

Aus der Abteilung Pädiatrische Pneumologie und Neonatologie
des Zentrums für Kinder- und Jugendmedizin, Humangenetik und Dermatologie
der Medizinischen Hochschule Hannover:
angefertigt im Rahmen der strukturierten Doktorandenausbildung StrucMed

**Analyse der *CLCA*-Genfamilie als
Modulator des Basisdefektes bei Patienten
mit Mukoviszidose**

Dissertation
zur Erlangung des Doktorgrades der Medizin
in der Medizinischen Hochschule Hannover

vorgelegt von
Ernst Wolfgang Kolbe
aus Göttingen

Hannover 2007

Angenommen vom Senat der Medizinischen Hochschule Hannover
am 26.03.2008

Gedruckt mit Genehmigung der Medizinischen Hochschule Hannover

Präsident:	Prof. Dr. Dieter Bitter-Suermann
Betreuer der Arbeit:	Prof. Dr. Dr. Burkhard Tümmler
Kobetreuer der Arbeit:	Prof. Dr. Hassan Y. Naim
Referent:	Prof. Dr. Tobias Welte
Korreferent:	Prof. Dr. Jörg Schmidtke

Tag der mündlichen Prüfung:	26.03.2008
-----------------------------	------------

Promotionsausschussmitglieder:	Prof. Dr. Thomas Werfel
	Prof. Dr. Georg Scheumann
	PD. Dr. Torsten Witte

Meinen Eltern

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Mukoviszidose – eine Erbkrankheit	1
1.2	Basisdefekt der Mukoviszidose	2
1.3	Hinweise für die genetische Modulation des klinischen Verlaufes bei der Mukoviszidose-Erkrankung	4
1.4	Hinweise für die Modulation des Basis-Defektes durch alternative Chloridionen-Kanäle	6
1.5	Die Gene der <i>CLCA</i> -Familie als potentielle Kandidaten für den Modulator des Basisdefektes bei Mukoviszidose-Patienten	6
1.5.1	Eigenschaften und Funktionen der <i>CLCA</i> -Proteinfamilie	7
1.6	Genetische Marker und deren Verwendung in Assoziationsstudien zur Identifikation von Kandidaten-Genen	10
1.7	Zielsetzung der Arbeit.....	13
2	Material und Methoden	14
2.1	Studienpopulationen	14
2.1.1	Patienten der Europäischen Mukoviszidose Zwillings- und Geschwisterstudie	14
2.1.2	Unverwandte und gesunde Kontroll-Probanden	14
2.2	Intestinale Ionenstromfluß-Messung – Klassifikation der gastrointestinalen Phänotypen	15
2.2.1	Prinzip der Messung	15
2.2.2	4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) als diskriminierendes Agens	16
2.3	Kriterien zur Auswahl der genetischen Marker	18
2.3.1	Größen zur Charakterisierung von genetischen Markern.....	18
2.3.2	Auswahlkriterien der verwendeten Marker	19
2.4	Verzeichnis der benötigten Reagenzien	20
2.5	Isolierung und Handhabung der DNA	21
2.5.1	DNA-Isolierung	21
2.5.2	Lagerung der DNA und Vorbereitung für die PCR.....	22
2.5.3	Parallele Analyse großer Probenzahlen.....	22
2.6	Typisierung der genetischen Marker.....	23
2.6.1	Theoretische Grundlagen der Polymerase Ketten Reaktion	23
2.6.2	Durchführung der PCR-Reaktion	24

2.6.3	Differenzierung der Allele	25
2.6.3.1	Allelunterscheidung durch Restriktionsverdau.....	25
2.6.3.1.1	Restriktionsverdau unaufgereinigter PCR-Produkte.....	27
2.6.3.1.2	Restriktionsverdau nach Ethanol-Fällung.....	27
2.6.3.1.3	Restriktionsverdau nach Adsorptionschromatographie.....	27
2.6.3.2	Allelspezifische PCR.....	27
2.6.4	Gelelektrophorese und Darstellung der PCR-Produkte	29
2.7	Rekonstruktion der Haplotypen	30
2.8	Konstruktion von stammesgeschichtlich fixierten Rekombinationsbruchpunkten	31
2.9	Charakterisierung der Patientenkollektive und Erläuterung der statistischen Auswertung der Assoziationsstudie.....	33
2.9.1	Charakterisierung der Patientengruppen	33
2.9.1.1	Patienten der Assoziationsstudie	33
2.9.1.2	Unverwandte und gesunde Kontroll-Probanden – Transkriptomanalyse.....	35
2.9.2	Monte-Carlo-Test zum Vergleich zweier Populationen bezüglich der Allelverteilung.....	37
2.10	DNA-Sequenzierung.....	38
2.10.1	PCR für die Sequenzierung.....	38
2.10.2	Prinzip der Sequenzier-Reaktion	39
2.10.3	Zusammensetzung der Sequenzdaten zu einem vollständigen Datensatz	40
2.10.4	Analyse der zusammengesetzten Sequenzdaten.....	41
2.11	Funktionelle Analyse mit der Genome-Atlas-Software	42
2.12	Vorhersage von Transkriptionsfaktor-Bindungsstellen.....	42
2.13	Abschätzung der Expression von Genen anhand von Transkriptomdaten aus Affymetrix-Chips.....	43
2.14	Abschätzung der Gen-Expression mittels der UniGene-Datenbank.....	44
2.15	Methoden zur Synthese und Testung von Antikörpern gegen das Homolog hCLCA4	45
2.15.1	Auswahl der Epitope zur Antikörper-Gewinnung	45
2.15.2	Immunisierung der Kaninchen	48
2.15.3	Immuno-Dotblot – Kontrolle der Immunisierung.....	48
2.15.4	Auftragen der für die Immunisierung verwendeten Peptide auf eine Membran	48
2.15.5	Durchführung der Immundetektion	48

3	Ergebnisse	50
3.1	Charakterisierung der typisierten Marker	50
3.1.1	Lokalisation der Marker im <i>CLCA</i> -Gen-Cluster	50
3.1.2	Vergleich der veröffentlichten und gemessenen Allel-Frequenzen.....	51
3.1.3	Informativität der Marker	53
3.1.4	Frequenzen der Genotypen – Hardy-Weinberg-Gleichgewicht	54
3.1.5	Allelfrequenzen und Informativität der etablierten Marker.....	56
3.2	Ergebnisse der Assoziationsstudie: Vergleich von CF-Patienten mit alternativer Chloridionenleitfähigkeit mit Patienten ohne Chloridionenleitfähigkeit im <i>CLCA</i>-Gen-Cluster	57
3.2.1	Auswertung der Assoziationsstudie: Vergleich der CF-Patienten an einzelnen Loci.....	58
3.2.2	Auswertung der Assoziationsstudie: Vergleich der Verteilung von Zwei-Marker Haplotypen zwischen den CF-Patienten	60
3.2.2.1	Gesamtes Studienkollektiv: Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den Gruppen der CF-Patienten mit residueller Chloridionenleitfähigkeit und ohne Chloridionenleitfähigkeit.....	61
3.2.2.2	Altersstratifizierte Gruppen: Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den Patienten mit alternativer Chloridionenleitfähigkeit und ohne meßbare Chloridionenleitfähigkeit.....	64
3.2.3	Stammesgeschichtlich fixierte Rekombinationsbruchpunkte im <i>CLCA</i> -Gen-Cluster.....	67
3.2.4	Auswahl des Zielbereiches und der DNA-Proben für die Resequenzierung.....	68
3.3	Ergebnisse der Teilresequenzierung des <i>CLCA</i>-Gen-Clusters	71
3.3.1	Qualität der Sequenzierung.....	71
3.3.2	Detektierte Variationen im Zielbereich der Sequenzierung: Ergebnisse der Rohdatenauswertung der Resequenzierung	74
3.3.3	Detektierte Variationen im Zielbereich der Sequenzierung: Verifizierung der Ergebnisse durch Nachtypisierung.....	76
3.4	Ergänzung der Assoziations-Studie: Typisierung der gesamten Studienpopulation an den durch Resequenzierung ermittelten Loci	82
3.4.1	Charakterisierung der ermittelten Marker.....	83
3.4.2	Stammesgeschichtlich fixierte Rekombinationsbruchpunkte im <i>CLCA</i> -Gen-Cluster – Ergänzung.....	85
3.4.3	Ergänzung der Assoziations-Studie – Verteilung der detektierten DNA-Variationen innerhalb der gesamten Studienpopulation	86
3.5	Ergebnisse der proteinchemischen Untersuchungen	90
3.6	Analyse der detektierten Marker – welche Einflüsse üben sie aus?	92
3.6.1	Auswirkung auf die physikalischen Eigenschaften der DNA – funktionelle Annotation mit der Genome-Atlas-Software	92
3.6.2	Auswirkung auf die Bindung von Transkriptionsfaktoren	97

4	Diskussion.....	103
4.1	Kartierung von Sequenzvariationen in der Kandidatengenregion mittels Allelvergleich von genetischen Markern zwischen Studienpopulationen mit kontrastierenden Phänotypen	103
4.2	Bewertung der Sequenziererergebnisse – Fehlerdiskussion	105
4.3	Mögliche Rolle der Sequenzvariationen bei der Kodierung des Modulators der alternativen Chloridionenleitfähigkeit	106
4.3.1	Differenzierung der Kernhaplotypen nach DNA-Struktur und makromolekularen Verhalten.....	107
4.3.2	Differenzierung der Kernhaplotypen nach potentiellen Transkriptionsfaktor-Bindungsstellen.....	108
4.4	Mögliche Rolle des Homologes hCLCA4 als Modulator der Mukoviszidose und dessen molekulare Funktion bei der Kodierung der alternativen Chloridionenleitfähigkeit.....	112
4.4.1	Die Rolle des Homologes hCLCA4 als genetischer Modulator der Mukoviszidose.....	112
4.4.2	Molekulare Funktion von hCLCA4 bei der Expression der alternativen Chloridionenleitfähigkeit.....	114
4.5	Schlußfolgerung und Ausblick.....	118
5	Zusammenfassung.....	120
6	Literatur.....	122
7	Anhang	133

Abkürzungsverzeichnis

ATP	Adenosintriphosphat
bp	Basenpaare
CaCC	calcium dependent chloride channels
cAMP	zyklisches Adenosinmonophosphat
CF	Cystische Fibrose
CFTR	cystic fibrosis transmembrane conductance regulator
CLCA	chloride channel, Ca ²⁺ activated
dbSNP	öffentlich zugängliche SNP-Datenbank
F508del	CFTR-Mutation: Dreibasenpaar-Deletion, resultierend in dem Verlust von Phenylalanin an der Position 508.
DIDS	4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
EDTA	Ethylendiamintetraacetat
ENaC	epithelial sodium channel
ICM	intestinal current measurement
kb	Kilo-Basen
kbp	Kilo-Basenpaare
kDa	Kilo-Dalton
Mbp	Mega-Basenpaare
n.s.	nicht signifikant
ORCC	outward rectifying chloride channels
PDZ	postsynaptic density-95, discs large, and ZO-1
PIC	polymorphism information content
pS	pico-Sievert
SDS	Sodiumdodecylsulfat
SNP	single nucleotide polymorphism
% (w/v)	Gewichtsprozent pro Volumen
% (v/v)	Volumenprozent pro Volumen

1 Einleitung

1.1 Mukoviszidose – eine Erbkrankheit [1, 2]

Die Mukoviszidose, alternativ auch als Cystische Fibrose bezeichnet, ist eine autosomal rezessiv vererbte Stoffwechselerkrankung. Mit einer Inzidenz von 1/2000 zählt sie zu den häufigsten schweren Erbkrankheiten der kaukasischen Bevölkerung.

Leitsymptom ist ein abnorm zähes Sekret aller exokriner Drüsen, welches sich klinisch sowohl an den Atemwegen als auch dem Gastrointestinaltrakt manifestiert.

Der Respirationstrakt weist bei Geburt eine weitgehend normale Morphologie auf. Infolge des hochviskösen Sekretes ist die mukoziliäre Clearance in den konduktiven Atemwegen vermindert, wodurch eine Neigung zur Obstruktion und Inflammation gegeben ist. Der Respirationstrakt wird dann im Laufe des Lebens mit pathogenen bakteriellen Keimen besiedelt. Initial handelt es sich dabei überwiegend um *Staphylokokkus aureus* und später um *Pseudomonas aeruginosa*. Dies resultiert in einer chronischen Entzündung mit konsekutiver Destruktion der Bronchien und des Lungenparenchyms. Neben der Obstruktion ist vor allem die Ausbildung von Bronchiektasen und der emphysematöse Umbau des Lungenparenchyms charakteristisch. Klinisch manifestiert sich dies vordergründig als obstruktive Ventilationsstörung. Im Endstadium kommt es schließlich zur pulmonalen Insuffizienz, vergesellschaftet mit einer pulmonalen Hypertonie, woraus sich ein Cor pulmonale mit konsekutiver Rechtsherzinsuffizienz entwickelt.

Charakteristisch ist zudem eine chronische Pansinusitis mit Ausbildung einer Polyposis nasi.

Im Gastrointestinaltrakt manifestiert sich die Mukoviszidose am Pankreas, an der Leber und am Darm.

So ist das Pankreas bei der Mehrzahl der Patienten bereits bei Geburt durch eine exokrine Insuffizienz gekennzeichnet. Dies wird auf eine Obstruktion der Drüsenausführungsgänge durch ein hochvisköses Sekret zurückgeführt. Bereits intrauterin kommt es dabei zu morphologischen Veränderungen. So sind die Pankreasgänge und Drüsenazini dilatiert, die Azinus-Epithelzellen abgeflacht und in ihrer Anzahl vermindert. Weite Teile des ortsständigen Pankreasgewebes sind durch Binde- und Fettgewebe ersetzt.

Klinisch tritt dies durch den intraluminalen Verdauungsenzymmangel im Darm als Malassimilation mit Fettstühlen in Erscheinung. Die Folge ist ein mangelhaftes Gedeihen der Patienten.

Später kann durch eine Insuffizienz des endokrinen Pankreasgewebes ein Diabetes mellitus hinzutreten.

Ein Teil der CF-Patienten erleidet in den ersten Lebenstagen einen Mekonium-Ileus. Dieses Ereignis wird auf eine mangelhafte Enzymsekretion durch das Pankreas und eine Dehydratation des Darminhaltes zurückgeführt. Dies tritt bei etwa 10-20% der CF-Patienten in Erscheinung und ist für die Diagnosefindung wegweisend.

Später können die Patienten ein distales intestinales Obstruktions-Syndrom (DIOS) entwickeln. Dies tritt meist im terminalen Ileum auf und wird als „Mekonium-Äquivalent“ gewertet, da es wie bei diesem Ereignis mit einem voluminösen, wandhaftenden und unvollständig verdauten Darminhalt vergesellschaftet ist.

In der Leber kommt es zur Ausbildung einer fokalen biliären Zirrhose, die auf eine Obstruktion der Gallekanalikuli durch ein hochvisköses Sekret zurückgeführt wird. Langfristige Folgen können eine portale Hypertension mit Ösophagusvarizen und Hypersplenismus sein, die mit Aszites und peripheren Ödemen vergesellschaftet sein können.

Weiterhin ist für die Mukoviszidose-Erkrankung ein Salzverlust-Syndrom charakteristisch. Dies beruht auf einem abnorm hohen Natriumchlorid-Gehalt des Körperschweißes, welcher durch ekkrine Schweißdrüsen abgesondert wird.

Nahezu alle männlichen Mukoviszidose-Patienten leiden aufgrund des beidseitig fehlenden Vas deferens unter Infertilität.

Derzeit wird die Lebenserwartung der Neugeborenen mit Mukoviszidose auf etwa 50 Jahre geschätzt (persönliche Mitteilung: Prof. Dr. Dr. Burkhard Tümmler). Die Prognose wird bei der Mehrzahl der Patienten vor allem durch die pulmonalen Komplikationen bestimmt.

1.2 Basisdefekt der Mukoviszidose

Mittels genetischer Kopplungsanalysen konnte der Genlocus bestimmt werden, welcher in Zusammenhang mit der Ausbildung des Krankheitsbildes der Mukoviszidose steht. Dieser befindet sich auf dem langen Arm des Chromosomes 7 (7q31) [3].

