 396
 1261 AAGCAAGAATATAAGACATTGGAATATAACTTAACGACTACAGAAGTAGTGATGGAGAAT
 V T A F W E E G F G E L F E K A K O N N 416
 1321 GTAACAGCCTTCTGGGAGGAGGATTTGGGAATTATTTGAGAAAGCAAAAACAAACAAT

FIG.1 (cont'd)

1381 N N R K T S N G D D S L F F S N F S L L 436
 AACAAATAGAAAACTTCTAATGGTGATGACAGCCTCTTCTTCAGTAATTTCTCACTTCTT

1441 G T P V L K D I N F K I E R G O L L A V 456
 GGTACTCCTGTCTGAAAGATATTAATTTCAAGATAGAAAAGAGGACAGTTGTTGGCGGTT

1501 A G S T G A G K I T S L L M H I M G E L E 476
 GCTGGATCCACTGGAGCAGGCAAGACTTCACTTCTAATGATGATTATGGGAGAAGCTGGAG

1561 P S E G K I K H S G R I S F C S O F S W 496
 CCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCAATCTGTTCTCAGTTTCTCTGG

1621 I M P G T I K E N I I F G V S Y D E Y R 516
 ATTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTCCTATGATGAATATAGA

1681 Y R S V I K A C O L E E D I S K F A E K 536
 TACAGAAGCGTCATCAAAGCATGCCAAGACTAGAAGAGGACATCTCCAAGTTTGCAGAGAAA

1741 D N I V L G E G G I T L S G G O R A R I 556
 GACAATATAGTTCTTGGAGAGGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAATT

1801 S L A N A V Y K D A D L Y L L D S P F G 576
 TCTTTAGCAAGAGCAGTATACAAAGATGCTGATTTGTATTTATTAGACTCTCCTTTTGGAA

1861 Y L D V L T E K E I F E S C V C K L M A 596
 TACCTAGATGTTTTAACAGAAAAAGAAATATTTGAAAGCTGTGTCTGTAAGCTGATGGCT

1921 N K T R I L V T S K M E H L K K A D K I 616
 AACAAAAGTAGGATTTTGGTCACTTCTAAAATGGAACATTTAAAGAAAGCTGACAAAATA

1981 L I L N E G S S Y F Y G T F S E L O N L 636
 TTAATTTTGAATGAAGGTAGCAGCTATTTTTATGGGACATTTTCAGAACTCCAAAATCTA

2041 Q P D F S S K L M G C D S F D Q F S A E 656
 CAGCCAGACTTTAGCTCAAAAAGCTGAGGATGTGATTCTTTTCGACCAATTTAGTGCAGAA

2101 R R N S I L T E Y L H R F S L E G D A P 676
 AGAAGAAATCAATCCTAACTGAGACCTTACACCGTTTCTCATTAGAAGGAGATGCTCCT

2161 V S W T E T K K Q S F K O T G E F G E K 696
 GTCTCCTGGACAGAAACAAAAACAATCTTTTAAACAGACTGGAGAGTTGGGGAAAAA

2221 R K N S I L N P I N S I R K F S I V Q K 716
 AGGAAGAATTCTATTCTCAATCCAATCAACTCTATACGAAAATTTCCATTGTGCAAAAG

2281 T P L O M N G I E E D S D E P L E R R L 736
 ACTCCCTTACAAATGAATGGCATCGAAGAGGATTCTGATGAGCCTTTAGAGAGAAGGCTG

2341 S L V P D S E Q G E A I L P R I S V I S 756
 TCCTTAGTACCAGATTCTGAGCAGGGAGAGGCGATACTGCCTCGCATCAGCGTGATCAGC

2401 T G P T L Q A R R R Q S V L N L M T H S 776
 ACTGGCCCCACGCTTCAGGCACGAAGGAGGAGTCTGTCCTGAACCTGATGACACACTCA

2461 V N Q G Q N I H R K T A S T R K V S L 796
 GTTAACCAAGGTCAGAACATTACCGAAAGACAACAGCATCCACACGAAAAGTCTCACTG

2521 A P Q A N L T E L D I Y S R R L S Q E T 816
 GCCCCTCAGGCAAACTTGACTGAAGTGGATATATATTCAAGAAGGTTATCTCAAGAACT

FIG.1 (cont'd)

2581 G L E I S E E I N E E D L N E C L F D D 836
 GGCTTGAAATAAGTGAAGAAATTAACGAAGAAGACTTAAAGGAGTGCCTTTTGTATGAT

2641 M E S I P A V T T W N T Y L R Y I T V H 856
 ATGGAGAGCATACCAGCAGTACTACATGGAACACATACCTTCGATATATTACTGTCCAC

2701 K S L I F V L L I W C L V I F L A E V A A 876
 AAGAGCTTAATTTTTGTGCTAATTTGGTGCTTAGTAATTTTTCTGGCAGAGGTGGCTGCT

2761 S L V V L W L L G N T P L O D K G N S T 896
 TCTTTGGTTGTGCTGTGGCTCCTTGAAACACTCCTCTTCAAGACAAAGGGAATAGTACT

2821 H S R N N S Y A V I I T S T S S Y Y V F 916
 CATAGTAGAAATAACAGCTATGCAGTGATTATCACCAGCACCAGTTCGTATTATGTGTTT

2881 Y I Y V G V A D T L L A M G F E R G L P 936
 TACATTTACGTGGGAGTAGCCGACACTTGTCTTGTATGGGATTCTTCAGAGGTCTACCA

2941 L V H T L I T V S K I L H H K M L H S V 956
 CTGGTGCATACTCTAATCACAGTGTGAAAATTTTACACCACAAAATGTTACATTCTGTT

3001 L O A P M S T L N T L K A E G I L N R F 976
 CTTCAAGCACCTATGTCAACCCTCAACAGCTGAAAGCAGTGGGATTCTTAATAGATTC

3061 S K D I A I L D D L L P L T I F D F I O 996
 TCCAAAGATATAGCAATTTGGATGACCTTCTGCCTCTTACCATATTTGACTTCATCCAG

3121 L L L I V I G A I A V V A V L O P Y I F 1016
 TTGTTATTAATTTGATTGGAGCTATAGCAGTTGTCCAGTTTACAACCCTACATCTTT

3181 V A T V P V I V A F I M L R A Y F L O T 1036
 GTTGCAACAGTGCCAGTGATAGTGGCTTTTATTATGTGAGAGCATATTTCTCCAAACC

3241 S O O L K O L E S E G R S P I F T H L V 1056
 TCACAGCAACTCAAACAACCTGGAATCTGAAGGCAGGAGTCCAATTTTCACTCATCTTGT

3301 T S L K G L W T L R A F G R O P Y F E T 1076
 ACAAGCTTAAAAGGACTATGGACACTTCGTGCCTTCGGACGGCAGCCTTACTTTGAAACT

3361 L F H K A L N L H T A N W F L Y L S T L 1096
 CTGTTCCACAAGCTCTGAATTTACATACTGCCAAGTGGTCTTGTACCTGTCAACACTG

3421 R W F O M R I E M I F V I F F I A V T F 1116
 CGCTGGTTCCAAATGAGAATAGAAATGATTTTTGTATCTTCTTATTGCTGTTACCTC

3481 I S I L T T G E G E G R V G I I L T L A 1136
 ATTTCCATTTTAAACAACAGGAGAAGGAGAAGGAAGAGTTGCTATTATCCTGACTTTAGCC

3541 M N I M S T L O W A V N S I D V D S I 1156
 ATGAATATCATGAGTACATTGCAGTGGGCTGTAAACTCCAGCATAGATGTGGATAGCTT

3601 M R S V S R V F K I D M P T E G K P T 1176
 ATGCGATCTGTGAGCCGAGTCTTTAAGTTCATTGACATGCCAACAGAAGGTAACCTACC

3661 K S T K P Y K N G Q L S K V H I I E N S 1196
 AAGTCAACCAAACCATACAAGAATGGCCAACTCTCGAAAGTTATGATTATTGAGAATTCA

3721 H V K K D D I W P S G G O M T V K D L T 1216
 CACGTGAAGAAAGATGACATCTGGCCCTCAGGGGGCCAAATGACTGTCAAAGATCTCACA

3781 A K Y T E G G N A I L E N I S F S I S P 1236
 GCAAAATACACAGAAGGTGGAAATGCCATATTAGAGAACATTTCTTCTCAATAAGTCTT

3841 G O R V G L L G R T G S G K S T L L S A 1256
 GGCCAGAGGTGGGCTCTTGGGAAGAACTGGATCAGGGAAGAGTACTTTGTTATCAGCT

FIG.1.(cont'd)