Später wurde das CF-Gen identifiziert, welches bei Mukoviszidose-Patienten krankheitsauslösende Mutationen trägt. Es handelt sich um das *CFTR* (Cystic Fibrosis Transmembrane Conductance Regulator) [4, 5, 6]. Das *CFTR* Gen kodiert für einen Chloridionen-Kanal, der vor allem in der apikalen Zellmembran von sekretorischen Epithelien lokalisiert ist.

Das *CFTR*-Gen erstreckt sich im Erbgut über 250 kbp und ist über 27 kodierende Bereiche (Exons) verteilt. Es wird in ein 6,5 kb langes Transkript (mRNA) umgeschrieben und anschließend in ein 1480 Aminosäuren großes Protein translatiert. Nach posttranslationaler Glykosylierung hat es eine Masse von etwa 170 kDa.

Das CFTR-Protein verfügt über fünf Domänen [7]. Die beiden Transmembrandomänen (MSD: membrane spanning domains) verankern es in der Plasmamembran und bilden gleichzeitig die Pore des Ionenkanals aus. Auch lassen sich zwei nukleotidbindende (NBD) und eine regulatorische (RD) Domäne abgrenzen. Erstere können Adenosintriphosphat binden und letztere verfügt über zahlreiche Phosphorylierungsstellen.

Aufgrund der Strukturhomologie wird das CFTR-Protein zur Klasse der ABC-Transporter gezählt. In Analogie zu diesen erfolgt energieabhängig unter ATP-Hydrolyse an den nukleotidbindenden Domänen die Öffnung der zentralen Pore, wenn die regulatorische Domäne durch Proteinkinasen phosphoryliert ist. Dann können gemäß des elektrochemischen Gradienten Chloridionen passiv durch die Kanalpore hindurchgeschleust werden.

Die Phosphorylierung erfolgt in erster Linie cAMP-abhängig über die Proteinkinase A, daneben wurde jedoch auch eine Phosphorylierung über die Diacylglycerol-Proteinkinase C Signalkaskade beschrieben.

Die Leitfähigkeit des CFTR-Proteins für Chloridionen beträgt etwa 6-10 pS und ist durch eine lineare Beziehung zwischen Spannung und Stromfluß charakterisiert [8]. Daneben kann das CFTR-Protein auch Glutathionen transportieren [9].

Am carboxyterminalen Ende verfügt das CFTR-Protein über ein PDZ-Bindungsmotiv. Über dieses besteht die Möglichkeit zur Interaktion mit zahlreichen anderen Proteinen [10]. So gelingt beispielsweise über Wechselwirkungen mit dem Zytoskelett die Lokalisation in der apikalen Plasmamembran [11, 12].

Neben seiner Hauptfunktion als Chloridkanal übt das CFTR-Protein auch regulatorische Einflüsse auf andere Ionenkanäle aus [13, 14]. Die Interaktion mit dem epithelialen Natriumionenkanal (ENaC) wurde bereits ausführlich untersucht. Dabei wurde überwiegend eine inhibitorische Wirkung des CFTR-Proteins beobachtet.

Daneben sind auch Einflüsse auf zwei weitere unterschiedliche Klassen von Chloridkanälen beschrieben worden. Diese sind bisher nur elektrophysiologisch charakterisiert worden. Die molekulare Entität dieser Kanäle ist noch nicht oder nicht eindeutig bekannt. Es handelt sich um die ORCC (outward rectifying chlorid channels) und die CaCC (calcium dependent chloride channels) Kanäle.

Während erstere, vermittelt über das CFTR-Protein, durch die Proteinkinase A aktiviert werden [14], führt bei letzteren eine Einhöhung der Ca^{2+} -Konzentration zur Öffnung der Kanäle. Bei diesen wirkt das CFTR-Protein zudem vorwiegend inhibitorisch [15, 16, 17]. Je

nach untersuchtem System unterscheiden sich die Charakteristika der CaCC etwas, dennoch kann man eine mittlere Leitfähigkeit von etwa 8-15 pS angeben [18].

Der durch Mutationen des *CFTR*-Gens verursachte zellphysiologische Basisdefekt wird für den erhöhten Natriumchloridionen-Gehalt im Sekret der ekkrinen Schweißdrüsen bei Mukoviszidose-Patienten verantwortlich gemacht [19]. So kann aus dem hypertonen Primärschweiß, welcher im sekretorischen Teil der Drüsen gebildet wird aufgrund der vorliegenden Defektmutation des CFTRs kein Chlorid aus dem Ausführungsgang reabsorbiert werden. Da auch die Natriumionen im Lumen zurückbleiben, ist der Körperschweiß bei CF-Erkrankten im Vergleich zu Gesunden hyperten.

Dieser Zusammenhang wird zur Diagnosefindung der Mukoviszidose mittels des Pilocarpin-Iontophorese-Schweißtest nach Gibson und Cooke ausgenutzt.

Der zellphysiologische Basisdefekt kann jedoch auch direkt anhand von Gewebeproben aus dem Intestinaltrakt mittels Ionenstromflußmessung nachgewiesen werden [20].

Der pathophysiologische Hintergrund für die erhöhte Viskosität der Sekrete aus den verschiedenen exokrinen Drüsen bei Mukoviszidosepatienten konnte bisher nicht allen Einzelheiten aufgeklärt werden. Sicherlich ist jedoch dem dysfunktionalen CFTR-Protein die tragende Rolle zuzuordnen.

1.3 Hinweise für die genetische Modulation des klinischen Verlaufes bei der Mukoviszidose-Erkrankung

Derzeit sind für das *CFTR*-Gen mehr als 1500 Mutationen beschrieben worden [21]. Unterscheidet man diese nach dem Mechanismus, welcher zur Dysfunktion des CFTR-Proteins führt, lassen sie sich in fünf Klassen einteilen [22, 23]:

Klasse I Mutationen führen primär zur verminderten Synthese des Genproduktes, Klasse II Mutationen zeichnen sich durch eine fehlgesteuerte Reifung und Prozessierung des prämaturen Proteins aus, zur Klasse III werden Mutationen gezählt, bei welchen die Regulation des CFTR-Proteins beeinträchtigt ist, Klasse IV-Mutationen zeichnen sich durch eine verminderte Leitfähigkeit des CFTR-Kanals aus und Klasse V Mutationen führen zu einer veränderten Protein-Stabilität.

Die häufigste Mutation in Europa ist F508del [24]. Während 40% aller Mukoviszidose-Patienten compound heterozygot für diese Mutation sind, tragen sie 50% aller Patienten auf beiden Chromosomen. Durch eine Dreibasenpaar-Deletion in der DNA-Sequenz enthält das CFTR-Protein an der Position 508 kein Phenylalanin. Dadurch wird es post-translational im

endoplasmatischen Retikulum und dem Golgiapparat unvollständig prozessiert und vorzeitig im Proteasom abgebaut [25]. Daher wird die Mutation zur Klasse II gezählt.

Untersuchungen an großen Patientenkollektiven haben gezeigt, daß ein gewisser Zusammenhang zwischen dem klinischen Verlauf der Erkrankung und der *CFTR*-Mutation gegeben ist. So sind die Mutations-Klassen I, II und III vorwiegend bei einer schweren Verlaufsform und die Klassen IV oder V überwiegend bei Patienten mit einer leichteren Erkrankungsform der Mukoviszidose beobachtet worden [26, 27]. Gleichzeitig haben die Untersuchungen gezeigt, daß die Variabilität des Schweregrades bei Patienten mit der gleichen Mutations-Klasse relativ hoch ist. Daher nimmt man an, daß neben Umgebungseinflüssen weitere genetische Faktoren den klinischen Verlauf determinieren.

Im Rahmen der Europäischen Mukoviszidose Zwillings- und Geschwister-Studie wurde der Einfluß der von genetischen Faktoren auf den klinischen Verlauf bei Patienten mit Homozygotie für die *CFTR*-Mutation F508del untersucht. Aufgrund der Tatsache, daß Geschwister den gleichen Umgebungseinflüssen ausgesetzt sind, sowie über ein ähnliches und im Falle von monozygoten Zwillingen sogar über ein identisches Erbgut verfügen, lassen sich Aussagen über den Einfluß der genetischen Faktoren auf den Krankheitsverlauf ableiten.

So wurden europaweit Geschwister und Zwillinge rekrutiert und anhand der Lungenfunktion sowie des Körpergewichtes der klinische Schweregrad der CF-Erkrankung beurteilt. Dabei wurden aus der Gesamtkohorte drei Gruppen von Geschwister-Paaren mit extremem klinischen Phänotyp ausgewählt: (1) Geschwister mit konkordant schwerer Verlaufsform, (2) Geschwister mit konkordant mildem klinischen Verlauf und (3) Geschwister mit diskordantem Schweregrad. Diese drei Gruppen sind bezüglich der genetischen Modulatoren besonders informativ.

Ein wichtiges Ergebnis dieser Studie war, daß der klinische Verlauf bei monozygoten Zwillingen konkordanter ist als bei dizygoten Zwillingen oder Geschwistern [28, 29, 30]. Damit wurde belegt, daß der genetische Hintergrund neben der *CFTR*-Mutation eine Rolle für den klinischen Verlauf der Mukoviszidose-Erkrankung spielt.

Zudem konnten bereits mehrere genetische Modulatoren des klinischen Verlaufes identifiziert werden, stellvertretend seien folgende genannt: Das Leptin, welches eine Schlüsselrolle im Energiehaushalt spielt [31] oder der $\text{TNF}\alpha$ -Rezeptor *TNFRSF1A*, welcher bei der Regulation der Immunantwort beteiligt ist und die beiden Untereinheiten *SCNN1B* und *SCNN1G* des epithelialen Natriumkanals *ENaC*, der für die Regulation des Flüssigkeitsfilmes in den Alveolen der Lunge wichtig ist [32].

1.4 Hinweise für die Modulation des Basis-Defektes durch alternative Chloridionen-Kanäle

Während der Europäischen Mukoviszidose Zwillings- und Geschwister-Studie wurden anhand von Gewebeproben aus dem Intestinaltrakt elektrophysiologische Messungen vorgenommen, um den zellphysiologischen Basisdefekt bei den Patienten zu untersuchen [29]. Dabei wurde mittels der intestinalen Ionenstromflußmessung (ICM: intestinal current measurement) [20] die Chloridionenleitfähigkeit charakterisiert.

Es wurde beobachtet, daß sich die Messkurven der einzelnen Mukoviszidose-Patienten substantiell voneinander unterscheiden. Diese lassen sich in drei Gruppen einteilen: Bei vielen Patienten (1) konnte keine Chloridionenleitfähigkeit nachgewiesen werden, bei einer weiteren Gruppe (2) gab es Hinweise für eine residuelle Chloridionenleitfähigkeit, die CFTR-vermittelt ist und bei einer dritten Kategorie (3) wurde eine residuelle Chloridionenleitfähigkeit gemessen, die durch 4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) hemmbar ist. Das Pharmakon DIDS kann mit Ausnahme des CFTR-Proteins [33] viele Chloridionenkanäle blockieren. Die Patienten der dritten Kategorie (3) exprimieren demnach einen anderen Chloridkanal als das CFTR-Protein. Aufgrund des verwendeten Stimulationsprotokolles kann es sich sowohl um cAMP- als auch Ca^{2+} -abhängige Kanäle handeln, wie z.B. die in Kapitel 1.2 beschriebene Kalzium-abhängige Chloridleitfähigkeit (CaCC).

Auch andere Arbeitsgruppen haben durch elektrophysiologische Untersuchungen am Epithelgewebe der Luftwege [34, 35] oder der Gallenblase [36] Hinweise für alternative Chloridionen-Kanäle bei Mukoviszidose-Patienten gewonnen.

Eine ausführliche Erläuterung des Messprotokolles und die repräsentative Darstellung der charakteristischen Messkurven erfolgt in Kapitel 2.2 des Methodenteiles.

1.5 Die Gene der *CLCA*-Familie als potentielle Kandidaten für den Modulator des Basisdefektes bei Mukoviszidose-Patienten

Als Kandidaten für die Vermittlung der beobachteten alternativen Chloridionen-Leitfähigkeit bei Mukoviszidose-Patienten kommen Kanal-Proteine in Betracht, die (1) ein ähnliches Expressionsmuster wie CFTR aufweisen, die (2) wie CFTR nicht an exzitatorischen Potentialen beteiligt sind und (3) gegebenenfalls mit CFTR in Wechselwirkung treten und im Falle des aberranten CFTRs aktiviert werden.

Prinzipiell kommen damit eine Reihe von molekularen Kandidaten in Betracht. Bisher konnte jedoch noch für keine der Proteine zweifelsfrei belegt werden, daß sie die Kalzium-abhängige Chloridionenleitfähigkeit (CaCC) kodieren [37]. Stellvertretend seien die Familien der Tweety-, der Bestrophin- und CLCA-Proteine genannt.

Die Proteine der *Tweety*-Gene sind erst vor kurzer Zeit entdeckt worden und weisen für die CaCC eine relativ hohe Leitfähigkeit auf [38]. Die Bestrophine haben eine mittlere Anionenleitfähigkeit, ihre Identität als Kanal oder Beteiligung als Untereinheit bei Kalzium-abhängigen Chloridkanälen wurde bereits belegt [39]. Die *CLCA*-Gene sind die am längsten bekannten molekularen Kandidaten für die CaCC. Sie haben eine mittlere Anionenleitfähigkeit. Bereits in den ersten Arbeiten wurde auf deren Eigenschaft als Kalzium-abhängige Chloridkanäle verwiesen. Im folgenden Kapitel erfolgt eine ausführliche Beschreibung dieser Proteinfamilie.

1.5.1 Eigenschaften und Funktionen der CLCA-Proteinfamilie

Die *CLCA*-Genfamilie ist bei verschiedenen Säugetieren einschließlich des Menschen und der Maus nachgewiesen worden. Das erste Gen *bCLCA1* wurde 1995 aus dem Rind (bovine) kloniert [40], nachdem dieses Protein bereits als Kalzium-abhängiger Chloridionenkanal charakterisiert worden war [41]. Dieses 140 kDa große Protein besteht aus 4 Untereinheiten, die je eine Masse von 38 kDa aufweisen und vermutlich über Disulfidbrücken miteinander verbunden sind. Diese gehen als Spaltprodukte aus einer etwa 100 kDa großen Vorstufe hervor.

Später wurde 1997 als zweites Homolog *bCLCA2* aus dem Rind isoliert [42], welches bereits zuvor als Zelladhäsionsmolekül charakterisiert worden war [43].

Derzeit sind beim Menschen insgesamt vier und bei der Maus sechs Homologe bekannt. Gemäß der internationalen Nomenklatur [44] werden die Proteine als Kalzium-abhängige Chloridionenkanäle bezeichnet (CLCA: **Cl** Channel, **Ca**²⁺ activated).

Die vier Homologe des Menschen sind als 230 kbp großes Cluster auf dem kurzen Arm des Chromosomes 1 (1p22-31) lokalisiert (siehe Abb. 1).

Physikalische Karte des CLCA-Gen-Clusters auf Chromosom 1

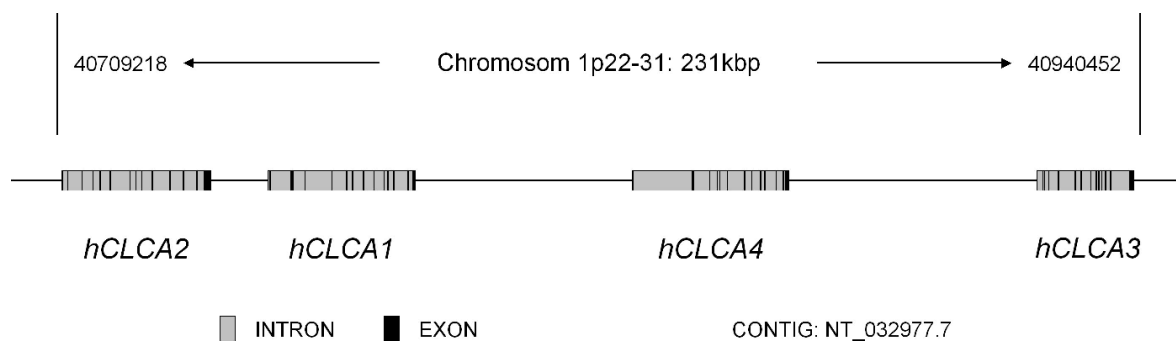


Abb. 1

Untersuchungen haben ergeben, daß sich die verschiedenen *CLCA*-Homologe durch ein spezifisches Expressionsmuster voneinander abgrenzen lassen: So ist hCLCA1 schwerpunktmäßig im Gastrointestinaltrakt nachweisbar, hCLCA2 vorwiegend in der Lunge und der Trachea, aber auch in der Glandula mammaria detektierbar und hCLCA4 wird vor allem im Kolon und in Teilen des Gehirns exprimiert. Die Variante hCLCA3 ist in verschiedenen Geweben wie Milz, Thymus, Lunge, Trachea und Glandula mammaria nachweisbar (siehe Tabelle 1).

Expressionsprofil der humanen CLCA-Homologe

hCLCA1	Dünndarm, Kolon [45] Magen, Dünndarm, Kolon, Appendix, Niere, Uterus, Testis, fetale Milz [48]
hCLCA2	Trachea, Lunge, Gl. mammaria [46] Trachea, Niere, Uterus, Testis, Prostata [48]
hCLCA3	Trachea, Lunge, Gl. mammaria, Milz, Thymus [47]
hCLCA4	Gehirn, Trachea, Magen, Dünndarm, Kolon, Appendix, Harnblase, Uterus, Testis, Prostata, Speicheldrüse, Gl. mammaria [48]

Tabelle 1

Während die Proteine hCLCA1, hCLCA2 und hCLCA4 eine posttranslationale Länge zwischen 914 und 943 Aminosäuren aufweisen [45, 46, 48], weicht das Protein hCLCA3 mit einer Größe von 262 Aminosäuren davon ab [47]. Letzteres stellt ein aminoterminals Homolog der übrigen dar.

Ausführliche strukturchemische Untersuchungen sind derzeit nur für die Homologe hCLCA1, hCLCA2 und hCLCA3 beschrieben:

Es wurde gezeigt, daß die Masse von hCLCA1 und hCLCA2 nach posttranslationaler Glykosylierung etwa 125 kDa beträgt und eine proteolytische Spaltung in zwei Untereinheiten mit der Größe von je etwa 40 kDa und 90 kDa erfolgt.

Sowohl für das aminoternale größere Fragment als auch für die carboxyterminale kleinere Untereinheit konnte eine Assoziation mit der Zellmembran nachgewiesen werden.

Während strukturchemisch beim Homolog hCLCA1 insgesamt vier Transmembrandomänen nachgewiesen werden konnten, die sich alle in dem großen Fragment befinden, sind bei der Variante hCLCA2 zwei in der großen und drei in der kleinen Untereinheit demonstriert worden.

Die Variante hCLCA3 weist nach posttranslationaler Glykosylierung abweichend von den übrigen lediglich eine Masse von 37 kDa auf und verfügt über keine Transmembrandomäne. Folglich wird diese Protein vornehmlich über den sekretorischen Weg in die Zellumgebung freigesetzt [47].

Neuere Methoden der Strukturvorhersage geben im Falle des Homologes hCLCA4 Hinweise für nur eine Transmembrandomäne im carboxyterminalen Ende [49].

Elektrophysiologische Untersuchungen anhand eines heterologen Expressions-Systemes (HEK293-Zellen) haben gezeigt [45, 46], daß die Homologe hCLCA1 und hCLCA2 die Chloridionenleitfähigkeit in Abhängigkeit von der intrazellulären Kalzium-Konzentration erhöhen. Dabei konnte eine Leitfähigkeit von 13,4 pS ermittelt werden [45]. Diese war in beiden Fällen durch den Chloridionenkanal-Hemmstoff 4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) inhibierbar.

Die dabei verwendeten Kalziumkonzentrationen lagen mit 1-2 mmol/l jedoch weit oberhalb des physiologischen Bereiches.

Untersuchungen am Homolog bCLCA1 haben gezeigt, daß die Kalzium-Empfindlichkeit durch Anwesenheit der Calmodulin-abhängigen Proteinkinase CaMKII mit Calmodulin und ATP in den mikromolaren Bereich gesenkt werden kann [41].

Neuere Untersuchungen geben zudem Hinweise, daß solche „unphysiologisch“ hohen Kalziumionen-Konzentrationen lokal durch Freisetzung aus intrazellulären Speichern und durch eine spezielle Anordnung des Tubulus-Systemes des glatten endoplasmatischen Retikulums erreicht werden können [37].

Auf der Grundlage von bioinformatischen Analysen wurde bei den CLCA-Proteinen eine von-Willebrand-Domäne vorhergesagt [49, 50]. Strukturchemisch konnte dieser Befund bisher nicht belegt oder die Relevanz für die Funktion der Proteine bestätigt werden. Von-Willebrand-Domänen sind jedoch häufig bei Protein-Protein-Interaktionen beteiligt, die teilweise über Kationen vermittelt werden [50].

Daneben wurde für verschiedene CLCA-Proteine eine Wechselwirkung mit dem β_4 -Integrin im Sinne der Vermittlung einer Zelladhäsion beschrieben [51]. Es wurde gezeigt, daß

Tumorzell-Linien, die nachweislich das β_4 -Integrin exprimieren, über Kontakte zu mCLCA1 im Lungenendothel der Maus adhären. Überdies wurde auch *in vitro* ein Kontakt dieser Zellen zum humanen Homolog hCLCA2 demonstriert. Dabei erfolgte eine Stimulation des Tumorzell-Wachstums über die Aktivierung des FAK-Signalweges (FAK: fokal adhesion kinasis) [52]. Als gemeinsames Bindungsmotiv der Homologe mCLCA1, mCLCA5 und hCLCA2 für die Wechselwirkung mit dem β_4 -Integrin wurde die Peptidsequenz F(S/N)R(I/L/V)(S/T)S abgeleitet [53].

Überdies wurde eine Korrelation zwischen der Rate des Zellumsatzes eines Gewebes und der Expression von CLCA-Proteinen beschrieben. So wurde gezeigt, daß das murine Homolog mCLCA2 vorwiegend in Geweben mit hohem Zellumsatz nachweisbar ist, welcher im Rahmen einer Proliferation oder Differenzierung auftritt. Daher vermutet man, daß das Protein mCLCA2 als Tumorsuppressor bei der Apoptose eine Rolle spielt [54].

Darüberhinaus wurde auch ein Zusammenhang zwischen der Asthma-Erkrankung und CLCA-Proteinen beobachtet: So wurde beispielsweise mittels eines Tiermodelles der Asthma-Erkrankung nachgewiesen, daß eine erhöhte mCLCA3-Expression eine Hyperreagibilität der Atemwege und eine erhöhte Mukus-Sekretion auslöst und umgekehrt eine verminderte Expression des genannten Genes diese reduziert [55].

Zudem zeigten *in vitro* Versuche anhand von menschlichen mukoepidermoiden Zell-Linien, daß die Expression von hCLCA1 oder mCLCA3 mit einer vermehrten Mukus-Sekretion einhergeht [56].

Es handelt sich bei den Homologen der *CLCA*-Genfamilie also um Proteine, die neben der Erhöhung der Chloridionenleitfähigkeit auch über viele andere Eigenschaften verfügen, aufgrund welcher sie als Kandidaten-Gene zur Kompensation des Basisdefektes bei Mukoviszidosepatienten interessant erscheinen.

1.6 Genetische Marker und deren Verwendung in Assoziationsstudien zur Identifikation von Kandidaten-Genen

Bei allen Eukaryonten ist das Erbgut als DNA im Zellkern gespeichert. Wie bei vielen Lebewesen ist dieses auch beim Menschen auf verschiedene Chromosomen verteilt. Somatische Zellen verfügen zudem von jedem Chromosom sowohl über ein mütterliches als auch ein väterliches Homolog. In diesen Zellen liegt damit ein diploider Chromosomensatz

vor, der insgesamt aus 44 Autosomen und 2 das Geschlecht determinierende Gonosomen besteht (44+2).

Während der Stammzellreifung zur Ausbildung der Gameten läuft als besondere Art der Zellteilung die Meiose ab. Hier erfolgt im Gegensatz zur gewöhnlichen Zellteilung eine Halbierung des Chromosomensatzes und zusätzlich werden Fragmente zwischen homologen Chromosomen durch Rekombination („crossing-over“) ausgetauscht. Reife Gameten verfügen deshalb über einen haploiden Chromosomensatz (22+1), bei dem jedes Chromosom sowohl väterliche als auch mütterliche Anteile aufweisen kann.

Das Erbgut unterscheidet sich zwischen verschiedenen Individuen einer Spezies in geringem Maße [57]. So gibt es Sequenzvariationen in Form von DNA-Polymorphismen, die keinen obligaten Einfluß auf die Funktion der kodierten Proteine haben. Man unterscheidet zwischen Einzelnukleotid-Polymorphismen (SNP: single-nucleotide-polymorphism) und Mikrosatelliten. Während erstere sich durch den Austausch einer einzelnen Base in der Sequenz auszeichnen, weisen letztere repetitive Sequenz-Motive auf, die durchschnittlich aus zwei bis vier Basen bestehen und sich unterschiedlich häufig wiederholen [58].

Damit haben die ersteren DNA-Marker lediglich zwei Allele und die letzteren etwa drei bis zehn. Man unterscheidet sie daher auch als dimorphe und polymorphe Marker. Beide lassen sich mit gängigen Labormethoden typisieren.

Der Genotyp bezeichnet die Allele eines Markers sowohl auf dem mütterlichen als auch auf dem väterlichen Chromosom. Dagegen beschreibt ein Haplotyp die Allele von mehreren Markern, die in Serie auf einem DNA-Strang angeordnet sind. Dieser bezieht sich daher nur auf eines der beiden homologen Chromosomen. Der Diplotyp beschreibt die Kombination von den beiden Haplotypen auf dem mütterlichen als auch dem väterlichen Chromosom (siehe Abb. 2).

Nomenklatur: Genotyp, Haplotyp und Diplotyp

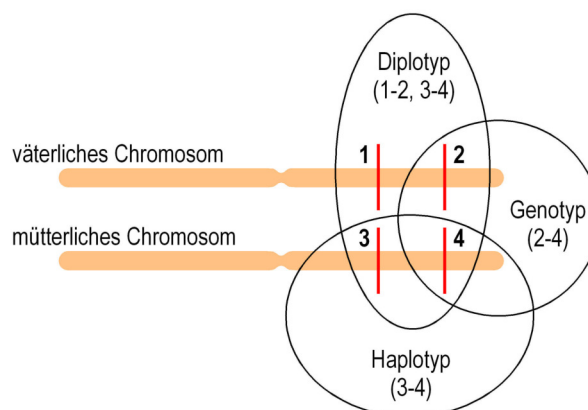


Abb. 2

In manchen Fällen erfolgt zwischen zwei Markern bei der meiotischen Zellteilung über viele Generationen keine Rekombination, sodaß diese voneinander nicht getrennt werden. Das bedeutet, diese Loci befinden sich im Kopplungsungleichgewicht. Dies kann mit Hilfe des Vier-Gameten-Testes [59, 60] abgeleitet werden (siehe Kapitel 2.8 im Methodenteil).

Werden dagegen bestimmte Allele von zwei benachbarten Markern häufiger als erwartet beobachtet, handelt es sich um eine genetische Assoziation. Dies kann neben dem Kopplungsungleichgewicht viele verschiedene Ursachen haben [61].

Bei genetischen Assoziationsstudien vergleicht man das äußere Erscheinungsbild (Phänotyp) mit der Information auf DNA-Ebene. Im einfachsten Falle handelt es sich dabei um eine Fall-Kontroll-Studie, bei der man die Studienpopulationen in einer bestimmten genetischen Region miteinander vergleicht.

Neben einer eindeutigen Klassifikation der Phänotypen und einer sorgfältigen Auswahl der Probanden für die beiden Studiengruppen ist überdies auch eine Definition der zu vergleichenden genetischen Region erforderlich. Diese wird durch die Lokalisation der Kandidaten-Gene vorgegeben, welche man auf der Grundlage von biologischen Hypothesen auswählt.

Übt eines der Kandidaten-Gene tatsächlich einen Einfluß auf den Phänotyp aus, so wird sich dies auch auf genetischer Ebene manifestieren. So wird mindestens eine kausale Sequenz-Variation vorliegen, aufgrund welcher der Einfluß des Kandidaten-Genes auf den Phänotyp bedingt sein wird. Daneben werden jedoch noch weitere Sequenzvariationen vorhanden sein, die mit der kausalen Variante aufgrund von genetischer Assoziation oder sogar eines Kopplungsungleichgewichtes gemeinsam weitervererbt werden.

Solche Sequenzvariationen können beispielsweise die Allele von DNA-Markern sein, sodaß man diese zur Eingrenzung der kausalen Sequenzvariation hinzuziehen kann.

1.7 Zielsetzung der Arbeit

Die vorliegende Arbeit befaßt sich mit der Rolle des *CLCA*-Gen-Clusters als genetischem Modulator des gastrointestinalen Basisdefektes bei Mukoviszidose-Patienten. Folgende Fragestellungen sollen bearbeitet werden:

A)

In Vorarbeiten wurde das *CLCA*-Gen-Cluster bereits als Modulator des gastrointestinalen Basisdefektes bei CF-Patienten identifiziert [62]: Vierzehn Mikrosatelliten auf Chromosom 1, die einen Bereich von 40 Mbp überspannen, wurden bei Patienten mit alternativer Chloridionenleitfähigkeit und ohne nachweisbare Chloridionenleitfähigkeit vergleichend typisiert. Die Verteilung der Markerallele zwischen den Patientengruppen war statistisch signifikant unterschiedlich für zwei Mikrosatelliten innerhalb und 4,1 Mbp stromaufwärts des *CLCA*-Gen-Clusters.

Ausgehend von dieser Assoziation zwischen Mikrosatellitengenotypen und der Expression einer alternativen Chloridionenleitfähigkeit setzte sich diese Arbeit zum Ziel, den Modulator des gastrointestinalen Phänotyps innerhalb des *CLCA*-Gen-Clusters weiter einzugrenzen und gegebenenfalls einem der vier Homologe zuzuordnen.

B)

Genetische Modulatoren zeichnen sich dadurch aus, daß sie nicht obligat den Phänotyp verändern, sondern nur fakultativ zumeist in Abhängigkeit von weiteren Faktoren das äußere Erscheinungsbild verändern.

In dieser Arbeit sollten für das *CLCA*-Gen-Cluster diese Faktoren näher bestimmt und die Sequenzvarianten identifiziert werden, unter denen der genetische Modulator den Phänotyp beeinflusst.

2 Material und Methoden

2.1 Studienpopulationen

In dieser Arbeit wurden Daten verwendet, die an zwei verschiedenen Studienkollektiven erhoben worden sind.

In diesem Abschnitt erfolgt eine allgemeine Beschreibung der Gruppen, eine genaue Charakterisierung ist dem Kapitel 2.9.1 zu entnehmen.

2.1.1 Patienten der Europäischen Mukoviszidose Zwillings- und Geschwisterstudie

Ein Großteil der Patienten dieser Arbeit sind dem Kollektiv der Europäischen Mukoviszidose Zwillings- und Geschwisterstudie [28] entnommen worden.

Diese umfaßt eine Gruppe von Mukoviszidose Erkrankten, die zumeist homozygote Träger der Mutation F508del sind. Sie besteht aus insgesamt 277 Geschwisterpaaren, 29 monozygoten und 12 dizygoten Zwillingen. Sie wurden über 158 europäische Mukoviszidose-Zentren in dem Zeitraum von 1996 bis 1998 rekrutiert.

Die im Rahmen der Zwillings-Studie vorgenommenen Untersuchungen wurden an den CF-Zentren in den Städten Hannover, Innsbruck, London, Rotterdam und Verona durchgeführt. Neben einer Blutentnahme bei den Patienten und deren Eltern erfolgte dabei auch eine intestinale Ionenstromfluss-Messung (siehe Kapitel 2.2).

2.1.2 Unverwandte und gesunde Kontroll-Probanden

Diese Studiengruppe besteht aus gesunden Kontroll-Probanden, die nicht an Mukoviszidose erkrankt sind. Sie wurden im Rahmen des Projektes SFB621-C7 über die Medizinische Hochschule Hannover rekrutiert. In diesem Projekt wurde anhand von rektalen Schleimhautbiopsien eine Transkriptom-Analyse durchgeführt (siehe Kapitel 2.13).