F L R L L N T E G E I O I D G V S W D S 1276
 3901 TTTTGAGACTACTGAACACTGAAGGAGAAATCCAGATCGATGGTGTGTCTTGGGATTCA

I T L O Q W R K A F G V I P O K V F I E 1296
 3961 ATAAC TTTGCAACAGTGGAGGAAAGCCTTTGGAGTGATACCACAGAAAGTATTTATTTT

S G T F R K N L D P Y E Q W S D Q E I W 1316
 4021 TCTGGAACATTTAGAAAAA ACTTGGATCCCTATGAACAGTGGAGTGATCAAGAAATATGG

K V A D E V G L R S V I E Q F P G K L D 1336
 4081 AAAGTTGCAGATGAGTTGGGCTCAGATCTGTGATAGAACAGTTTCTGGGAAGCTTGAC

F V L V D G G C V L S E G H K O L M C L 1356
 4141 TTTGTCCTTGTGGATGGGGGCTGTGTCTAAGCCATGGCCACAAGCAGTTGATGTGCTTG

A R S V L S K A K I L L L D E P S A H L 1376
 4201 GCTAGATCTGTTCTCAGTAAGGCGAAGATCTTGCTGCTTGATGAACCCAGTGCTCATTG

D P V T Y Q I I R R T L K O A F A D C T 1396
 4261 GATCCAGTACATACCAATAATTAGAAGA ACTCTAAAACAAGCATTGCTGATTGCACA

V I L C E H R I E A M L E C O O F L V I 1416
 4321 GTAATTCTCTGTGAACACAGGATAGAAGCAATGCTGGAATGCCAACAAATTTTGTGTCATA

E E N K V R O Y D S I O K L L N E R S L 1436
 4381 GAAGAGAACAAAGTGGGGCAGTACGATTCCATCCAGAAACTGCTGAACGAGAGGAGCCTC

F R O A I S P S D R V K L F P H R N S S 1456
 4441 TTCCGGCAAGCCATCAGCCCTCCGACAGGGTGAAGCTCTTCCCCACGGAACTCAAGC

K C K S K P Q I A A L K E E T E E E V O 1476
 4501 AAGTGCAAGTCTAAGCCCAGATTGCTGCTCTGAAAGAGGAGACAGAAGAAGAGGTGCAA

D T R L = 1480
 4561 GATACAAGGCTTTAGAGAGCAGCATAAATCTTGACATGGGACATTGCTCATGGAATTGG
 4621 AGCTCGTGGGACAGTCACTCATGGAATTGGAGCTCGTGGAACAGTTACCTCTGCCTCAG
 4681 AAAACAAGGATGAATTAAGTTTTTTTTAAAAAGAAACATTTGGTAAGGGGAATTGAGG
 4741 ACACTGATGGGTCTTGATAAATGGCTTCCTGGCAATAGTCAATGTGTGAAAGGTAC
 4801 TTCAAATCCTTGAAGATTTACCACCTGTGTTTTGCAAGCCAGATTTCTTGAAAACCTT
 4861 GCCATGTGCTAGTAATTGGAAGGCAGCTCTAAATGTCAATCAGCCTAGTTGATCAGCTT
 4921 ATTTGCTAGTGAAACTCGTTAATTTGTAGTGTGAGAGAAGACTGAAATCATACTTCTTA
 4981 GGGTTATGATTAAGTAATGATAACTGGAACTTCAGCGGTTTATATAAGCTTGTATTCTT
 5041 TTTTCTCTCTCCTCCCATGATGTTTAGAAACACA ACTATATTGTTGCTAAGCATTCCA
 5101 ACTATCTCATTTCGAAGCAAGTATTAGAATACCACAGGAACCACAGACTGCACATCAA
 5161 ATATGCCCCATTCAACATCTAGTGAGCAGTCAGGAAAGAGA ACTTCCAGATCCTGGAAT
 5221 CAGGGTTAGTATTGTCCAGGCTACCAAAAATCTCAATATTTTCAGATAATCACAATACAT
 5281 CCCTTACCTGGGAAAGGGCTGTTATAATCTTTCACAGGGGACAGGATGGTTCCTTGATG
 5341 AAGAAGTTGATATGCCTTTTCCCAACTCCAGAAAGTGACAAGCTCACAGACCTTTGAAT
 5401 AGAGTTTAGCTGGAAAAGTATGTTAGTGCAAAATGTCACAGGACAGCCCTTCTTCCACA
 5461 GAAGCTCCAGGTAGAGGGTGTGTAAGTAGATAGGCCATGGGCACTGTGGGTAGACACACA
 5521 TGAAGTCCAAGCATTTAGATGTATAGTTGATGGTGGTATGTTTTAGGCTAGATGTATG
 5581 TACTTCATGCTGTCTACACTAAGAGAGAATGAGAGACACACTGAAGAAGCACC AATCATG
 5641 AATTAGTTTTATATGCTTCTGTTTTATAATTTGTGAAGCAAAATTTTCTCTAGGAAA
 5701 TATTTATTTAATAATGTTTCAACATATATTACAATGCTGTATTTAAAAGAATGATTA
 5761 TGAATTACATTTGTATAAATAATTTTATATTTGAAATATTGACTTTTATGGCACTAG
 5821 TATTTTATGAAATATTATGTTAAA ACTGGGACAGGGGAGAACCTAGGGTGATATTAACC
 5881 AGGGCCATGAATCACCTTTTGGTCTGGACGGAAGCCTTGGGGCTGATCGAGTTGTTGCC
 5941 CACAGCTGATGATTTCCAGCCAGCACAGCCTCTTAGATCCACTCTGAAGAAGATGTT
 6001 ACCACCAGTCTGACTGTTTCCATCAAGGGTACTGCTTCTCAACTCCAAACTGACTCT
 6061 TAAGAAGACTGCATTATATTTACTGTAAGAAAATATCACTTGTCAATAAATCCATA
 6121 CATTGTGT (A) n

FIG.2.

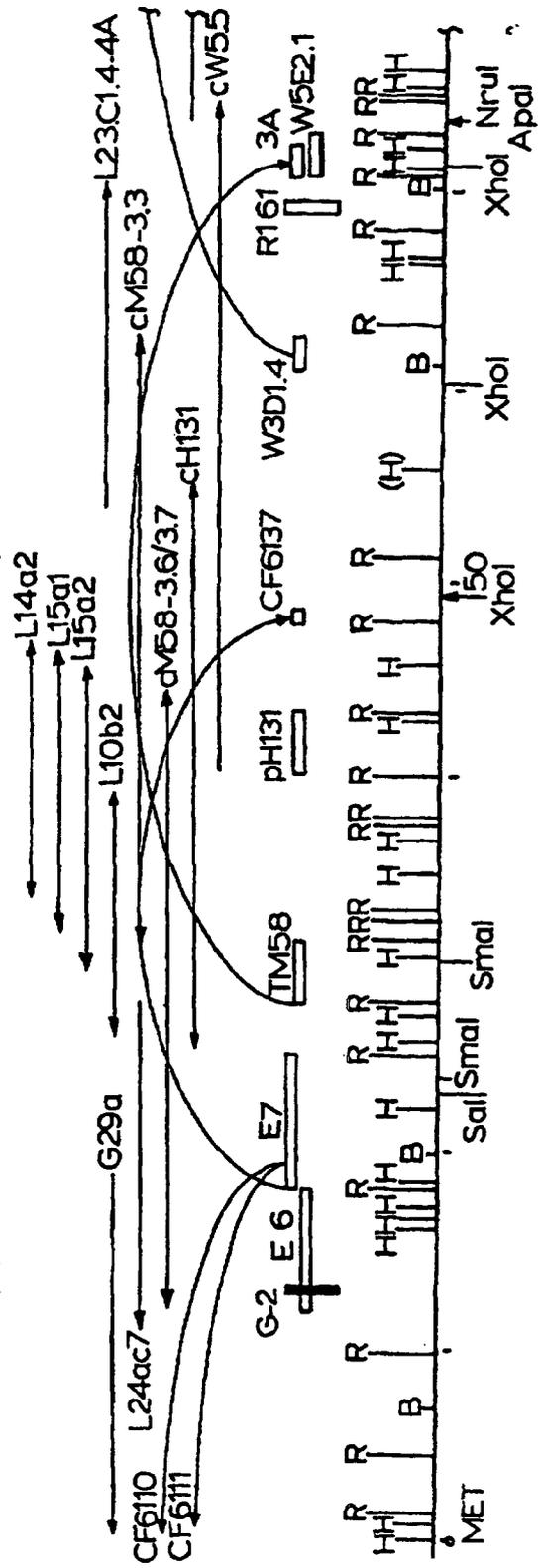
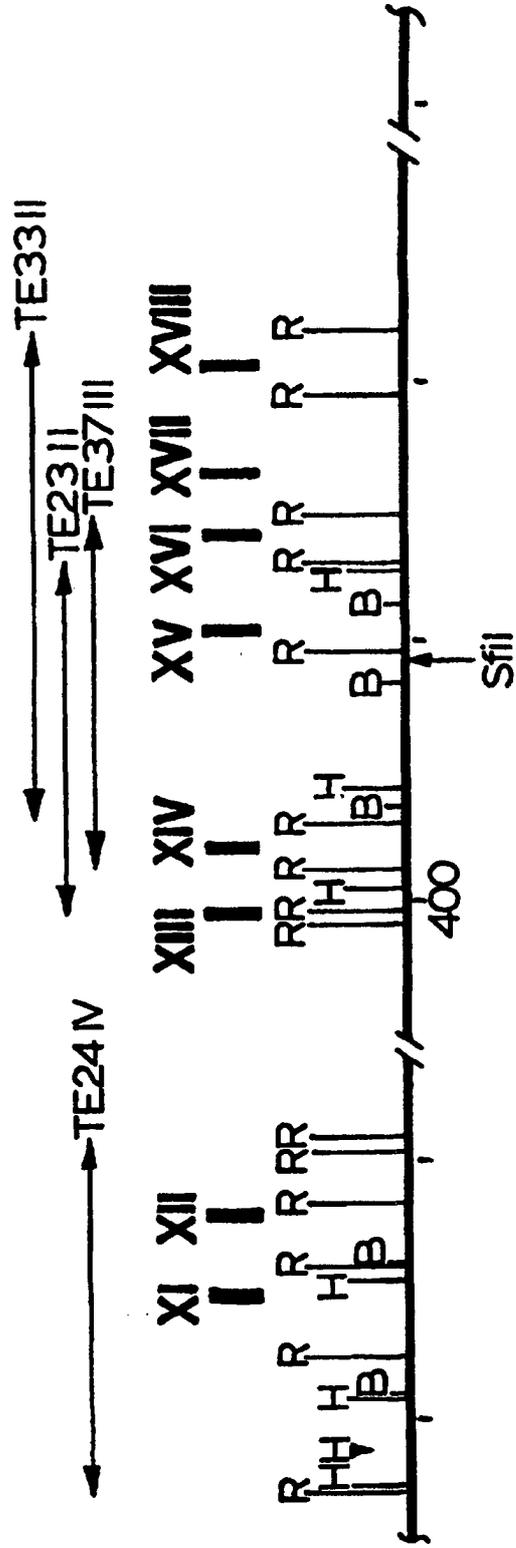