2.2 Intestinale Ionenstromfluß-Messung – Klassifikation der gastrointestinalen Phänotypen

Bei der intestinalen Ionenstromfluß-Messung (ICM: intestinal current measurement) wird ein transepithelialer Stromfluss in Abhängigkeit von verschiedenen applizierten Sekretagoga aufgezeichnet. Bei Kenntnis der Wirkweise dieser Substanzen kann auf die beteiligten Ionenkanäle geschlossen werden. So ist es möglich, das CFTR-Protein von alternativen Chloridionenkanälen zu differenzieren.

Auf der Grundlage dieser Messkurven wurde die Klassifikation der gastrointestinalen Phänotypen der CF-Patienten vorgenommen. Die dafür notwendigen Messungen sind während der Europäischen Mukoviszidose Zwillings- und Geschwisterstudie [29] durchgeführt worden.

Im folgenden wird zunächst das Prinzip der Messung erläutert (Kapitel 2.2.1) und dann gemeinsam mit Abb. 3 die Wirkweise von 4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) als diskriminierendes Agens erläutert (Kapitel 2.2.2).

2.2.1 Prinzip der Messung

Epitheliales Gewebe wird durch eine rektal durchgeführte Saugbiopsie entnommen und in eine Mikro-Ussing-Kammer eingespannt. Unter Supplementation des Zellstoffwechsels mit Sauerstoff und Glukose (10 mM) in Meyler-Puffer, erfolgt eine Kurzschließung des Gewebes nach dem Voltage-clamp-Verfahren (deutsch: Spannungsklemme). So kann ein transepithelialer Stromfluß in Korrelation zu den applizierten Substanzen und in Abhängigkeit der Zeit aufgezeichnet werden (siehe Abb. 3).

Die Stimulation des Gewebes mit Carbachol oder Histamin führt über eine Erhöhung der intrazellulären Ca^{2+} -Konzentration über basolaterale Kanäle zu einem Kaliumioneneinwärtsstrom. Daraus resultiert ein elektrochemisches Ungleichgewicht, welches die treibende Kraft für eine Chloridionensekretion ist. So strömen die Chloridionen über apikale Kanäle wie das CFTR-Protein aus der Zelle heraus, welches in der Messkurve als ein Trend in die Richtung zu positiven Werten erkennbar ist. Sind keine funktionellen Chloridionenkanäle vorhanden, wird der Kaliumioneneinwärtsstrom in der Messkurve als ein Trend zu negativen Werten sichtbar.

Für die Einsicht des ausführlichen Protokolls sei auf die Referenzen [20], [29] und [30] verwiesen.

2.2.2 4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) als diskriminierendes Agens

4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) ist ein Pharmakon, welches verschiedene Chloridionen-Kanäle blockiert [63]. Dies ist nachgewiesen worden für die Klasse der ORCC [64] und CaCC [65] sowie stellvertretend für die Familie der CLCA-Proteine für die Homologe hCLCA1 [45] und hCLCA2 [46].

Da das Kanalprotein CFTR durch DIDS, appliziert von der extrazellulären Seite, keine Inhibition erfährt [33], ist hiermit eine Möglichkeit zur Unterscheidung zwischen diesem und den übrigen Chloridionenkanälen gegeben.

Dazu wird das Gewebe zunächst mit Carbachol zur Chloridionensekretion stimuliert, dann nach Aufzeichnung des Antwortsignals mit DIDS inkubiert und anschließend mit Histamin erneut zur Chloridionensekretion angeregt.

Weist das zweite Signal im Vergleich zum ersten in die negative Richtung, so muß es sich um Kanäle handeln, die eine andere Entität haben, als das CFTR-Protein. (siehe Abb. 3 D).

Bleibt das Signal bei der zweiten Stimulation dagegen in die positive Richtung erhalten, so ist die Chloridionenleitfähigkeit CFTR vermittelt (siehe Abb. 3 A und 3 C).

Ist bereits nach der ersten Stimulation kein Trend in positive Richtung zu verzeichnen, so liegt keine meßbare Chloridionenleitfähigkeit vor (siehe Abb. 3 B).

Das genaue Protokoll ist gemäß Referenz [30] der Legende von Abb. 3 zu entnehmen.

Interpretation von einigen typischen Kurven der intestinalen Ionenstromflußmessung

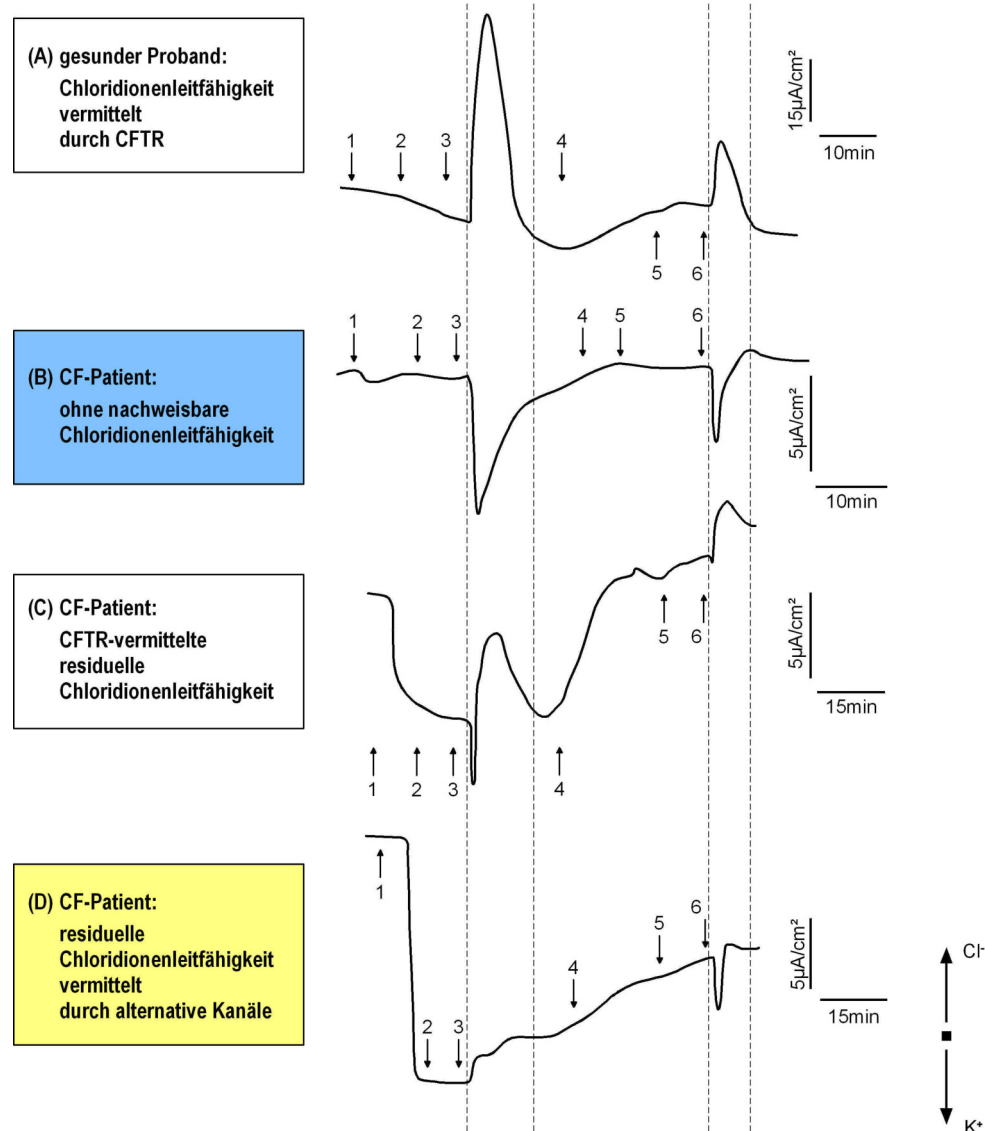


Abb. 3: Exemplarische Darstellung typischer Kurven der intestinalen Ionenstromflußmessung

PROTOKOLL: Nach der Äquilibration des Gewebes werden folgende Substanzen appliziert (M: mukosale Seite, S: serosale Seite): (1) Amilorid (10^{-4} mol/l, M), (2) Indomethacin (10^{-5} mol/l, M+S) (3) Carbachol (10^{-4} mol/l, S), (4) Forskolin (10^{-5} mol/l, M+S) gemeinsam mit 8-bromo-cAMP (10^{-3} mol/l), (5) DIDS ($2 \cdot 10^{-4}$ mol/l, M) 10min Inkubation, (6) Histamin ($5 \cdot 10^{-4}$ mol/l, S)

DIDS wird als diskriminierendes Agens eingesetzt, es blockiert mit Ausnahme von CFTR viele Chloridionenkanäle (siehe Text).

Ein Trend der Kurve nach Stimulation mit Carbachol oder Histamin in die positive Richtung zeigt einen Chloridionenauswärtsstrom und damit intakte Chloridionenkanäle an und ein Trend in die negative Richtung weist auf einen Kaliumioneneinwärtsstrom in die Zellen hin und zeigt damit funktionsuntüchtige oder fehlende Chloridionenkanäle an.

INTERPRETATION: Es sind Messungen unterschiedlicher Probanden gezeigt, **A** ist repräsentativ für eine gesunde Kontrollperson, und die Kurven **B**, **C** sowie **D** sind anhand von Gewebeproben unterschiedlicher CF-Patienten mit dem Mutationstyp F508del/F508del ermittelt worden.

Bei Kurve (**A**) folgt sowohl nach Stimulation mit Carbachol (3) als auch nach Applikation von DIDS (5) und erneuter sekretorischer Anregung mit Histamin (6) ein positives transientes Signal, dies entspricht einem normalen CFTR-vermittelten Antwortpotential.

Bei der Messung (**B**) weist das Signal sowohl nach Applikation von Carbachol (3) als auch Histamin (6) in die negative Richtung, hier liegt keine messbare Chloridionenleitfähigkeit vor, da nicht genügend intakte Chloridionenkanäle vorhanden sind. (→)

Die Kurve **(C)** zeigt dagegen nach beiden Stimulationen (Carbachol (3) und Histamin (6)) ein initial abfallendes und anschließend ansteigendes Potential, dies zeigt eine Chloridionenrestleitfähigkeit an. Da nach Applikation von DIDS (5) das Antwortsignal erneut in die positive Richtung weist, ist diese CFTR vermittelt.

Dagegen ist bei Kurve **(D)** nach Stimulation mit Carbachol (3) ein ansteigendes Potential zu erkennen sowie nach Applikation von DIDS (5) und erneuter Stimulation durch Histamin (6) ein abfallendes. Hier liegt ebenfalls eine Chloridionenrestleitfähigkeit vor, die jedoch durch DIDS-inhibierbare alternative Kanäle vermittelt ist. Adaptiert nach [19]

2.3 Kriterien zur Auswahl der genetischen Marker

Als genetische Marker wurden Einzelnukleotid-Polymorphismen verwendet.

In diesem Kapitel erfolgt nach einer Darstellung einiger statistischer Größen zur Charakterisierung genetischer Marker (Kapitel 2.3.1) eine Beschreibung der Kriterien, anhand welcher die Marker zur Typisierung der Patienten ausgewählt wurden (Kapitel 2.3.2).

2.3.1 Größen zur Charakterisierung von genetischen Markern

Die im folgenden gemachten Ausführungen beziehen sich auf dimorphe Marker:

Allelfrequenzen:

Werden die Allele a_1 und a_2 eines Markers mit der Wahrscheinlichkeit $p(a_1)$ und $p(a_2)$ auf einem zufällig ausgewählten Chromosom beobachtet, so ergibt deren Summe 1:

$$1 = p(a_1) + p(a_2) \quad (1)$$

Heterozygotität:

Diese Größe gibt die Wahrscheinlichkeit an, mit der ein zufällig ausgewähltes Individuum einer Population einen heterozygoten Genotypen trägt. Dies entspricht dem Anteil der heterozygoten Personen an der gesamten Gruppe. Die Größe wird nach folgender Gleichung berechnet:

$$\text{Heterozygotität}_{\text{erwartet}} = 2p(a_1)p(a_2) \quad (2)$$

Das Maximum für diallelische Marker beträgt 0,5. Diese Größe ist nur für eine Population gültig, die sich im Hardy-Weinberg-Gleichgewicht befindet.

Hardy-Weinberg-Gleichgewicht:

Dieses beschreibt die Genotypfrequenzen einer Population, für welche die Bedingung der Panmixie (zufällige Paarung) erfüllt ist und bei der kein Selektionsvorteil bezüglich der Genotypen gegeben ist.

Der Zusammenhang der Allelfrequenzen unter diesen Bedingungen ist durch folgenden Term gegeben:

$$1 = p(a_1)^2 + 2p(a_1)p(a_2) + p(a_2)^2 \quad (3)$$

PIC-Wert (engl: polymorphism information content):

Diese Größe definiert den Informationsgehalt eines Markers. Sie ist ein Maß für die Wahrscheinlichkeit, daß eine Rekombination aufgrund einer Betrachtung der Allele bei den Eltern und dem gemeinsamen Nachkommen eindeutig nachgewiesen werden kann [66].

Die Größe wird nach folgender Formel berechnet:

$$PIC = 1 - (p(a_1)^2 + p(a_2)^2) - 2p(a_1)^2 p(a_2)^2 \quad (4)$$

Ausgehend von dem Betrag Eins, bei welchem eine Rekombination in jedem Fall nachgewiesen werden kann, wird derjenige Anteil subtrahiert, bei welchem dieser Nachweis nicht möglich ist.

So ist mit dem zweiten Term der Anteil der homozygoten Eltern gegeben und mit dem dritten die Hälfte der Eltern mit heterozygotem Genotyp.

Für einen diallelischen Marker mit den Allelfrequenzen $p(a_1)=0,5$ und $p(a_2)=0,5$ beträgt das Maximum 0,375.

2.3.2 Auswahlkriterien der verwendeten Marker

Die Marker wurden der öffentlich zugänglichen SNP-Datenbank (dbSNP) [67] entnommen. Die Auswahl erfolgte gemäß den Kriterien:

Heterozygotität

Soweit für die untersuchte Region bekannt, wurden in erster Linie statistisch validierte SNPs mit einer Heterozygotität größer als 0,3 ausgewählt. Für diese ist ein entsprechend großer PIC-Wert zu erwarten.

Typisierbarkeit:

Die verwendeten SNPs sollten mit Hilfe von Restriktionsenzymen typisierbar sein (siehe Kapitel 2.6.3). Dazu wurde die umgebende DNA-Sequenz mit der Basenreihenfolge von verschiedenen Restriktionsschnittstellen verglichen, um ein passendes Enzym zu bestimmen.

Dies erfolgte entweder manuell anhand einer Liste (NEB-Produkt-Katalog 2005/2006) oder mit Hilfe des Suchprogrammes NEBcutter [68].

Bestätigung der Informativität:

Die Lesbarkeit und Informativität der Marker wurde durch testweise Anwendung an Kontroll-DNA-Proben (n=8...11) bestätigt, bevor der Einsatz an dem Patientenkollektiv erfolgte. Es wurden nur Marker mit erzielten PIC-Werten von 0,3 oder besser akzeptiert.

2.4 Verzeichnis der benötigten Reagenzien

Alle verwendeten Chemikalien entsprachen dem analytischen Reinheitsgrad und wurden soweit nicht anders angegeben von den Firmen Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) oder Sigma (Steinheim) bezogen.

Lysispuffer:

- 100 mM Tris-HCl
- 0,32 M Saccharose
- 1% (w/v) Triton X-100
eingestellt auf pH 7,5

TE:

- 10 mM Tris-HCl
- 1 mM Na₂EDTA
- eingestellt auf pH 8,0

STE-Puffer:

- 100 mM NaCl₂
- 50 mM Tris-HCl
- 1 mM Na₂EDTA
eingestellt auf pH 7,5

TBE (10x):

- 0,9 M Tris-HCl
- 0,9 M Borsäure
- 0,02 M Na₂EDTA

TBS (10x):

- 0,2 M Tris-HCl
- 1,37 M NaCl₂
eingestellt auf pH 7,6

Blocklösung:

- 1xTBS
- 0,05% (v/v) Tween 20
- 2% (w/v) I-Block-Reagenz

T-TBS:

- 1xTBS
- 0,05% (v/v) Tween 20

2.5 Isolierung und Handhabung der DNA

Die Typisierung der Marker erfolgte anhand genomischer DNA. Die Isolierung aus den Blutproben ist in Kapitel 2.5.1 und der Umgang inklusive Lagerung in den Kapiteln 2.5.2-3 erläutert.

2.5.1 DNA-Isolierung

Die Isolierung von hochmolekularer genomischer DNA aus Leukozyten erfolgte gemäß einem modifizierten Protokoll nach Gross-Bellard [69].

Um die Zellen aufzuschließen, wurden 5-10 ml K-EDTA Blut mit 40 ml Lysispuffer vermenget. Nach 30 minütiger Inkubation im Eiswasserbad wurden die Kerne durch Zentrifugation (600 g, 15 min) und Abdekantieren des Überstandes als Pellet von den Zelltrümmern getrennt. Es schloss sich eine Wiederholung dieses Vorganges mit 25 ml Lysispuffer und einer Inkubationszeit von 15 min an.

Der proteolytische Abbau der Histone sowie der weitere Aufschluß der Zellkerne gelang mittels Proteinase K der Firma Roche. Dazu wurde das Pellet in 4 ml STE-Puffer gelöst, mit 400 µl einer 5%igen SDS-Lösung versetzt und 50 µl Proteinase K-Lösung (50 mg/ml) hinzugegeben. Es schloss sich eine Inkubation im Schüttelwasserbad für 8-14 h bei 56°C an.

Jetzt erfolgte eine Trennung der DNA von den restlichen Proteinen. Dazu wurden 3 ml Chloroform / Isoamylalkohol (29:1) und 3 ml TE-gesättigtes Phenol hinzugegeben, für 10 min auf einer Schwenk-Wippe inkubiert und anschließend für 10 min abzentrifugiert (600 g). Nach dem Absaugen der unteren proteinhaltigen organischen Phase wurde dieser Schritt einmal wiederholt.

Um das restliche Phenol in der oberen wässrigen Phase zu entfernen, folgte eine abschließende Extraktion mit 6 ml Chloroform.

Zur Präzipitation der DNA wurde der wässrige Überstand zu einem Zehntel seines Volumens mit 3 M Na-Azetat-Lösung versetzt und anschließend das Gemisch mit eiskaltem Ethanol (99%ig) auf 50 ml aufgefüllt.

Nach 1-2 h Inkubation im Eiswasserbad wurde die ausgefallene DNA zweimal mit 70%igem Ethanol gewaschen. Dazu wurde in einem 1,5 ml-Eppendorf-Gefäß 1 ml Ethanol vorgelegt, die DNA mit einem Häkchen überführt und bei 4°C für 5 min bei 12000 g abzentrifugiert.

Anschließend wurde das Pellet nach dem Abdampfen des restlichen Ethanols in 300 µl TE-Puffer gelöst.

Nach einer mindestens vierwöchigen Lagerung bei 4°C erfolgte eine photometrische Abschätzung der DNA-Konzentration [70]. Dazu wurde die Extinktion bei 260 nm sowie 280 nm bestimmt.

Im Absorptionsmaximum für Nukleinsäuren (260 nm) entspricht der Extinktion von Eins einer Doppelstrang-DNA-Konzentration von etwa 50 µg/ml.

Das Absorptionsmaximum von Proteinen liegt bei 280 nm. Der Quotient $A_{260\text{nm}}/A_{280\text{nm}}$ lässt somit Aussagen bezüglich Verunreinigungen zu. Ein Wert von 1,8-1,9 zeigt Reinheit an, kleinere Werte hingegen geben einen Hinweis auf einen erhöhten Proteingehalt der DNA-Probe.

2.5.2 Lagerung der DNA und Vorbereitung für die PCR

Die Lagerung der DNA erfolgte nach der Isolierung ohne weitere Zusätze als Stammlösungen bei 4°C. Die so in TE-Puffer gelöste DNA war über viele Jahre haltbar (etwa zehn Jahre bis zum Zeitpunkt diese Arbeit).

Für PCR-Reaktionen wurden diese Lösungen in TE-Puffer auf eine Konzentration von 10 ng/µl verdünnt. Durch eine zweiwöchige Zwischenlagerung bei 4°C konnte eine gleichmäßige Verteilung der DNA sichergestellt werden.

2.5.3 Parallele Analyse großer Probenzahlen

Um eine große Anzahl von Analysen parallel durchzuführen, sind 96 Multiwellplatten (Fa. Greiner, Frickenhausen) verwendet worden.

Für die PCR-Reaktionen wurden mehrere Platten zu einem früheren Zeitpunkt mit Proben-DNA beschichtet und bei 4°C eingelagert. Die dabei eingetrocknete und damit in einen stabilen Zustand überführte DNA wurde vor Beginn der PCR-Reaktion durch Hinzugabe des entsprechenden Volumens an bidestilliertem Wasser wieder gelöst.

Sowohl für PCR Reaktionen als auch für die Durchführung des diagnostischen Verdaues wurde das Reaktionsgemisch mit 15 µl Paraffin überschichtet, die Platte mit einer selbklebenden Aluminiumfolie abgedichtet und der Deckel der Platte aufgelegt.

Bei PCR-Reaktionen wurden zudem unbenutzte Wells im Randbereich der Platte mit einem äquivalenten Volumen an Wasser befüllt.

Für alle Pipettierschritte wurde dabei eine Multikanal-Pipette verwendet.

2.6 Typisierung der genetischen Marker

2.6.1 Theoretische Grundlagen der Polymerase Ketten Reaktion

Mit Hilfe der Polymerase-Ketten-Reaktion werden Abschnitte eines DNA-Stranges zur weiteren Analyse vervielfältigt.

Die Reaktion wird durch die Temperatur nach folgendem repetitiv angewendeten Profil gesteuert:

- Denaturierung: Auseinandertrennung des Matrizen-doppelstranges (95°C)
- Annealing: Anlagerung der Primer an die Matrize (empirisch ermittelte Temperatur)
- Elongation: Synthese der Kopie im Temperaturoptimum der DNA-Polymerase (72°C)

Die temperaturstabile Taq-DNA-Polymerase erstellt eine komplementäre Kopie des Matrizenstranges durch Verlängerung der 3'-Enden der beiden Primer. Vorgegeben durch ihre Basenabfolge binden diese Oligonukleotide an den jeweiligen 3'-Enden des codogenen und des nicht-codogenen Template-Stranges und definieren somit den zu amplifizierenden Bereich.

Die Spezifität des Produktes hängt daher im entscheidenden Maße von der Genauigkeit ab, mit welcher der Primer an der vorgegebenen Position hybridisiert.

Unter Vorgabe der Basenzusammensetzung des Template-Stranges und bei Ausschluß der im übrigen Genom häufig vorkommenden Sequenzen wird dies für Primer mit einer Länge von etwa 20 Basen [71] erreicht. Zur Bestimmung dieser Oligonukleotid-Sequenzen wurde das Programm Primer3 [72] verwendet.

Die Bindung der Primer an das Template kann durch die Ionenkonzentration des Puffers beeinflusst werden: Nukleotidpolymere haben ein negativ geladenes Rückgrat, somit erfolgt bei der Hybridisierung dieser Stränge eine wechselseitige Abstoßung. Diese kann durch Kationen als Gegenionen vermindert werden. Zur Einstellung der Ionenkonzentration wurde $MgCl_2$ verwendet. Magnesiumionen sind zudem ein Co-Faktor der DNA-Polymerase.

Auf die Stringenz der Reaktion hat daneben nicht nur die Temperatur im Anlagerungsschritt an sich einen Einfluß, sondern auch die Geschwindigkeit, mit welcher ausgehend von der Denaturierungsphase abgekühlt wird. Um diesen Einfluß so gering wie möglich zu halten, wurde für die Abkühlung eine Rampe mit ausreichend großer Zeitrage definiert, die für alle Anlagerungstemperaturen gleich ist.

Des Weiteren kann Dimethylsulfoxid (DMSO) als chaotropes Reagenz die Wechselwirkungen der Nukleotidpolymere vermindern und so die effektive Schmelztemperatur senken [73]; insbesondere in GC-reichen Regionen kann dadurch die Spezifität verbessert werden [74].

Da für die Amplifikation theoretisch die DNA-Menge einer Zelle ausreicht, ist dieses Verfahren anfällig für Kontaminationen. Eine Leerkontrolle wurde daher bei jeder Reaktion mitgeführt, also anstatt von Template Wasser hinzugegeben.

2.6.2 Durchführung der PCR-Reaktion

Für alle PCR-Reaktionen wurde ein Thermocycler der Firma peqlab (Erlangen) verwendet. Die Synthese der Primer wurde bei den Firmen Invitrogen (Karlsruhe) und Invitex (Berlin) in Auftrag gegeben. Die Amplifizierung der Proben zur Typisierung der Marker erfolgte mit der Taq-DNA-Polymerase von der Firma Invitex (Berlin).

Für die Reaktionen wurde folgendes Temperaturprofil verwendet:

	Temperatur	Zeit	Zyklenzahl
Denaturierung	92°C	180 s	1x
Anlagerung	AT*	60s	
Elongation	72°C	60 s	35x
Denaturierung	92°C	30 s	
Anlagerung	AT*	60 s	1x
Letzte Elongation	72°C	180 s	1x
Abkühlschritt	35°C	30 min	1x
	20°C		

* Die Abkühlung ausgehend von der Denaturierung erfolgte mit einer Geschwindigkeit von 1°C/s. Die Anlagerungstemperaturen (AT) wurden für jedes Primerpaar empirisch ermittelt.

Die Sequenzen der Primer sind aus Tabelle 15 im Anhang zu entnehmen.

Protokoll:

5 µl	DNA (10 ng/µ, gelöst in 1x TE)	50 ng
3 µl	10x-PCR-Puffer	InVitek
3 µl	dNTP	2 mM jedes Nukleotid
1-3 µl	MgCl ₂ **	50 mM
0; 1 µl	DMSO**	unverdünnt
3 µl	Primer A	5 µM
3 µl	Primer B	5 µM
0,08 µl	Taq-Polymerase	5 u/µl, InVitek

Das Reaktionsgemisch wurde mit bidestilliertem Wasser zu einem Volumen von 30 µl ergänzt.

** Die notwendige Menge von MgCl₂ sowie die Anwendung von DMSO wurden im Rahmen der Optimierung für jedes Primerpaar empirisch ermittelt.

2.6.3 Differenzierung der Allele

Die Allele der SNPs wurden durch Anwendung von zwei verschiedenen Methoden unterschieden. Der Restriktionsverdau ist in Kapitel 2.6.3.1 beschrieben und die allelspezifische PCR in 2.6.3.2 erläutert.

2.6.3.1 Allelunterscheidung durch Restriktionsverdau

Wie in Kapitel 2.3.2 beschrieben, war Grundlage der Auswahl der Marker, daß sie mit der umgebenden DNA-Sequenz eine Erkennungsstelle für Restriktionsenzyme bilden. Auf diese Weise können solche Endonukleasen zur Allelunterscheidung herangezogen werden.

Jedes dieser Enzyme spaltet doppelsträngige DNA an spezifischen, zumeist palindromischen Sequenzen. Diese sind etwa vier bis acht Basen lang. Schließt eine solche Nukleotidabfolge einen SNP ein, wird sie den Allelen entsprechend verändert. Dadurch ist die Erkennungssequenz für das Restriktionsenzym nur bei einem der beiden allelischen Varianten vorhanden.

Als Allel 2 wurde diejenige Base des SNPs definiert, bei der eine Spaltung erfolgt und in Analogie dazu wurde als Allel 1 dasjenige Nukleotid bezeichnet, welches die Erkennungssequenz maskiert und somit keine Spaltung zuläßt.

In manchen Fällen, bei denen durch die umgebende Sequenz des Markers eine Erkennungstelle nur unvollständig definiert war, konnte durch Mutagenese-Primer eine artifizielle Restriktionsschnittstelle in die PCR-Produkte eingeführt werden. Dazu wurden Primer nahe des SNPs positioniert, welche die Basenreihenfolge des Templates zur vollständigen Erkennungssequenz des Restriktionsenzymes ergänzten. Diese waren demnach nicht vollständig komplementär zur Matrize, sondern enthielten bis zu zwei fehlerhaft gepaarte Nukleotide. Um der DNA-Polymerase einen adäquaten Startpunkt zu bieten, wurde bei der Positionierung der bezeichneten Nukleotide stets ein Mindest-Distanz von einer Base zum 3'-Ende des Primers eingehalten.

Die so eingeführte Spaltstelle befindet sich nahe am Ende des PCR-Produktes.

Damit die Allele durch Elektrophorese (siehe Kapitel 2.6.4) eindeutig diskriminiert werden können, ist ein relativer Längenunterschied zwischen gespaltenen und ungespaltenen Produkten von mindestens 10% notwendig. So kann mit einem 4%igen Agarosegel der Längenunterschied von 200 bp zu 180 bp gut dargestellt werden.

Bestimmung der Allele eines SNPs mittels Restriktionsverdau

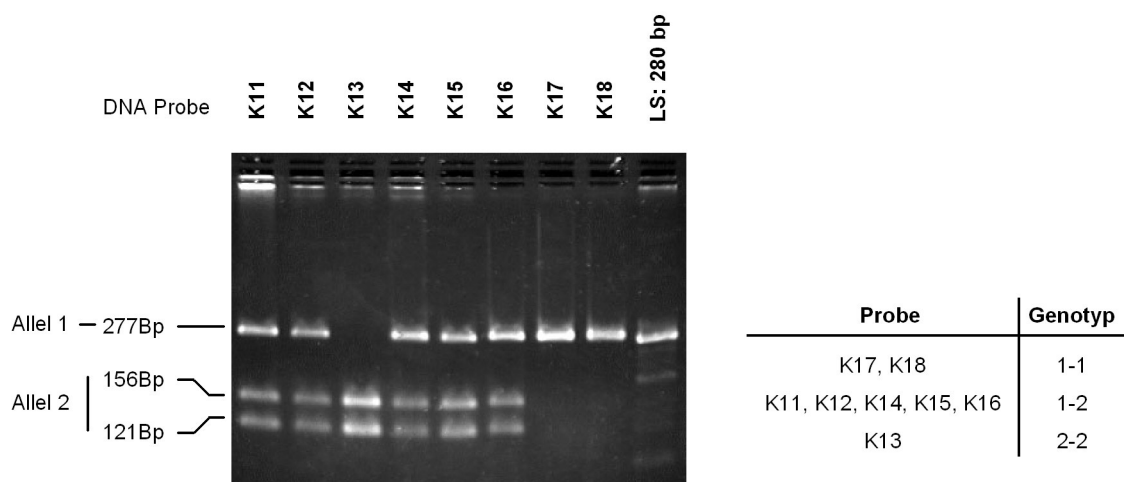


Abb. 4: Exemplarische Darstellung der Typisierung eines SNPs mittels PCR und Restriktionsverdau. Über die Länge der Produkte kann man Rückschlüsse auf die Allele des SNPs ziehen. Der Verdau der PCR-Produkte von DNA-Probe K17 und K18 resultiert in einer Bande, bei K13 sind es zwei kleine und bei K11, K12, K14, K15 und K16 sind neben der großen Bande beide kleinen zu sehen. Durch den Verdau ist das Produkt in zwei Hälften gespalten worden. So sind zwei schneller laufende Banden zu erkennen, die in der Summe die Länge des Ausgangsproduktes bilden. Da ein diploider Chromosomensatz vorliegt, können beide Varianten in einer Spur beobachtet werden, d.h. das Enzym konnte nur auf einem der beiden amplifizierten Strängen die Erkennungssequenz wiederfinden.

2.6.3.1.1 Restriktionsverdau unaufgereinigter PCR-Produkte

Jeweils 10 µl PCR-Produkt wurde mit der katalytischen Mengeneinheit 1 U des entsprechenden Enzyms inkubiert. Dazu wurde 10 µl Verdau-Puffer für ein Reaktionsvolumen von 20 µl vorbereitet, das Enzym hinzugegeben und dem vorgelegten PCR-Produkt hinzugefügt. Das Reaktionsgemisch wurde mit 15 µl Paraffin überschichtet und über Nacht (8-14 h) bei der entsprechenden Temperatur inkubiert.

2.6.3.1.2 Restriktionsverdau nach Ethanol-Fällung

Jeweils 30 µl des PCR-Produktes wurden zu einem Zehntel des Volumens mit 3 M Na-Azetat-Lösung versetzt und mit 120 µl reinem Ethanol vermengt.