FIG.2 (cont'd)



2

FIG. 3

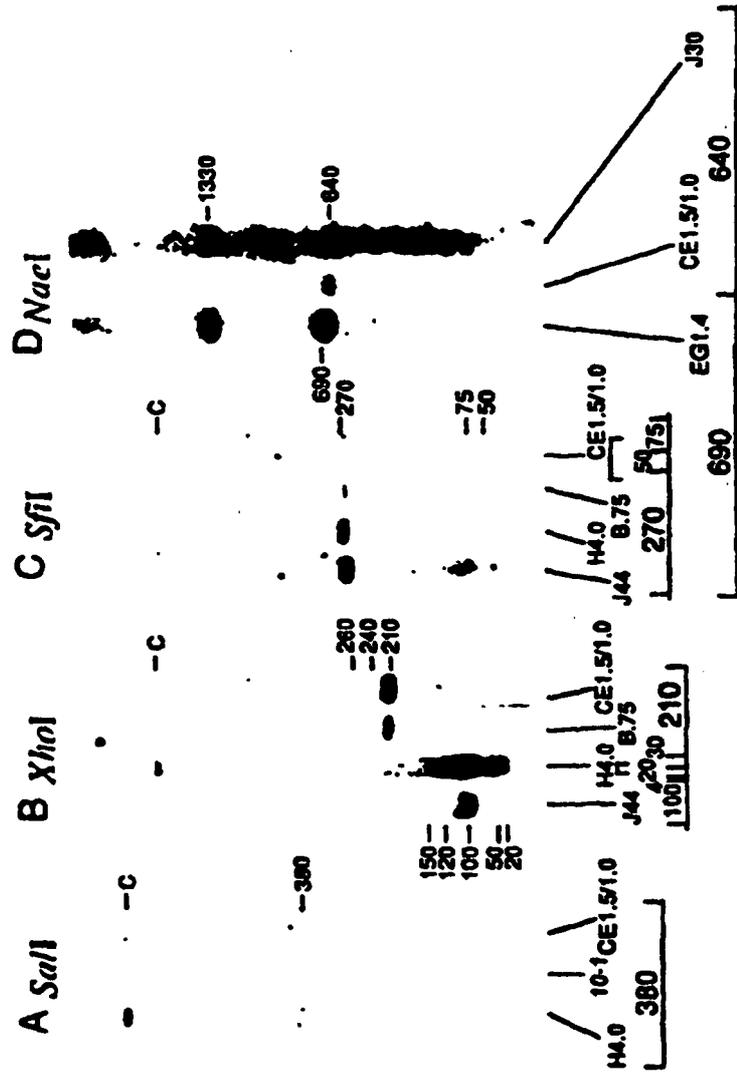


FIG. 4

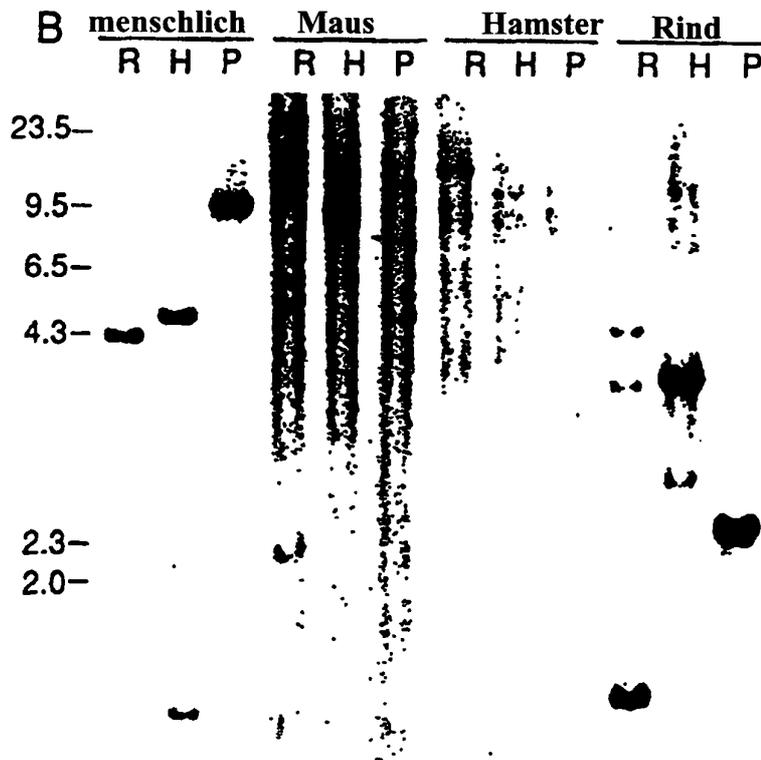
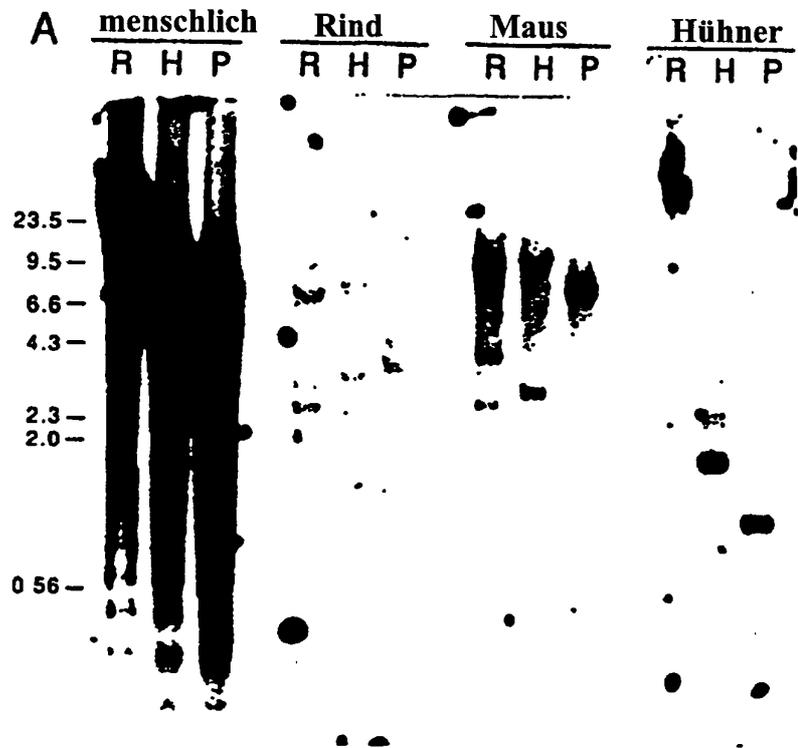


FIG. 4

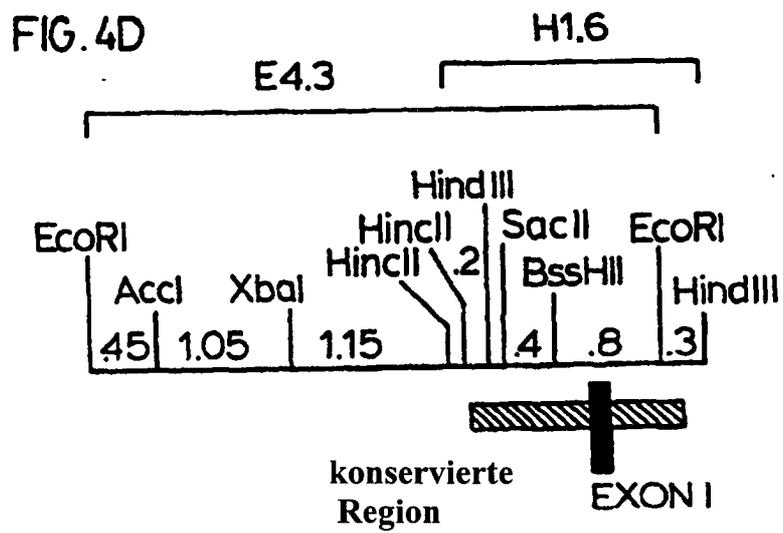
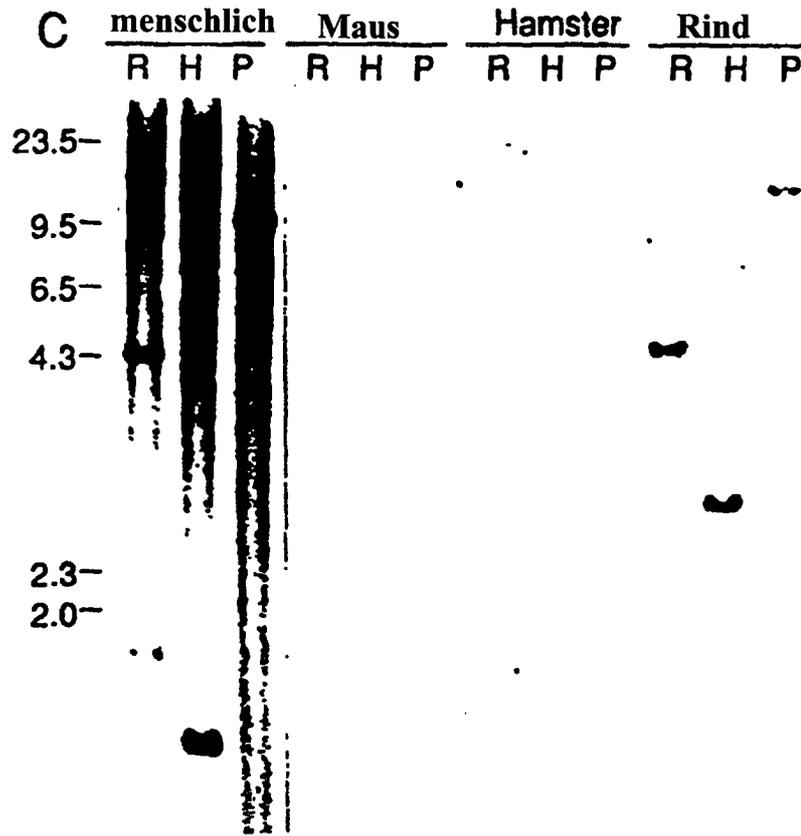