Nach einer ein- bis zweistündigen Inkubation bei 4°C wurden die ausgefallenen Produkte für 15 min mit einer Eppendorf-Zentrifuge bei maximaler Drehzahl pelletiert.

Nach Abdekantieren des Überstandes wurde das Pellet zweimal gewaschen. Dies erfolgte durch Hinzugabe von 120 µl 70%igen Ethanols mit einer anschließenden Zentrifugation unter gleichen Bedingungen.

Nach dem Abdekantieren des Überstandes wurde das Pellet schließlich in 10 µl destilliertem Wasser aufgenommen und der Verdau nach dem oben beschriebenen Verfahren durchgeführt.

2.6.3.1.3 Restriktionsverdau nach Adsorptionschromatographie

Die Aufreinigung der PCR-Produkte mit dem QIAquick-PCR-Purification Kit (Fa. Qiagen, Hilden) erfolgte gemäß QIAquick Spin Handbook (Stand: Juli 2002).

Dabei wurde mit jeder Säule 30 µl des PCR-Produkts aufgereinigt.

Als Elutionsmittel wurde dabei der Puffer verwendet, welcher für das nachfolgend eingesetzte Restriktionsenzym empfohlen ist.

2.6.3.2 Allelspezifische PCR

Mit der allelspezifischen PCR können SNPs genotypisiert werden. Dabei wird für jedes Allel ein Primer zusammen mit einem gemeinsamen Gegenprimer verwendet. Für jedes Allel wird somit ein gesonderter Reaktionsansatz notwendig.

Die Primer sind so ausgelegt, daß die Base ihres 3'- Endes auf der Position des SNP-Nukleotides zu liegen kommt. Sind diese komplementär zueinander, hybridisiert der Primer über seine ganze Länge mit der Matrize, und die DNA-Polymerase wird ein Produkt bilden, dessen Länge durch die Position des gemeinsam verwendeten Gegenprimers bestimmt wird.

Wenn die Basen nicht komplementär zueinander sind, wird der Abstand des 3'-Endes des Primers vom Template-Strang zu groß sein, um einen Startpunkt für die DNA-Polymerase zu bilden, die Amplifikation bleibt aus.

Als Polymerase wurde die wenig promiskuitive Goldstar-DNA-Polymerase von Eurogentec (Fa. Seraing, Belgien) eingesetzt. Sie hat aufgrund einer „proof-reading“-Aktivität eine geringere Fehlerrate als gewöhnliche Taq-Polymerasen und liefert damit reproduzierbare Ergebnisse, sodaß sie u.a. für die Applikation in Sequenzieranwendungen geeignet ist (Datenblatt Eurogentec).

Die Bedingungen der Reaktion wurden für jeden allelspezifischen Primer so gewählt, daß einerseits ausreichend viel Produkt bei Komplementarität gebildet wird, andererseits aber keine Amplifikation der DNA stattfindet, wenn das 3'-Ende des Primers zur Matrize nicht passt.

Die Optimierung erfolgte anhand von DNA mit bekannten homozygoten Genotypen.

Dabei wurde zunächst die Reproduzierbarkeit eines Rezeptes durch mehrfache (drei mal) Durchführung der allelspezifischen PCR unter den ermittelten Bedingungen überprüft.

Ergab diese Austestung jeweils dieselben eindeutig interpretierbaren Ergebnisse, erfolgte eine Typisierung der gesamten Studienpopulation bei diesen Bedingungen.

Um insbesondere falsch positive Ergebnisse auszuschließen, wurden für eine initiale Amplifikation die Bedingungen sehr stringent gewählt. Gegeben durch Assay-zu-Assay Variationen verschiedener Reaktionskomponenten (effektive DNA-Konzentration, $MgCl_2$ -Konzentration) sind dabei Ausfälle relativ häufig; sie werden an der fehlenden Produktbildung in den Reaktionsansätzen beider allelspezifischer Primer erkannt. Bei der Nachtypisierung dieser Proben wurde eine niedrigere Anlagerungstemperatur vorgegeben sowie die $MgCl_2$ -Konzentration verändert.

Eine Kontrolle der Genotypisierung erfolgte durch Analyse der Vererbung eines Alleles innerhalb jeder Familie und durch einen Vergleich der Genotypen bei monozygoten Zwillingspaaren. Sind beide Kriterien konform mit den Mendelschen Regeln, so wurde die Typisierung als richtig bewertet.

Dieses Verfahren kann alternativ zum Restriktionsverdau eingesetzt werden, wenn dieser keine eindeutig interpretierbaren Resultate ergibt oder prinzipiell nicht anwendbar ist. Aufgrund möglicher falsch positiver oder negativer Resultate wurde dieses Verfahren bei nur einem SNP angewendet.

Allelspezifische PCR, vergleichende Darstellung der Ergebnisse unter optimalen und bei zu wenig stringenten Reaktionsbedingungen

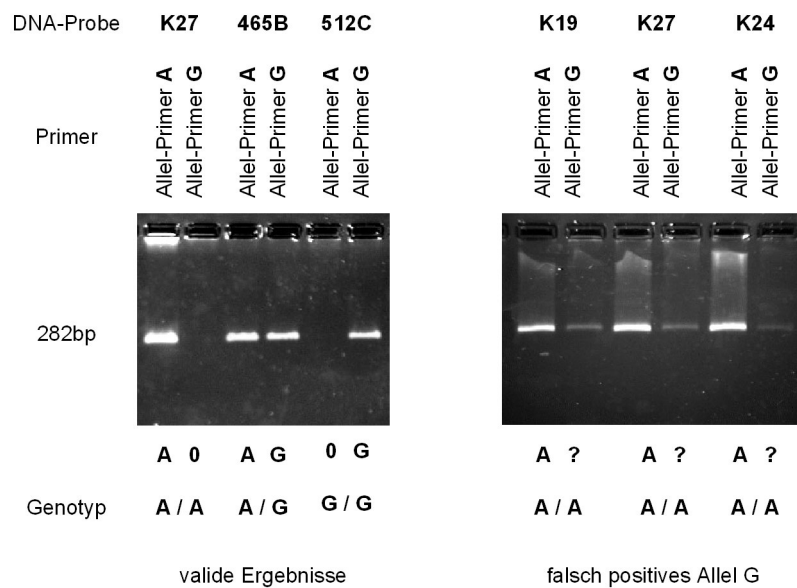


Abb. 5: Exemplarische Darstellung der Typisierung eines SNPs mittels allelspezifischer PCR. Es sind zwei Gele gezeigt. Im linken Teil sind Ergebnisse unter optimalen Reaktionsbedingungen und im rechten bei zu wenig stringenten Bedingungen dargestellt.

Für jede Probe sind zwei Spuren vorhanden, die linke ist stellvertretend für das Allel A und die rechte für die Allel G.

So ist bei Probe **K27** im rechten Gel im Vergleich zum linken eine schwache Bande für das Allel G zu sehen, dies ist ein typisches falsch positives Ergebnis.

2.6.4 Gelelektrophorese und Darstellung der PCR-Produkte

Geladene Moleküle wandern in einem elektrischen Feld zur Elektrode mit der gegensätzlichen Ladung. Da die Wanderungsgeschwindigkeit von der Form und Größe der Moleküle sowie von der Beschaffenheit der Matrix abhängig ist, in welcher sich diese bewegen, kann damit eine Trennung von DNA nach der Länge erfolgen. Dieses Prinzip wurde für die Längenbestimmung der PCR- und Spaltprodukte ausgenutzt.

Als Medium wurde Agarose gelöst in 1xTBE eingesetzt. Je nach gefordertem Auflösungsvermögen wurden Massenkonzentrationen bis 4 % verwendet. Um ein möglichst homogenes Gel zu erhalten, wurde die Agarose vor dem Erhitzen auf einem Magnetrührer bei Raumtemperatur für 1 h präinkubiert.

Die schlierenfreie Lösung wurde dann in Kammern mit speziell für Mehrkanalpipetten angepasstem Spurbstand gegeben. Nachdem sie erstarrte, wurden die Produkte, vermengt in einem Verhältnis von 1:1 mit Auftragspuffer (je eine Spatelspitze Bromphenolblau und Xylencyanol gelöst in 60%iger wässriger Glycerinlösung), in die Taschen des Geles gegeben und die Elektrophorese bei 4°C mit einer Feldstärke von etwa 7 V/cm durchgeführt.

Im Anschluß wurde die DNA mittels der interkalierenden Substanz Ethidiumbromid angefärbt. Dazu wurde für jedes Gel in einem frisch angesetzten Bad (40 µl einer Stammlösung (10 mg/ml) in 1 l demineralisiertem Wasser) für 30 min inkubiert und anschließend zweimal für je 20 min mit Wasser entfärbt. Zur besseren Anfärbung und Erhöhung des Kontrastes wurden diese Zeiten bei hochprozentigen Gelen auf je bis zu 1 h verlängert.

Die angefärbten Nukleinsäurebanden wurden, angeregt durch UV-Licht (302 nm), als orangefarbene Fluoreszenz sichtbar. Jedes Gel wurde so mit einem Geldokumentationssystem photographiert.

Die Größe der Produkte wurde durch Vergleich anhand eines aufgetragenen Längenstandards bestimmt.

2.7 Rekonstruktion der Haplotypen

Rekonstruktion von Haplotypen mit Hilfe der elterlichen Genotypen

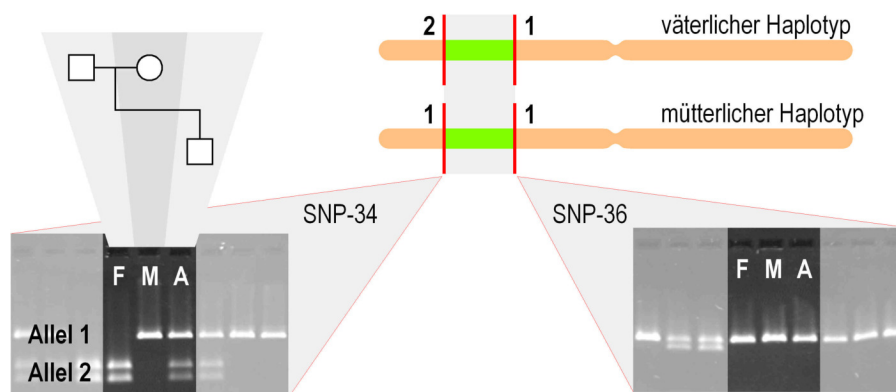


Abb. 6: Zur Ableitung der Haplotypen wurden die Genotypen der Eltern am gleichen Locus wie beim Patienten ermittelt. Unter der Annahme, daß in der elterlichen Generation während der Keimzellentwicklung keine Rekombination zwischen den untersuchten Markern stattgefunden hat, kann so eine Zuordnung der Allele des Patienten auf die beiden Chromosomen erfolgen.

So kann das Allel 2 des Markers SNP-34 beim Patienten nur auf dem väterlichen Chromosom liegen, das Allel 1 dagegen kann ausschließlich Teil des mütterlichen Chromosomes sein.

Beim SNP-36 ist eine Zuordnung ebenfalls möglich, hier jedoch läßt sich nicht genau festlegen, welches Allel 1 von welchem Elternteil vererbt wurde.

LEGENDE: Spurbezeichnung: **F:** Vater, **M:** Mutter, **A:** Patient

Die Ableitung der Haplotypen erfordert bei diploiden Individuen die Zuordnung der Allele eines Markers zu einem bestimmten Chromosom. Dies wurde nach dem in Abb. 6 erläuterten Prinzip mit Hilfe der elterlichen Genotypen vorgenommen.

Einschränkend ist anzumerken, daß eine Rekonstruktion der Haplotypen bei dimorphen Markern nach diesem Prinzip unmöglich ist, wenn sowohl der Patient als auch dessen Eltern heterozygote Genotypen tragen. In einem solchen Fall blieb die Rekonstruktion unvollständig.

2.8 Konstruktion von stammesgeschichtlich fixierten Rekombinationsbruchpunkten

Die stammesgeschichtlich fixierte Rekombination ist neben der Punktmutation ein wesentliches Element der Entstehung von Haplotypen.

Im Gegensatz zur gewöhnlichen Rekombination, tritt diese nicht in der Eltern-Generation der Studienpopulation auf. Denn sie hat sich bei einem stammesgeschichtlich gemeinsamen Vorfahren zu einem in der Vergangenheit weit zurückliegenden Zeitpunkt ereignet. Der durch die Rekombination verursachte Bruchpunkt läßt sich mit dem Vier-Gameten-Test [59, 60] konstruieren und lokalisieren.

Dazu werden folgende Annahmen gemacht:

An jeder Stelle eines gegebenen DNA-Stranges können Mutationen auftreten, diese sind selten, sodaß wiederholte Ereignisse an demselben Locus nicht eintreten. Die Mutationen werden an die Nachkommen weitergegeben und benachbarte können zudem als Haplotypen kombiniert werden. Dabei kann es sich beispielsweise um SNPs handeln, diese genetischen Marker sind dimorph und haben sehr stabile Allele.

Daraus folgt: Ist ein beobachtetes Allel häufig, so ist es über viele Generationen von einem gemeinsamen Vorfahren ausgehend von einem in der Vergangenheit weit zurückliegenden Zeitpunkt an eine große Zahl von Nachfahren weitergegeben worden. Umgekehrt lassen sich seltene Allele auf ein Mutationsereignis in der jüngeren Vergangenheit zurückführen. Dies läßt sich analog auf die abgeleiteten Haplotypen übertragen.

Es folgt weiterhin: Für Zwei-Marker-Haplotypen, bestehend aus dimorphen Markern, lassen sich vier verschiedene Kombinationen konstruieren. Durch Mutationsereignisse kann die Entstehung von höchstens drei Haplotypen erklärt werden. Die vierte Variante geht dabei aus einer Rekombination von zwei der bereits existierenden Haplotypen hervor. Der Bruchpunkt liegt dann zwischen den beiden Markern.

Unter der Annahme, daß jeder Haplotyp mit der gleichen Wahrscheinlichkeit an die Nachfahren vererbt wird, kann die Reihenfolge ihrer Entstehung aus den beobachteten Häufigkeiten abgeleitet werden: Der Haplotyp mit der größten Frequenz ist dann zuerst

entstanden und die übrigen sind gemäß ihrer Häufigkeit nach in absteigender Folge sukzessiv aus diesem hervorgegangen.

Diese Annahme setzt jedoch voraus, daß kein Träger eines dieser Varianten einem Selektionsnachteil unterliegt.

Für den Beweis einer stammesgeschichtlich fixierten Rekombination reicht dabei eine einzelne Beobachtung des vierten Haplotypen innerhalb der untersuchten Studienpopulation aus.

In Abb. 7 ist der beschriebene Prozess durch eine Darstellung der möglichen Haplotypen illustriert.

Entstehung von Haplotypen durch Mutationsereignisse und stammesgeschichtlich fixierte Rekombinationen

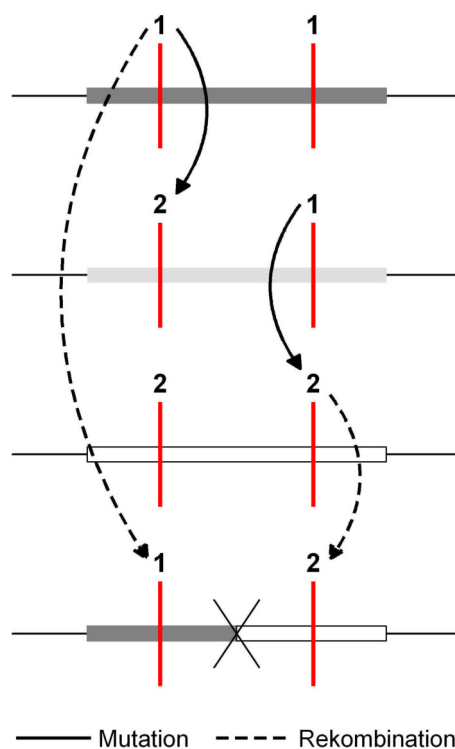


Abb. 7: Ableitung eines stammesgeschichtlich fixierten Rekombinationsbruchpunktes. Die Haplotypen 2-1 und 2-2 sind durch Punktmutationen aus der Variante 1-1 hervorgegangen. Dagegen ist der Haplotyp 1-2 durch eine ancestrale Rekombination entstanden. Die Haplotypen sind ihrer Häufigkeit nach geordnet, die Variante 1-1 hat die größte Frequenz und 1-2 ist am seltensten. Das Ereignis einer stammesgeschichtlich fixierten Rekombination ist daher für 2-Marker-Haplotypen bestehend aus dimorphen Markern bereits mit der Beobachtung aller vier möglichen Varianten nachgewiesen.

2.9 Charakterisierung der Patientenkollektive und Erläuterung der statistischen Auswertung der Assoziationsstudie

Bei der im Rahmen dieser Arbeit durchgeführten Assoziationsstudie wurden Patienten mit alternativer Chloridionenrestleitfähigkeit in der genetischen Region der *CLCA*-Proteine mit Mukoviszidose-Erkrankten verglichen, bei denen keine Chloridionenleitfähigkeit nachgewiesen werden konnte.

Dazu wurden genetische Marker in der Region des *CLCA*-Clusters typisiert, welche nach den in Kapitel 2.3.2 beschriebenen Kriterien ausgewählt worden sind und die Verteilung der Allele dieser Marker zwischen den beiden Patientengruppen verglichen.

In Kapitel 2.9.1 erfolgt eine genaue Beschreibung der Patientengruppen, welche in dieser Studie miteinander verglichen worden sind. Und in Kapitel 2.9.2 wird die statistische Methode erläutert, mit welcher die genetischen Daten der Studie ausgewertet worden sind.

Darüberhinaus wird in Kapitel 2.9.1.2 die Gruppe der unverwandten Einzelpatienten beschrieben, anhand welcher die in Kapitel 2.13 beschriebene Transkriptomanalyse durchgeführt worden ist.

2.9.1 Charakterisierung der Patientengruppen

2.9.1.1 *Patienten der Assoziationsstudie*

Die Patienten dieser Studien wurden der Europäischen Mukoviszidose Zwillings- und Geschwisterstudie entnommen [28].

Die Auswahl erfolgte gemäß den Ergebnissen der intestinalen Ionenstromfluß-Messung (siehe Kapitel 2.2).

Es wurden nur die DNA-Proben solcher Probanden genotypisiert, bei denen eindeutig interpretierbare Resultate der intestinalen Ionenstromfluß Messung vorlagen. Dabei wurden die Patienten mit alternativer Chloridionenleitfähigkeit in einer Gruppe I und CF-Erkrankte ohne meßbare Chloridionenleitfähigkeit zu einer Gruppe II zusammengefaßt.

Abweichend davon wurde bei den Gewistern dieser Patienten, unabhängig von den Ergebnissen der intestinalen Ionenstromfluß-Messung, stets eine Genotypisierung der Marker vorgenommen, um mehr Informationen über die genetische Region des *CLCA*-Clusters zu erhalten.

Monozygote Zwillinge wurden nur bei konkordanten Ergebnissen der intestinalen Ionenstromfluß-Messung eingeschlossen. Um einen systematischen Fehler bezüglich doppelt

gezählter Genotypen zu vermeiden, wurde jeweils nur ein Zwilling in der statistischen Auswertung berücksichtigt.

Soweit verfügbar, wurden zudem die DNA-Proben von den Eltern der Patienten typisiert, um die Haplotypen korrekt ableiten zu können.

Durch Untersuchungen anhand von gesunden Kontroll-Personen wurde beobachtet, daß das Antwortpotential bei der intestinalen Ionenstromflußmessung mit zunehmenden Alter der Probanden kleiner wird [75].

Daher wurden hier die Patienten mit alternativer Chloridionenleitfähigkeit (Gruppe I) mit zwei unterschiedlichen Gruppen von CF-Erkrankten ohne Chloridionenleitfähigkeit verglichen. So wurden letztere einerseits zu einer Gruppe IIa zusammengefaßt, in welcher alle verfügbaren Patienten mit dem bezeichneten Phänotyp vertreten waren und andererseits erfolgte eine Stratifizierung nach dem Alter, sodaß die Anzahl der Probanden in den miteinander verglichenen Gruppen gleich groß war. Dazu wurde ausgehend von der kleineren Gruppe I, jedem CF-Erkrankten mit alternativer Chloridionenleitfähigkeit ein Patient ohne nachweisbare Chloridionenleitfähigkeit zugeordnet, sodaß die Abweichung des Alters beider Probanden möglichst gering war. Die auf diese Weise ermittelten Patienten ohne residuelle Chloridionenleitfähigkeit wurden zu der Gruppe IIb zusammengefaßt.

Ingesamt wurden die Proben von 26 Kernfamilien sowie von 5 weiteren Familien mit nur teilweise verfügbaren DNA-Proben der Eltern analysiert.

Damit wurden 119 DNA-Proben typisiert, von denen 54 von den Eltern und 65 von den Patienten stammten. Gemäß des intestinalen Phänotypen teilt sich die Gruppe der Patienten in 26 ohne meßbare, in 14 mit alternativer, in drei mit CFTR-vermittelter Chloridionenleitfähigkeit sowie in 22 CF-Erkrankte mit Meßkurven, welche Hinweise für gemischte Phänotypen geben.

Durch die Stratifizierung der Patienten nach dem Alter ergaben sich folgende Gruppen (siehe Abb. 8, Seite 36):

So besteht die Gruppe I aus 13 Patienten mit alternativer Chloridionenleitfähigkeit. Von den neun weiblichen und vier männlichen CF-Erkrankten tragen acht den CFTR-Genotyp F508del auf beiden Chromosomen, während vier Patienten diesen Genotyp neben einer seltenen Mutation nur auf einem Chromosom tragen. Bei einem Patient entsprechen beide chromosomale Kopien jeweils einer seltenen CFTR-Mutation.

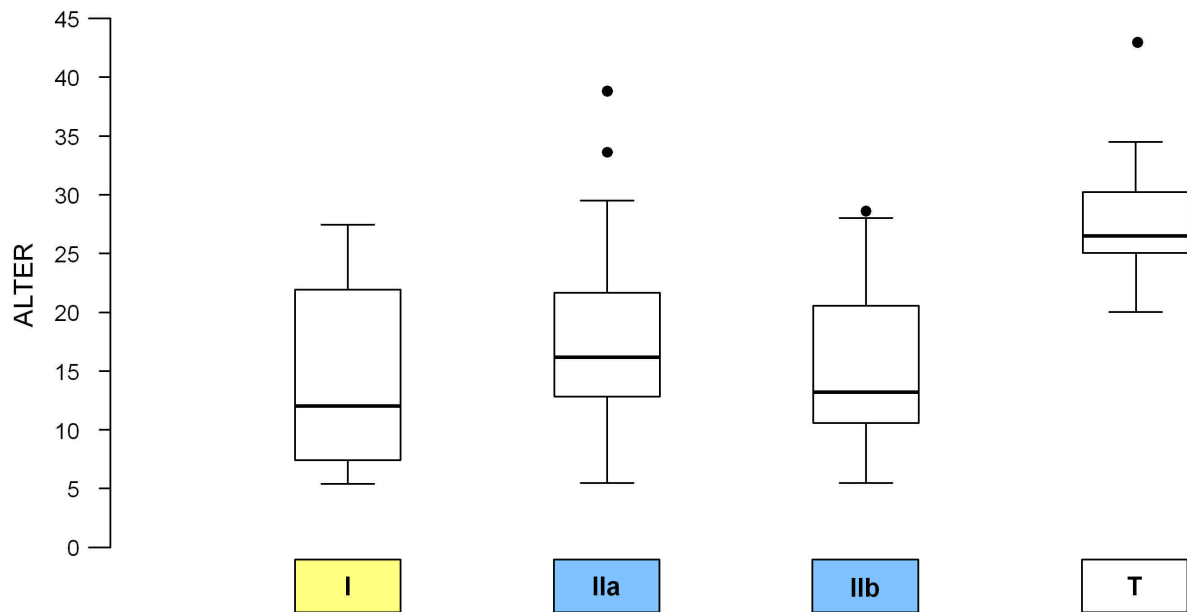
Die Gruppe IIa umfaßt 22 Patienten ohne meßbare Chloridionenleitfähigkeit, von denen 12 weiblich und zehn männlich sind. Den homozygoten CFTR-Mutationsgenotyp F508del tragen

17 Patienten dieser Gruppe, während fünf Patienten diesen Genotyp neben einer seltenen Mutation nur auf einem Chromosom tragen.

Die Gruppe IIb besteht aus 13 Patienten ohne nachweisbare Chloridionenleitfähigkeit, von denen sieben weiblich und sechs männlich sind. Von diesen sind 11 Patienten homozygot für den CFTR-Mutationsgenotyp F508del und zwei CF-Erkrankte tragen diesen Genotyp neben einer seltenen Mutation nur auf einem Chromosom.

2.9.12 Unverwandte und gesunde Kontroll-Probanden – Transkriptomanalyse

Diese Studiengruppe wurde im Rahmen des Projektes SFB621-C7 rekrutiert und umfaßt vier männliche und vier weibliche Kontroll-Probanden. Sie sind gesund und miteinander nicht verwandt. Anhand dieser acht Personen wurde die in Kapitel 2.13 beschriebene Transkriptomanalyse durchgeführt. Die Charakterisierung dieser Gruppe T bezüglich des Alters erfolgt in Abb. 8.

Boxplot-Darstellung des Alters der einzelnen Studienpopulationen

	I	IIa	IIb	T
Maximum	27,5	38,8	28,6	43,0
75% Quartil	22,0	21,6	20,6	30,75
Median	12,1	16,2	13,1	26,5
25% Quartil	7,4	12,8	10,6	25
Minimum	5,4	5,4	5,4	20
Patientenzahl	13	22	13	8

- I:** Patienten mit alternativer Chloridionenleitfähigkeit
IIa: Patienten ohne meßbare Chloridionenleitfähigkeit, größte Patientenzahl
IIb: Patienten ohne meßbare Chloridionenleitfähigkeit, Subgruppe von IIa, mit Gruppe I nach dem Alter stratifiziert
T: unverwandte, gesunde Kontroll-Probanden

Mann-Whitney IIa gegen I: $p=0,1$, einseitig
Mann-Whitney IIb gegen I: $p=$ n.s. einseitig
nach Referenz [76]

Abb. 8: Es sind die Alterswerte der einzelnen Studienpopulationen als Boxplots gezeigt. Die Darstellungen I, IIa und IIb repräsentieren die unterschiedliche Gruppen der Patienten von der Europäischen Mukoviszidose Zwillings- und Geschwisterstudie und die Darstellung T gibt die Werte der unverwandten gesunden Kontroll-Probanden wieder.

2.9.2 Monte-Carlo-Test zum Vergleich zweier Populationen bezüglich der Allelverteilung

Der klassische Ansatz, die Verteilung von Haplotypen oder Allelen zwischen zwei Studiengruppen zu vergleichen, besteht in der Aufstellung von Kontingenztafeln und deren Analyse mit dem Chi-Quadrat Test. Hier wird unter Annahme der Nullhypothese, daß die beiden Gruppen derselben Grundgesamtheit angehören, die Abweichung der beobachteten Allelverteilung von der theoretisch erwarteten bewertet.

Die Anwendung des Chi-Quadrat Testes ist jedoch bei erwarteten Zellzahlen kleiner als fünf unzulässig. Der traditionelle Weg zur Umgehung dieser Einschränkung besteht in der Zusammenlegung von Spalten der Kontingenztafel, sodaß die Zellzahl von fünf nicht unterschritten wird. Dies ist jedoch nur bei einer aus dem theoretischen Hintergrund des Experimentes abgeleiteten *a priori* Hypothese zulässig.

Aus diesem Grunde wurde hier für die statistische Analyse das Programm CLUMP eingesetzt. Dabei wird die stochastische Unabhängigkeit zweier Gruppen bezüglich der Verteilung von m Allelen unter Anwendung der Monte-Carlo-Simulation abgeschätzt [77].

Dieses Programm erzeugt eine große Anzahl von Tabellen mit $m \times 2$ Zellen, welche die selben Spalten- und Zeilensummen aufweisen, wie die Kontingenztafel der beobachteten Allelverteilung. Diese geschieht durch zufällige Verteilung der Allele auf die einzelnen Zellen. Des weiteren wird sowohl für die Kontingenztafel der gemessenen Allelverteilung als auch für die erzeugten Tabellen die Prüfgröße gemäß des Chi-Quadrat Testes ermittelt.

Der p -Wert der beobachteten Allelverteilung ergibt sich dabei durch Bestimmung des Anteiles derjenigen simulierten Tabellen, welche einen gleichen oder größeren Betrag für die Chi-Quadrat-Prüfgröße aufweisen, als die der im Experiment ermittelte Kontingenztafel.

Die Anzahl der zu simulierenden Tabellen kann dabei vorgegeben werden. Beginnend mit einem Wert von 1000, wurde bei kleinen p -Werten ($<20/1000$) die Vorgabe um den Faktor Zehn erhöht, um die statistische Signifikanz exakter abzuschätzen.

Das Programm CLUMP berechnet verschiedene Testcharakteristika, in dieser Arbeit wurde jedoch lediglich derjenige p -Wert für die weitere Auswertung verwendet, welcher nicht auf einer Zusammenlegung der ermittelten Kontingenztafel zu 2×2 Tabellen oder einer Subgruppierung von Allelen beruht.

2.10 DNA-Sequenzierung

Die Sequenzierung ist eine Methode zur Auslesung der Basenreihenfolge in einem vorgegebenen DNA-Abschnitt. Die Vorgehensweise zur Bestimmung des Zieles der Sequenzierung ist in Kapitel 3.2.4 des Ergebnisteiles beschrieben, hier erfolgt die Darlegung der technischen Methoden.

2.10.1 PCR für die Sequenzierung

Für eine genauere Untersuchung von Abschnitten im Genom nach dem Verfahren der Sequenzierung ist eine Amplifikation dieser Bereiche mittels einer PCR notwendig.

Dafür wurde die DAp-Goldstarpolymerase der Firma Eurogentec (Seraing, Belgien) gewählt. Sie ist aufgrund einer 3' → 5'-Exonuklease-Aktivität in der Lage, während der Synthese der komplementären Kopie des Templates falsch eingebaute Nukleotide zu erkennen, herauszuschneiden und nach Anknüpfung der korrekten Base mit der Synthese fortzufahren. Die Fehlerrate ist aufgrund dieser „proof-reading“-Aktivität 7 mal besser als die der gewöhnlichen Taq-Polymerase (Datenblatt, Eurogentec).

Es wurden Produkte mit einer Länge von 2965 Basenpaaren amplifiziert. Dies erforderte die Berücksichtigung zweier Punkte:

So ist (1) der Zeitbedarf, welche die Polymerase zur Synthese des Stranges über die ganze Länge benötigt, recht groß. Zudem ändert sich im Verlaufe der PCR-Reaktion das Verhältnis zwischen Template und Polymerase-Aktivität durch die steigende Anzahl der Kopien zugunsten des Templates hin.

Dieser Tatsache wurde Rechnung getragen durch eine lange Elongationszeit von 3 min, die ab dem 20. Zyklus schrittweise um 5 s verlängert wurde, sodaß im 35. Zyklus die Extensionszeit 4 min und 15 s betrug.

Zum anderen (2) sind DNA-Fragmente mit dieser Länge aufgrund der Einwirkung von Scherkräften und dadurch verursachten Strangbrüchen seltener als kürzere Abschnitte. Da dies insbesondere für Verdünnungen gilt, die sich schon länger in Gebrauch befinden, wurden die DNA-Stammlösungen für die hier durchgeführte PCR jeweils frisch verdünnt.

Protokoll:

10 µl	DNA (10 ng/µ, gelöst in 1xTE)	100 ng
5 µl	10x-Puffer	Eurogentec
3 µl	MgCl ₂	50 mM
5 µl	dNTP	2 mM jedes Nukleotid
5 µl	PrimerA	5 µM
5 µl	PrimerB	5 µM
0,5 µl	DAP Polymerase	4 U/µl, Eurogentec

Das Reaktionsgemisch wurde mit bidestilliertem Wasser zu einem Volumen von 50 µl ergänzt.

Die Primersequenzen sind aus Tabelle 16 des Anhangs zu entnehmen.

2.10.2 Prinzip der Sequenzier-Reaktion

Die Sequenzierung erfolgte nach dem Strangabbruch-Verfahren durch die Firma Qiagen (Hilden).