Fig. 5

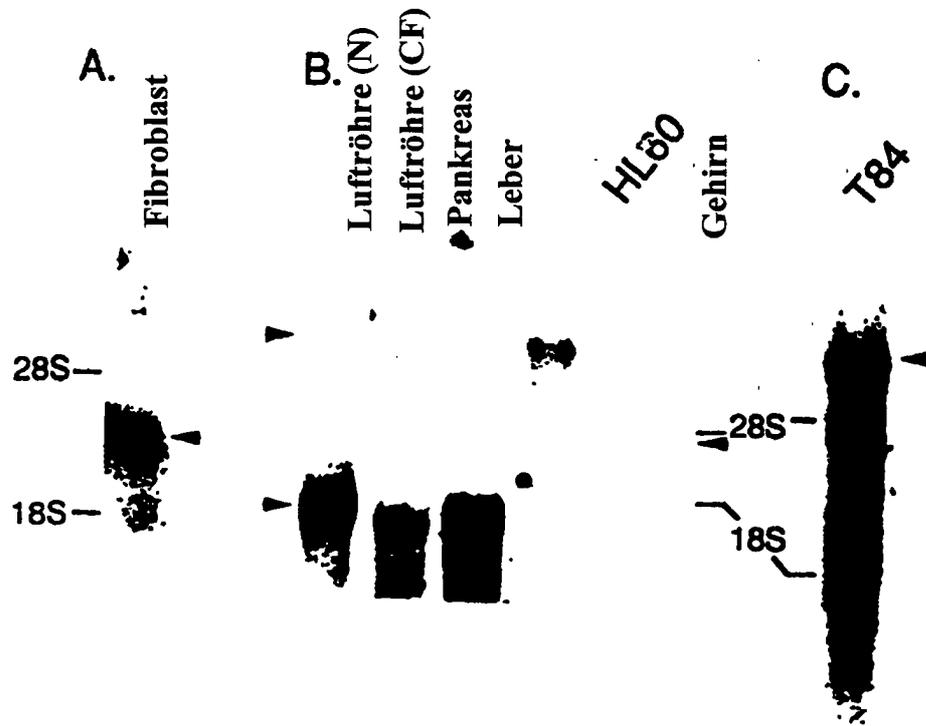


FIG. 6

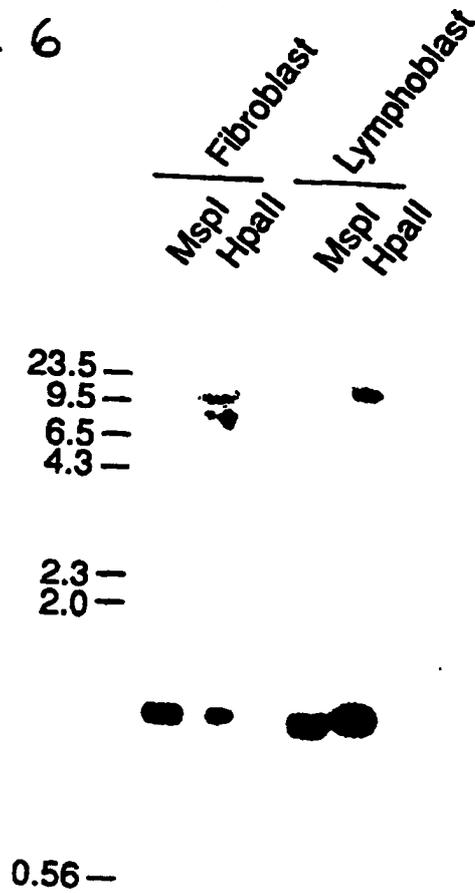


FIG. 7.

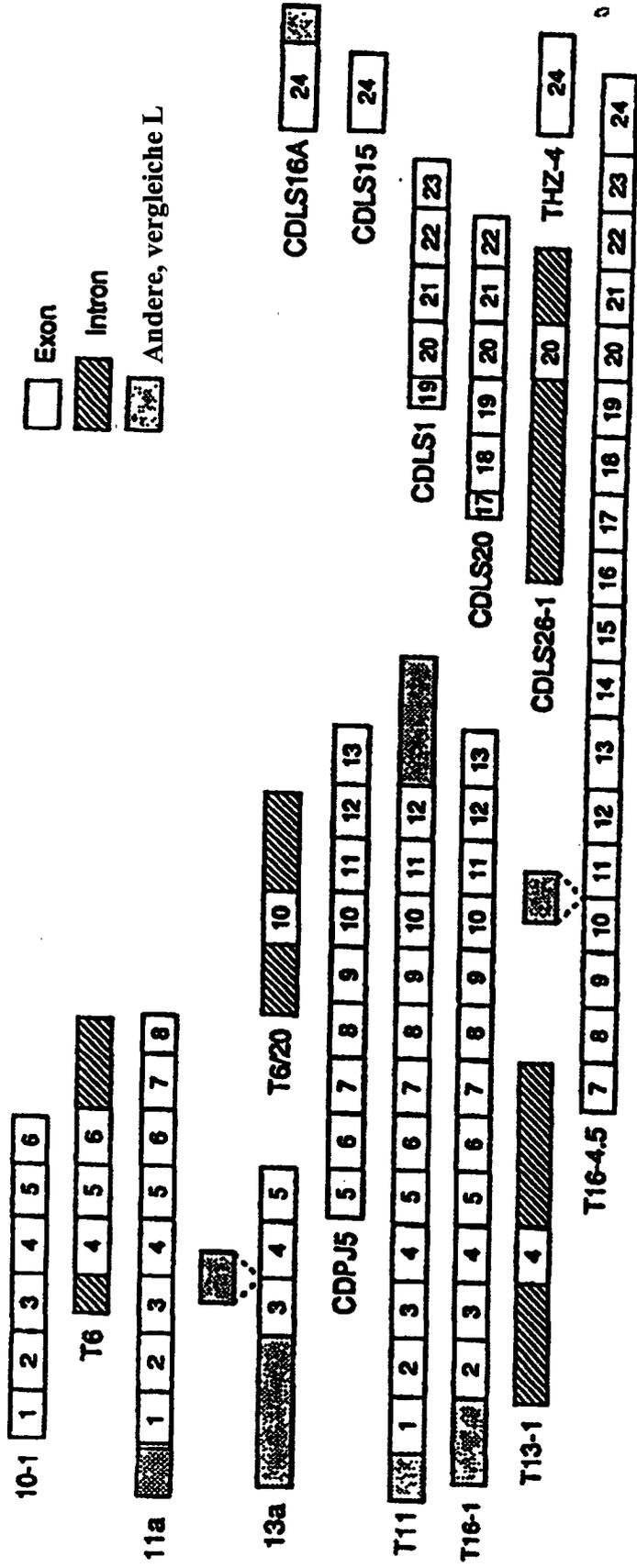
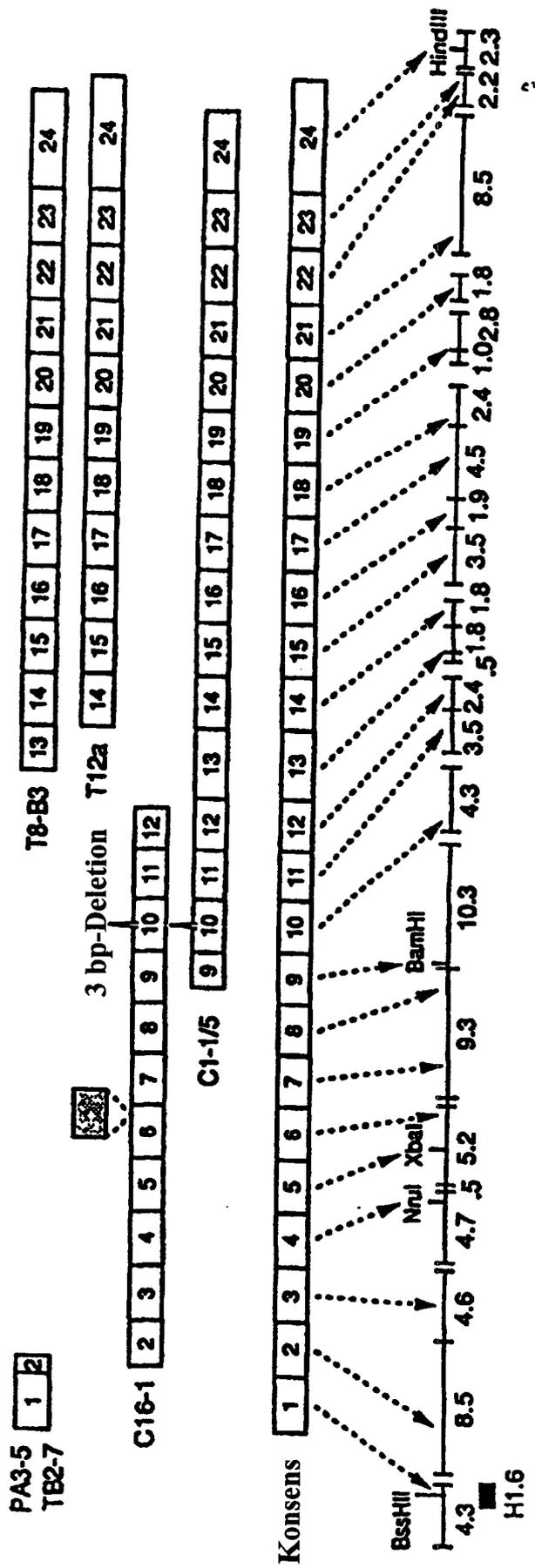


FIG 7. (cont'd)



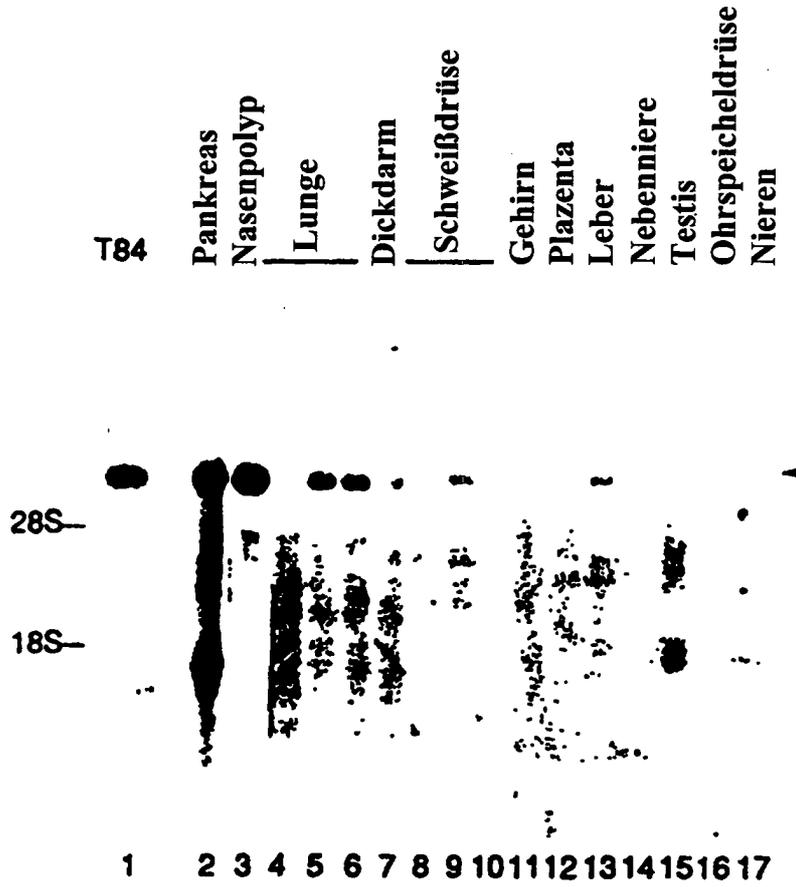


Fig. 8

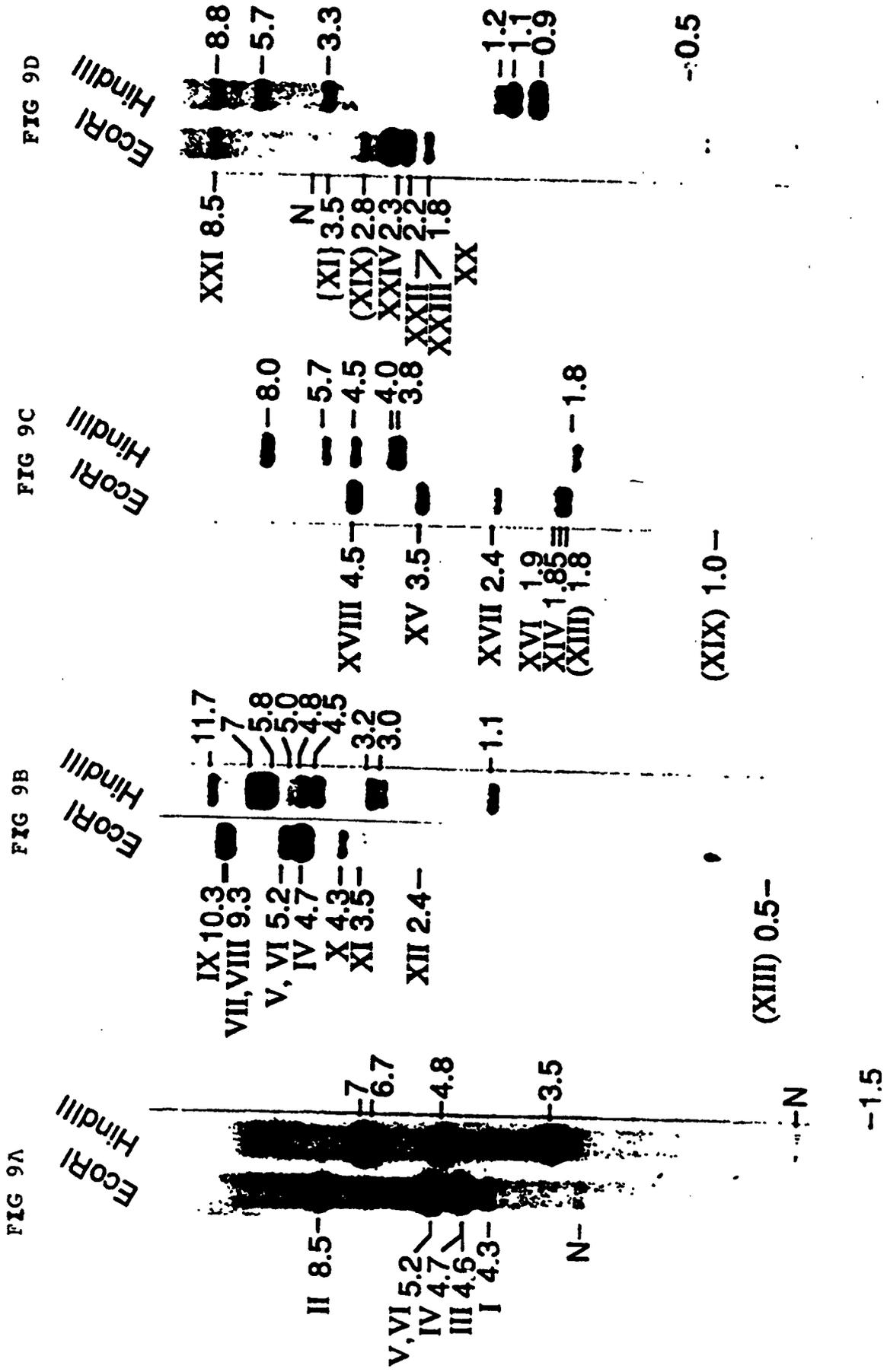


FIG. 10A

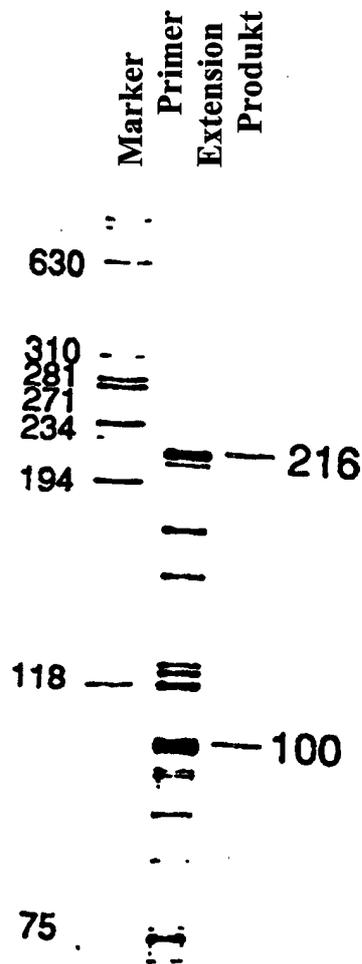


FIG. 10 B

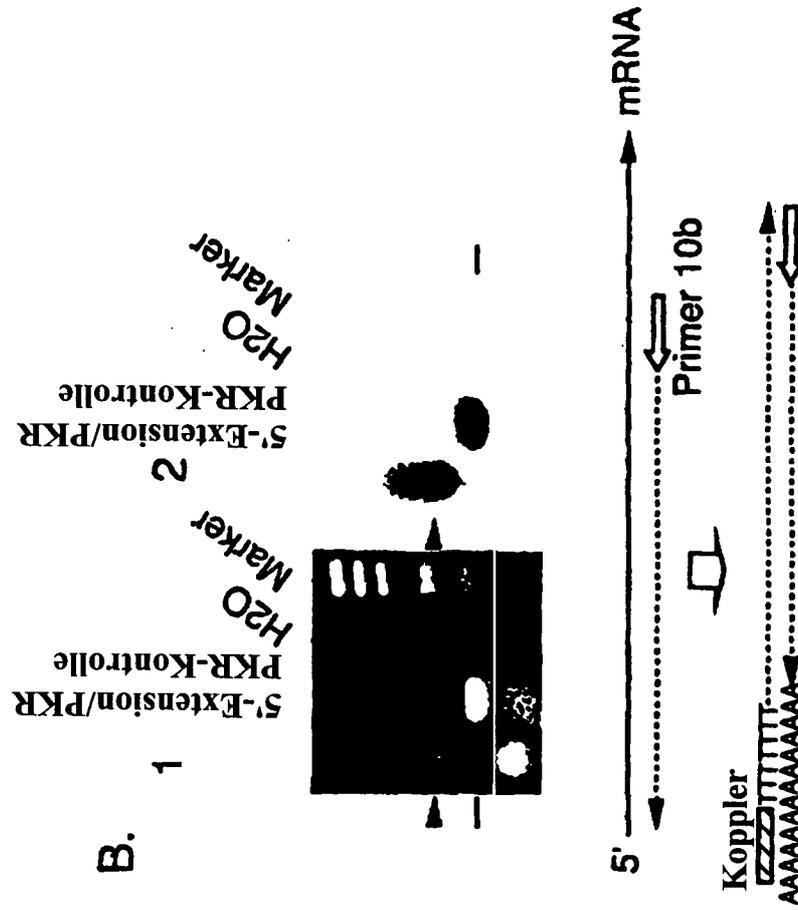


FIG.11.

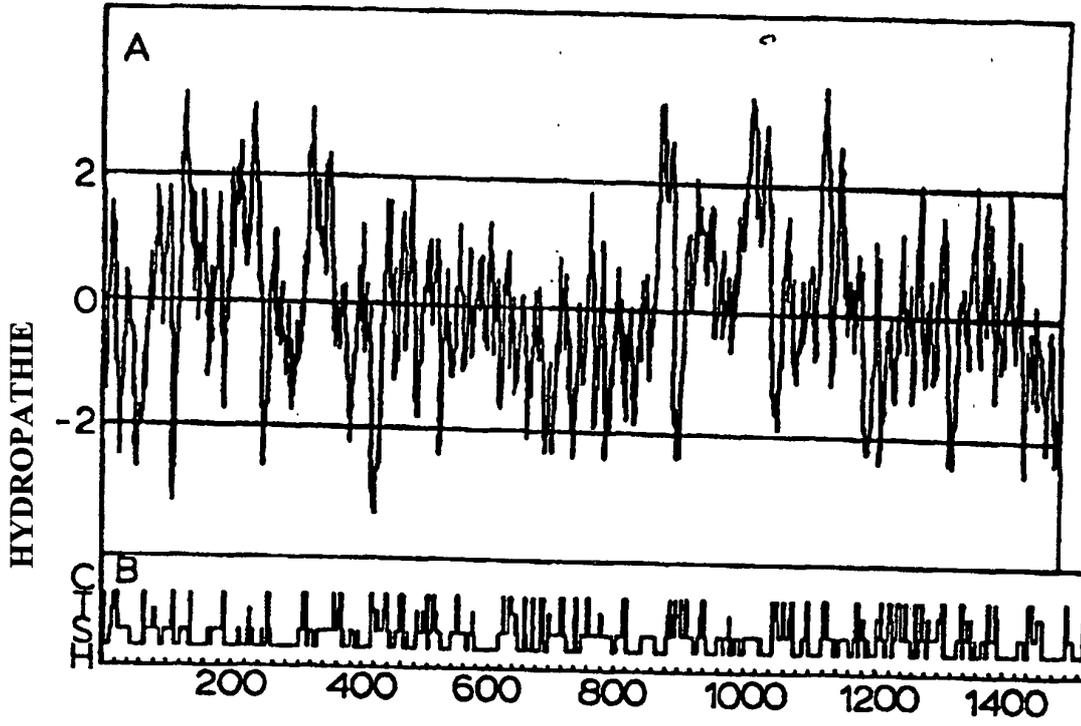


FIG. 12

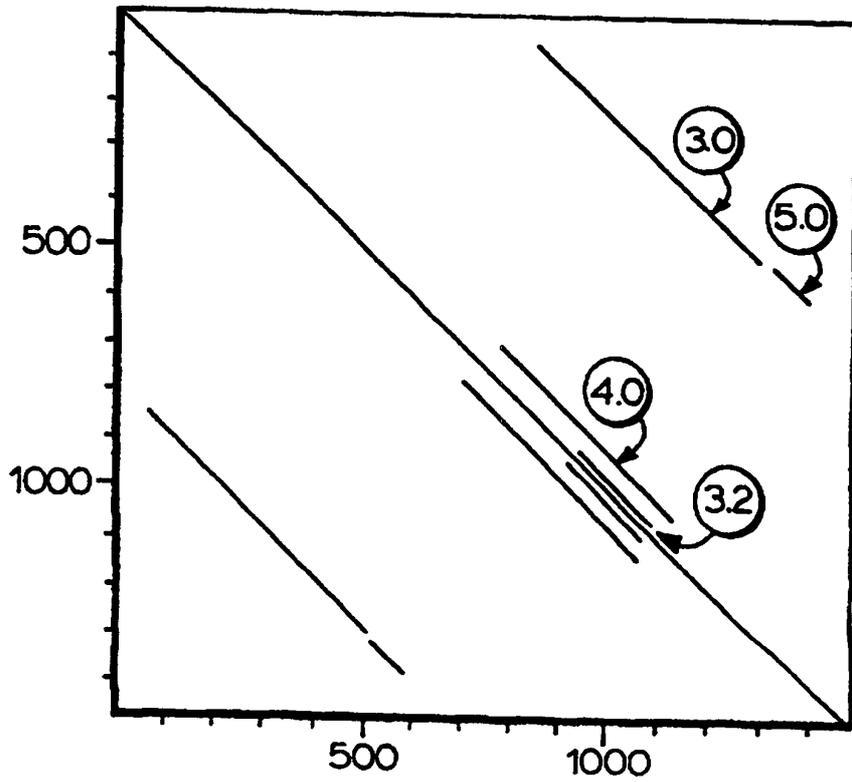


FIG. 13

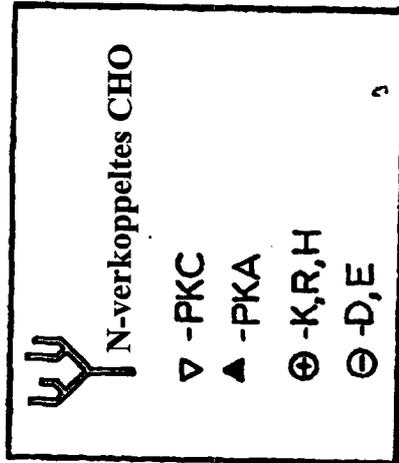
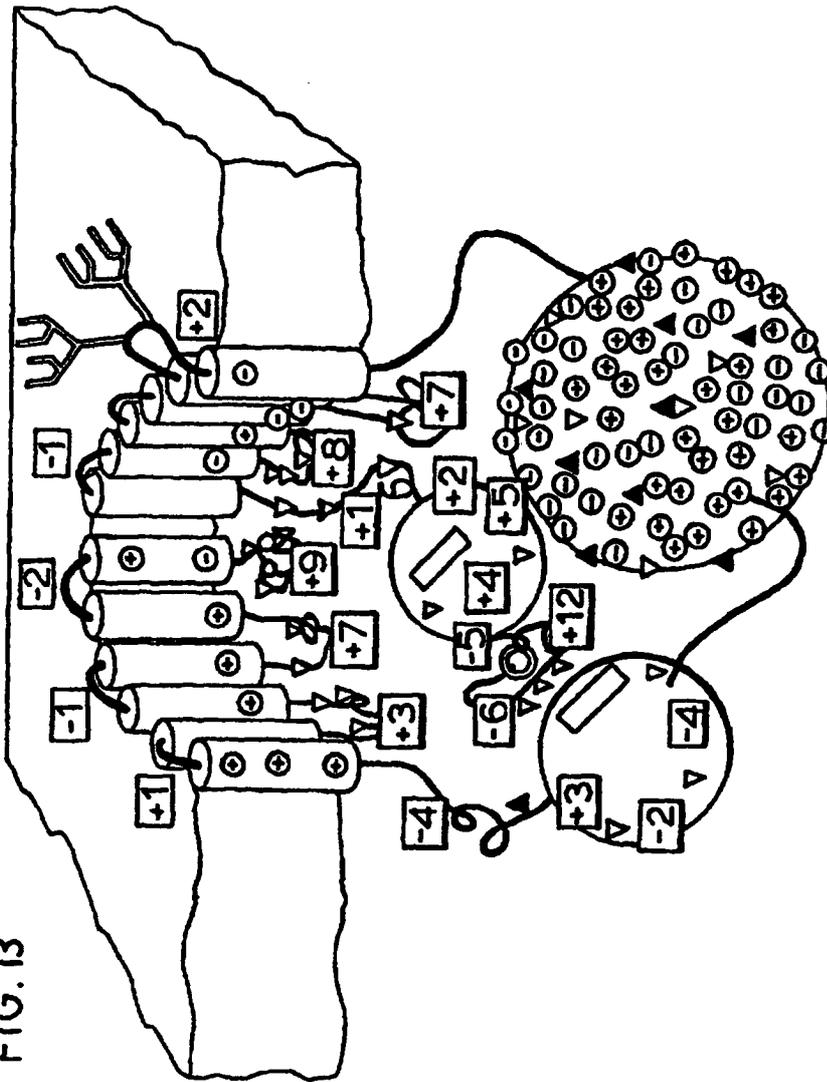
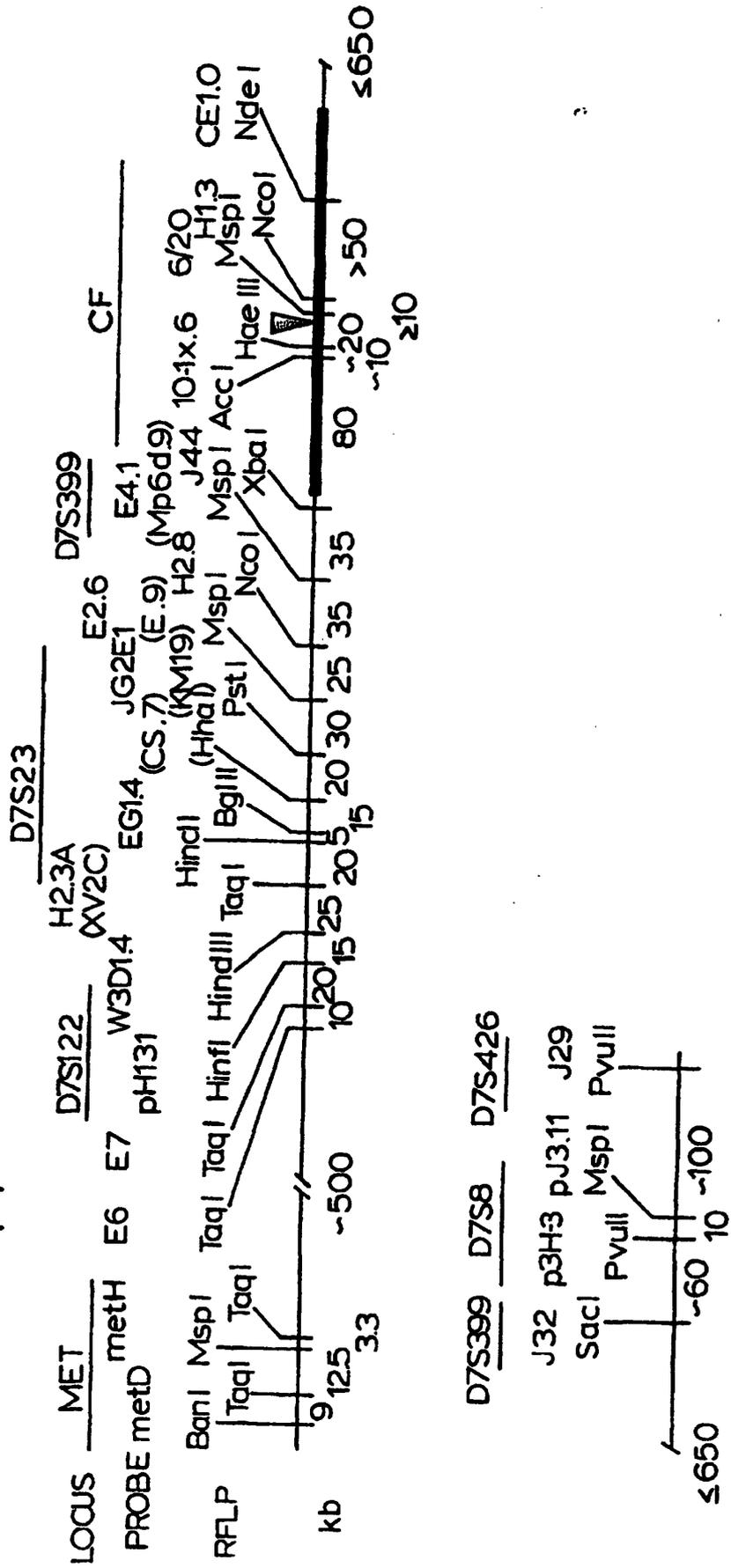


FIG.14.



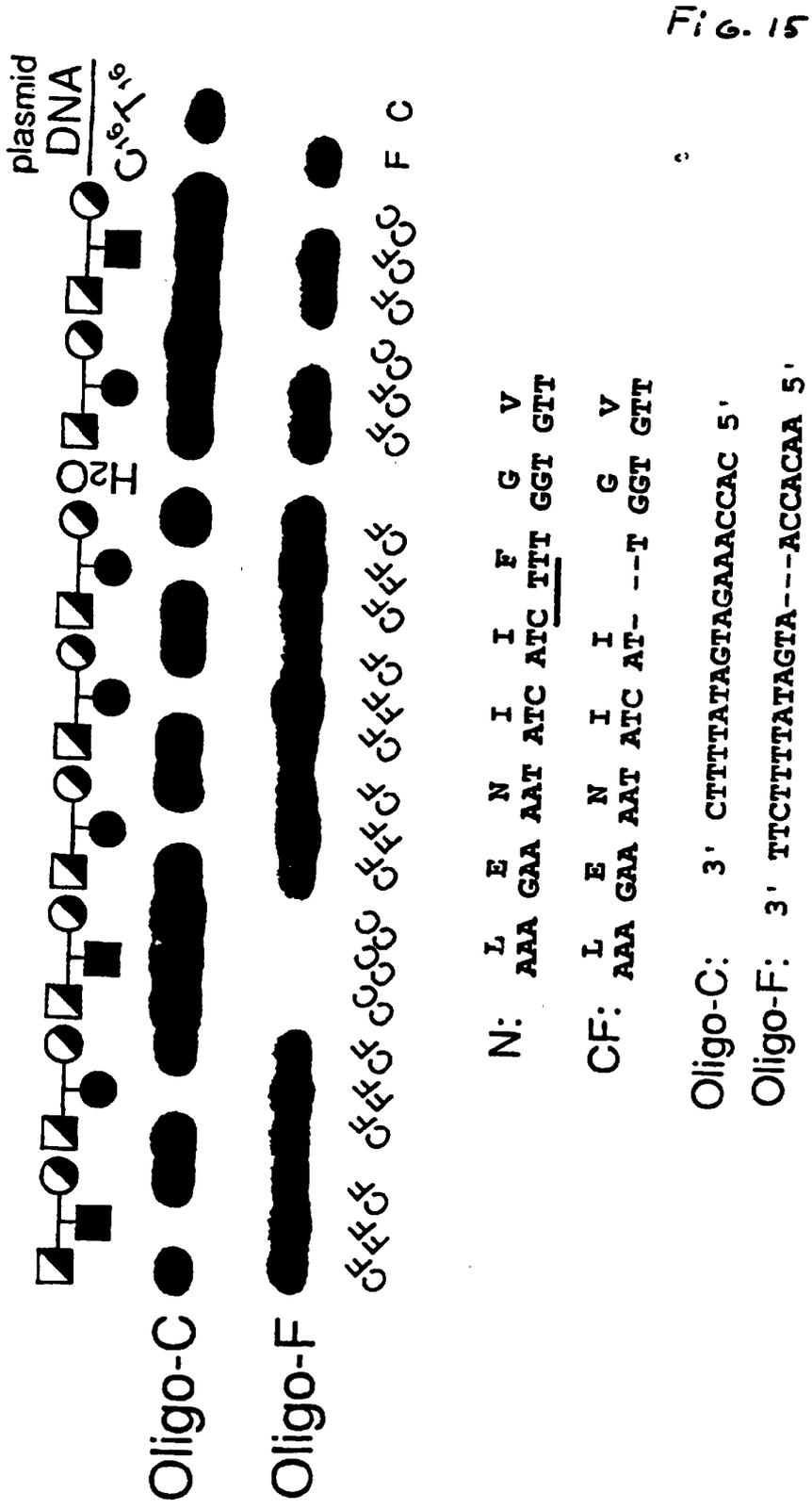


Fig. 15

FIG.16. (cont'd)

CFTR (N)	■■■■■	GEGITLSSGGQRARISLARAVYKDADLYLLDSPFFGYLDVLTEK
CFTR (C)		VDGGCVLSHGKQLMCLARSVLSKAKILLLDEPSAHLDPVTYQ
hmdr1 (N)		GERGAQLSSGGQKQRIARALVRNPKILLDEATSALDTESEA
hmdr1 (C)		GDKGTLSSGGQKQRIARALVRQPHILLDEATSALDTESEK
mmdr1 (N)		GERGAQLSSGGQKQRIARALVRNPKILLDEATSALDTESEA
mmdr1 (C)		GDKGTQLSSGGQKQRIARALVRQPHILLDEATSALDTESEK
mmdr2 (N)		GDRGAQLSSGGQKQRIARALVRNPKILLDEATSALDTESEA
mmdr2 (C)		GDKGTQLSSGGQKQRIARALVRNPKILLDEATSALDTESEK
pfmdr (N)		GSNASKLSSGGQKQRIARAIMRNPKILLDEATSALDTESEK
pfmdr (C)		PYGKS - LSSGGQKQRIARALREP KILLDEATSALDTESEK
STE6 (N)		GTGGVTLSSGGQKQRIARAFIRDTPILFLDEAVSALDIVHRN
STE6 (C)		RIDTTLSSGGQKQRIARALRKS KILLDEATSALDTESEK
hlyB		GEQAGLSSGGQKQRIARALYNNPKILLDEATSALDTESEK
White		PGRVKGLSSGGERKRLAFASEALDTPPLLICDEPTSGLDSFTA
MbpX		FYPAQLSSGGQKQRIARSLAIQPDLLL - DEFFGALDDELRR
Btud		GRSTNQLSSGGEWQRVLAAVLQITLILLDEPMNSLDVAQOSA
PatB		HQSGYLSGGQKQRIARALRNP KILLDEATSALDTESEK
hisp		GKYPVHLSGGQKQRIARALAMEPDLVLFDEPTSALDPELVG
malK		DRKPKALSSGGQKQRIARALRNP KILLDEATSALDTESEK
oppD		KMYPHEFSGGMRQRMIAAMALLCRPKLLIADDEPTTALDVTVOA
oppF		NRYPHFSGGQKQRIARALRNP KILLDEATSALDTESEK
RbsA (N)		DKLVGDLSIGDQQMVEIAKVLSEKVIIMDEPTCALIDTETE
RbsA (C)		EQAILLSSGGNQQKVAIARGLMTRPKVILLDEPTGVDVGA
UvrA		GQSATTLSSGGEAQRVKLARELSKRGVYLLDEPTTGLHFADIQ
NodI		NTRVADLSSGGMKRLTLAGALINDPQLLILDEPTTGLDPHARH
FtsE		KNFPIQLSSGGEQQRVGIARAVVKNKPAVLLADEPTGNLDDALSE

FIG. 17

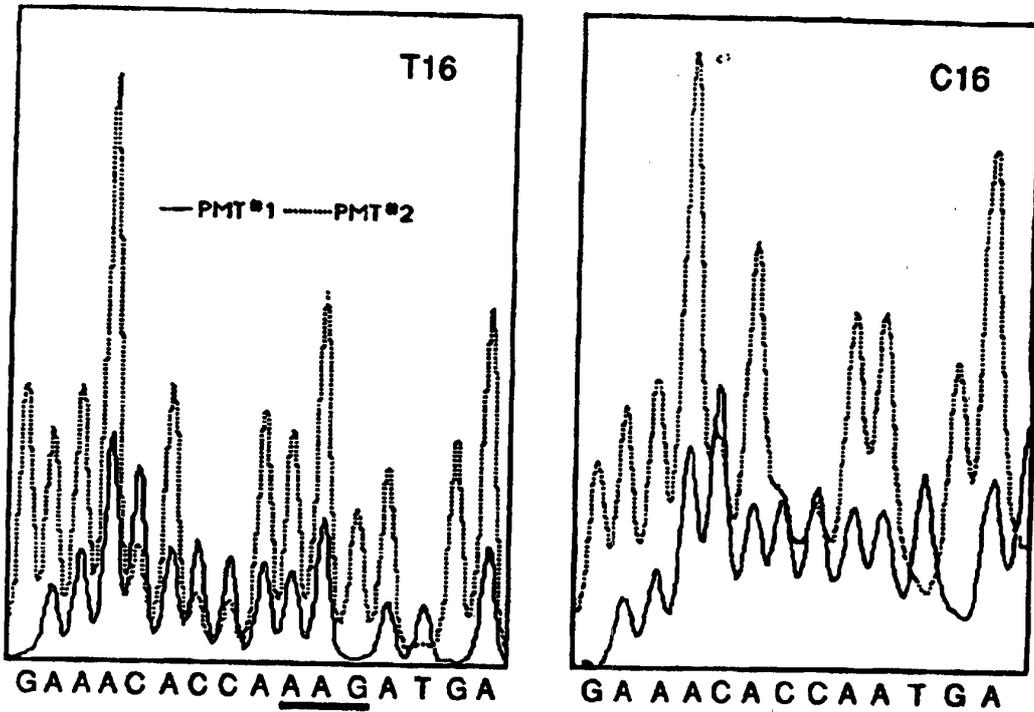


FIG. 18 A

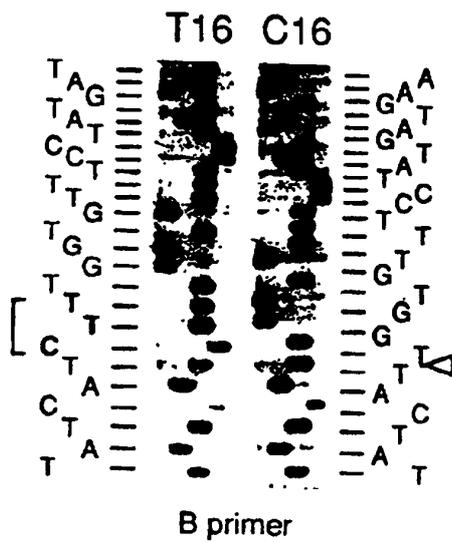


FIG. 18 B

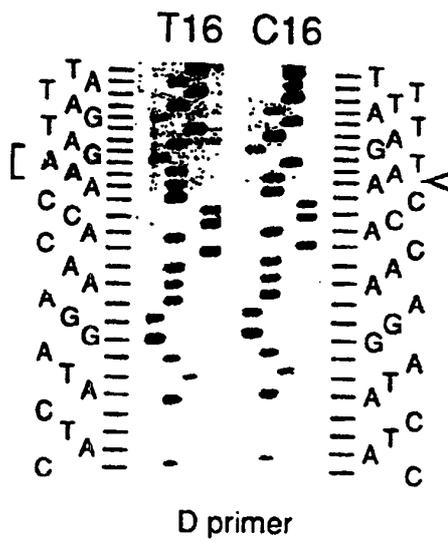


FIG. 19A

Mr 1 2 3 4 5 6

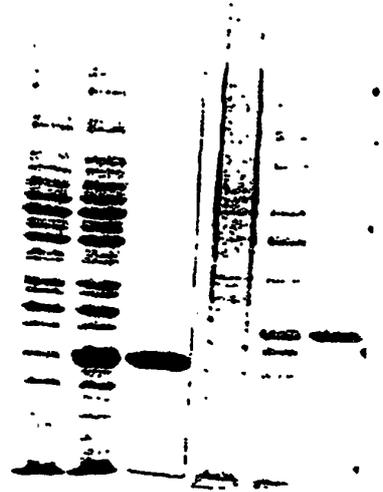
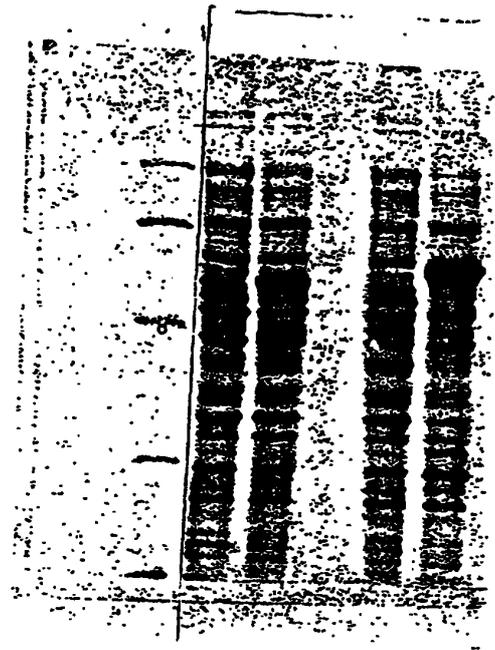


FIG. 19B

1 2 3 4

97 -
66 -
42 -
31 -



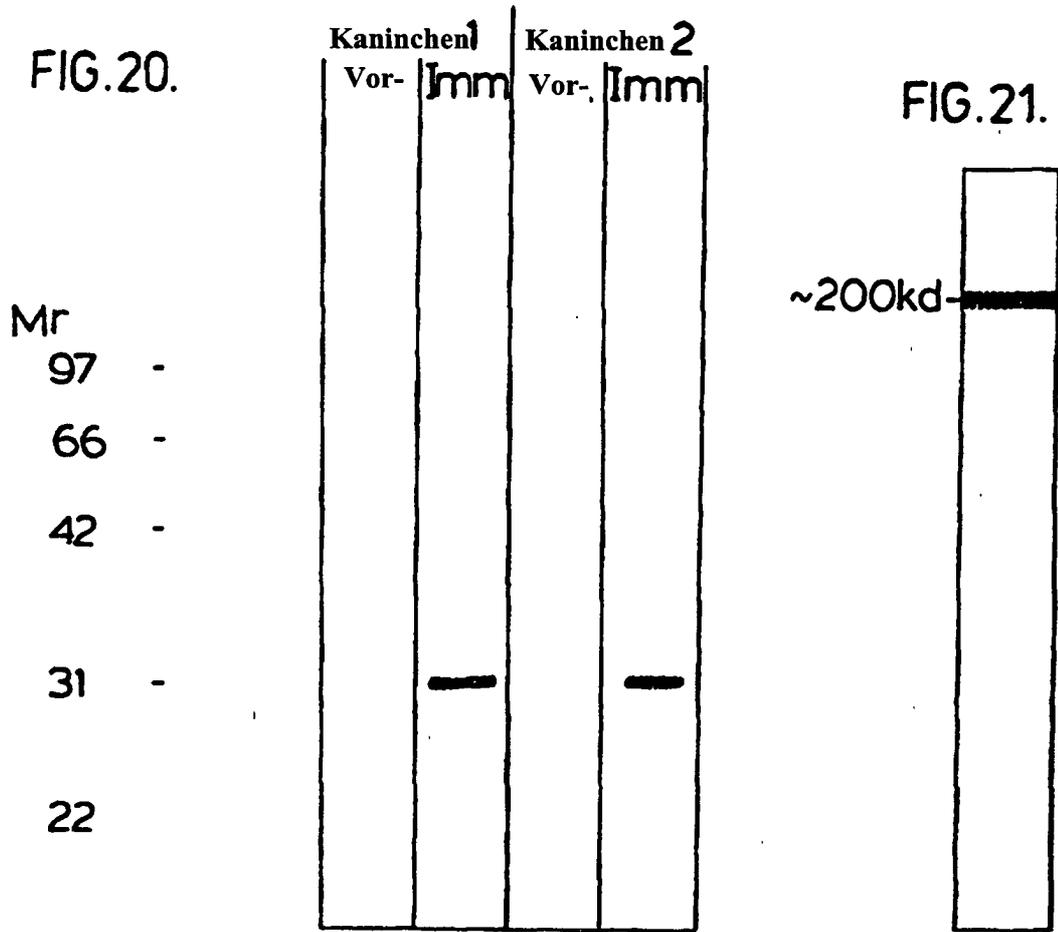


FIG.22.

Vorimmun			Immun		
ng			ng		
100			100	•	•
10			10	•	•
1			1		•
0.1			0.1		
	H	P ₁ P ₂		H	P ₁ P ₂