Nach diesem Verfahren wird das Template von einer Sequenzierpolymerase abgelesen, und es entsteht wie bei einer PCR eine Kopie. Durch die Verwendung von Didesoxynukleotiden, welche in 3'-Position keine OH-Gruppe besitzen, kann die DNA-Polymerase nach dessen Einbau die Synthese nicht fortführen, es kommt zum Strangabbruch. Diese Terminatoren werden mit unterschiedlichen Fluorophoren markiert und im Unterschluß zusammen mit gewöhnlichen Desoxynukleotiden der Reaktion hinzugegeben. So werden Sequenzprodukte mit unterschiedlicher Länge erstellt. Trägt man sie auf einem Gel elektrophoretisch auf, kann man mit Hilfe eines Laserfluorometers aus der Anordnung der Fluorophore auf die Basensequenz des Templates schließen. Ein Beispiel für eine solche Auswertung ist Abb. 9 zu entnehmen.

Exemplarische Darstellung von Spurdaten einer Sequenzierung nach dem Strangabbruchverfahren unter der Verwendung von Fluorophoren

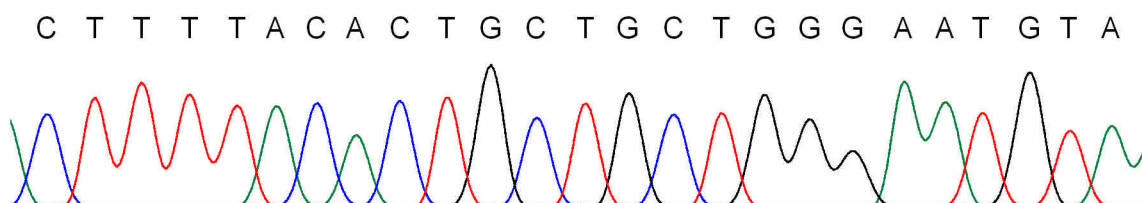


Abb. 9: Den detektierten Intensitätsmaxima kann der Wellenlänge der Fluorophoren entsprechend eine Base zugeordnet werden.

Aufgrund des limitierten Auflösungsvermögens der Gele und der Signalintensität der Reaktion wurden Sequenzierprimer mit Hilfe des Programms Primer3 [72] definiert, die in Abständen von etwa 500 bp auf dem codogenen und auf dem nichtcodogenen Template-Strang des zu sequenzierenden PCR-Produktes binden. Sie bilden jeweils einen Startpunkt der Sequenzierung, sodaß man überlappende Daten (Reads) sowohl in 3'→5' als auch 5'→3' Richtung des Templates herhält. Eine Übersicht der Anordnung dieser Primer gibt Abb. 10 wieder. Detaillierte Informationen sind der Tabelle 16 im Anhang zu entnehmen.

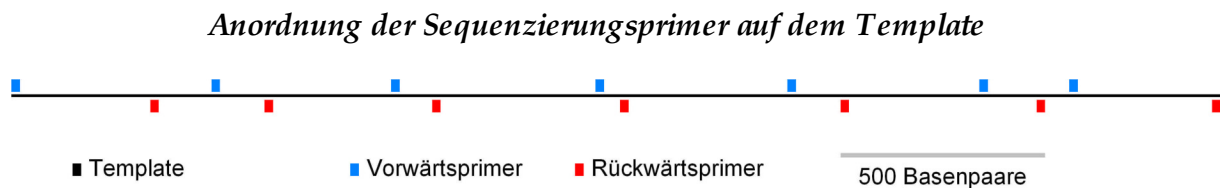


Abb. 10: Die Primer sind auf dem Template so angeordnet, daß die Hybridisierungs-Positionen nicht überlappend besetzt werden.

2.10.3 Zusammensetzung der Sequenzdaten zu einem vollständigen Datensatz

Die Auslesung der Sequenzdaten aus den Elektropherogrammen als auch die Abschätzung der Qualität erfolgte mit dem Phred-Algorithmus (persönliche Mitteilung: Dr. Henne, Qiagen).

Dabei wird angenommen, daß die in der Sequenzierung generierten Fragmente auf dem Gel in gleichen Abständen zueinander angeordnet sind [78]. Unter dieser Voraussetzung werden mittels Fourieranalysen die theoretischen Positionen der Fragmente berechnet. Als nächstes werden die Spurdaten der Elektropherogramme ausgelesen und so die Orte der tatsächlichen Intensitätsmaxima für die Sequenzprodukte bestimmt. Da jeder der vier Spuren eines Elektropherogrammes einer bestimmten Base zugeordnet ist, kann so die Sequenz abgeleitet werden. In einem weiteren Schritt werden diese durch verschiedene Methoden bestimmten Orte miteinander abgeglichen. Die vorhergesagten Positionen erlauben dabei eine leichtere und eindeutige Zuordnung der Basen in Bereichen mit gering aufgelösten Spurdaten.

Grundlage für die Zuordnung von Qualitätswerten einer jeden Base bildet die Analyse der umgebenden Spurdaten [79]. Dabei erfolgt eine Charakterisierung durch Berechnung von vier oder fünf Parametern. Eine Meßgröße dafür ist beispielsweise das Verhältnis des größten Intensitätsmaximums ohne zugeordneter Base zu dem kleinsten Maximum mit zugeordneter Base. Als nächstes erfolgt anhand einer Tabelle auf der Grundlage dieser Parameter eine Zuordnung der Qualitätswerte. Diese Rangliste ist durch Analysen von Sequenzdaten mit bekannten Fehlern erstellt worden, sie enthält die gemessenen Fehlerraten für verschiedene

Kombinationen der Parameter. Die so geschätzte Fehlerwahrscheinlichkeit wird nach folgender Formel skaliert:

$$Q = -10 \times \log_{10}(P) \quad (5)$$

P ist die geschätzte Fehlerwahrscheinlichkeit und Q der zugeordnete Qualitätswert. Große Qualitätswerte gehen mit einer geringen Fehlerwahrscheinlichkeit einher und kleine mit einer erhöhten Falschbestimmungsrate. Ein Wert von 20 bedeutet: In einem auf 100 Fällen liegt für die Base an dieser Position eine Falschbestimmung vor.

Die Zusammenstellung der einzelnen Sequenzdaten zu dem vollständigen Produkt erfolgte in Annotation dieser Qualitätswerte, d.h. lag für eine Position verschiedene Information vor, so wurde diejenige Base mit dem besseren Wert verwendet. Diese Assemblierung wurde mit dem Programm CodonCode Aligner (Fa. CodonCode Cooperation, USA) durchgeführt.

2.10.4 Analyse der zusammengesetzten Sequenzdaten

Mögliche Unterschiede zwischen den sequenzierten Proben wurden durch einen direkten Vergleich der assemblierten Sequenzen untersucht. Dazu wurde das Programm Bioedit [80] verwendet. Auf diese Weise erfolgte ebenfalls ein Vergleich mit der veröffentlichten Referenzsequenz.

Untersuchungen, die sich jeweils auf die Daten eines Patienten beschränkten, wurden mit Hilfe des Programms CodonCode Aligner durchgeführt.

SNPs mit heterozygotem Genotyp können in den Elektropherogrammen erkannt werden. In einem solchen Fall sind die etwa gleich großen Intensitätsmaxima für zwei verschiedene Nukleotide an derselben Position zu beobachten.

Mit Hilfe einer integrierten Programmfunktion wurden solche Positionen angezeigt und anschließend manuell überprüft. Sie wurden bestätigt, wenn in allen Elektropherogrammen mit Qualitätswerten um etwa 20 in der Umgebung des fakultativen SNPs diese doppelten Signale ersichtlich waren.

2.11 Funktionelle Analyse mit der Genome-Atlas-Software

In Kooperation mit Prof. David W. Ussery (The Technical University of Denmark, Center for Biological Sequence Analysis, Department of Biotechnology, Lyngby, Denmark) erfolgte eine bioinformatische Analyse des DNA-Abschnittes, welcher durch eine Sequenzierung (siehe Kapitel 2.10 und 3.2.4) genauer untersucht wurde.

Es wurden mit der DNA-Genome Atlas-Software folgende Parameter berechnet:

- ‚Basenstapelungsenergie‘ (Ornstein) [81]
- ‚Krümmung des DNA-Rückgrates‘ [82]
- ‚Stress induzierte DNA-Doppelstrang Öffnungswahrscheinlichkeit‘ [83]

2.12 Vorhersage von Transkriptionsfaktor-Bindungsstellen

Zur Vorhersage von Transkriptionsfaktor-Bindungsstellen wurde die Datenbank Genomatix [84] mit der Matrix-Family-Library Version 6.0 verwendet. Sie basiert auf dem Konzept der positionsgewichteten Matrix (pwm: position weight matrix). Dabei werden die für jeden Transkriptionsfaktor bekannten DNA-Bindungssequenzen in einer Matrix zusammengefasst und den Nukleotiden an jeder Position ein gewichteter Wert zugeordnet, welcher aus den beobachteten Häufigkeiten berechnet wird. Transkriptionsfaktoren mit ähnlicher Bindungssequenz sind als Familien gruppiert.

Mit Hilfe dieser Datenbank wurde der Einfluß von SNP-Allelen eines Sequenzabschnittes auf das Spektrum der an diesen DNA-Doppelstrang bindenden Transkriptionsfaktoren untersucht. Dazu wurden für jeden SNP des Sequenzabschnittes die Position und die Allele an die Datenbank gesandt und mit Hilfe der Teilfunktion MatInspector (Version 7.4.2) und SNPInspector (Version 2.0) untersucht.

Für alle einstellbaren Parameter wurden dabei die Standartvorgaben akzeptiert.

2.13 Abschätzung der Expression von Genen anhand von Transkriptomdaten aus Affymetrix-Chips

Microarrays dienen dazu, die RNA-Menge bestimmter Gene nachzuweisen.

Diese Microarrays sind Elemente, welche aus synthetisch hergestellten Oligonukleotiden bestehen, die auf einem Trägermaterial an definierten Positionen eines Rasters aufgetragen sind.

Diese Nukleotide werden als Sonden eingesetzt.

Je nach Art des verwendeten Microarrays wird die aus dem Gewebe extrahierte RNA zunächst in cDNA oder cRNA umgeschrieben, um anschließend eine Hybridisierung mit dem Array durchzuführen. So bleiben nach einem erfolgten Waschvorgang lediglich die zu den Sonden komplementären Nukleotid-Polymere der Probe haften. Da diese mit einem Fluoreszenzmarker gekoppelt sind, kann das beladene Microarray mit Hilfe einer Laserkamera ausgelesen werden und so die RNA-Menge quantitativ bestimmt werden.

Nach einer anschließenden Normierung des Primärsignales, welche u.a. die Qualität der RNA-Isolierung und Abbauvorgänge berücksichtigt, kann das Genexpressionsmuster des untersuchten Gewebes beschrieben werden.

In dieser Arbeit wurden Daten einer Microarray-Analyse verwendet, welche im Rahmen des Projektes SFB621-C7 erhoben worden sind. Im nachfolgenden werden die Details beschrieben, wie diese Daten ermittelt wurden (gemäß einer persönlichen Mitteilung von Dr. Andrea van Barneveld):

Für die Signal-Gewinnung wurden Chips der Firma Affymetrix mit der Produktbezeichnung HG-U133A eingesetzt. Hier besteht ein „Probenset“ für jedes Gen aus je 11 „Match- und Mismatcholigonukleotiden“. Insgesamt verfügen diese Chips über Datenpunkte von 22284 verschiedenen Genen.

Für die Hybridisierung wurde RNA aus Rektumschleimhautbiopsien isoliert.

Die dazu notwendigen Proben wurden an der Medizinischen Hochschule Hannover den Probanden der in Kapitel 2.9.1.2 beschriebenen Studienpopulation entnommen.

Aus dem Gewebe erfolgte zunächst eine RNA-Totalextraktion an der MHH, dann wurde an der „Transcript-Labeling-Einheit“ der Uni-Jena eine Umschreibung in cDNA und anschließend wurde eine Konvertierung in eine mit Biotin markierte cRNA vorgenommen.

Mit dieser erfolgte schließlich die Hybridisierung und Auslesung der Chips in der „Affymetrix Gene Chip Core Facility“ der Medizinischen Fakultät der Universität Leipzig.

Dafür wurde die Affymetrix-Microarray Suite v5.1 verwendet. Hier wurde zudem die Normalisierung der Daten durchgeführt.

Insgesamt konnten so von 8 gesunden Kontroll-Probanden die Expressionsdaten von 22284 verschiedenen Genen in der Rektumschleimhaut ermittelt werden.

Diese Daten wurden für die weitere Auswertung freundlicherweise von Dr. Andrea van Barneveld zur Verfügung gestellt.

2.14 Abschätzung der Gen-Expression mittels der UniGene-Datenbank

Zur Abschätzung der Gen-Expression wurde auch die UniGene-Datenbank [85] verwendet. Hier sind ESTs (expressed sequence tags) verwaltet, welche die Sequenz des 5'- oder 3'- Endes von cDNA-Klonen darstellen. Diese cDNA-Klone entsprechen dabei umgeschriebener mRNA von bestimmten Genen.

Nimmt man an, daß eine erhöhte Gen-Expression mit einer hohen mRNA-Kopienzahl einhergeht, kann man auf die Expressionsrate zurückschließen, indem man die Anzahl der cDNA-Klone oder ESTs bestimmt.

In der UniGene-Datenbank sind die ESTs nach den Genen zu Clustern zusammengefaßt und gemäß den verschiedenen untersuchten Geweben geordnet.

Um die relative Expressionsrate des betreffenden Genes im Kolon im Vergleich zu den übrigen Geweben zu beurteilen, wurde eine Normierung nach folgender Formel vorgenommen:

$$N_i = \frac{A_i / B_i}{\sum_{k=1}^{k=46} A_k / B_k} \quad \text{mit } i \in k \quad (6)$$

A_i : Anzahl der Klone des betreffenden Genes im Gewebe i

B_i : Anzahl aller verfügbaren Klone im Gewebe i

k : verschiedene nicht pathologisch veränderte Gewebe von adulten Probanden

N_i : normierte Häufigkeit des Klones vom betreffenden Gen im Gewebe i

Dieses Verfahren der Normalisierung zur Abschätzung der relativen Genexpression anhand von Daten aus der UniGene-Datenbank wurde bereits bei der SOURCE-Datenbank angewendet [86]. Grundlage der Auswertung war die Datenbank in der Version UniGene-build #199.

2.15 Methoden zur Synthese und Testung von Antikörpern gegen das Homolog hCLCA4

Im Rahmen dieser Arbeit war die Verfügbarkeit von Antikörpern gegen das Homolog hCLCA4 notwendig. Für eine genaue Begründung der Auswahl des Genproduktes hCLCA4 sei auf das entsprechende Kapitel im Ergebnisteil verwiesen. In diesem Abschnitt erfolgt lediglich eine Darlegung der technischen Methoden.

Die Synthese der polyklonalen Antikörper erfolgte in Kooperation mit der Firma Eurogentec (Seraing, Belgien). So wurde die Auswahl der Peptidepitope (Kapitel 2.15.1) und deren Synthese sowie die Immunisierung der Nutztiere (Kapitel 2.15.2) bei Eurogentec in Auftrag gegeben.

Die Antikörperbildung der Nutztiere wurde anhand von Immunodotblots kontrolliert (siehe Kapitel 2.15.3 bis 2.15.5).

2.15.1 Auswahl der Epitope zur Antikörper-Gewinnung

Eine Repräsentative der veröffentlichten Proteinsequenzen für das Homolog hCLCA4 (AAQ88834) wurde an die Firma Eurogentec gesendet. Unter Anwendung verschiedener Algorithmen wurden vier Bereiche der Aminosäuresequenz als Epitope vorgeschlagen, die für eine Immunisierung und damit zur Antikörpersynthese geeignet waren. Dabei wurde unter anderem die Oberflächenwahrscheinlichkeit und Antigenität der Sequenzabschnitte berechnet (persönliche Mitteilung: Dr. R. Corman, Eurogentec). Um die Eignung der empfohlenen Peptide für eine Immunisierung abzuschätzen, wurden anschließend verschiedene Analysen durchgeführt:

So wurde die Lage der vier Epitope I, II, III, und IV in einem Hydrophobizitätsprofil nach Kyte & Doolittle des Proteins überprüft (siehe Abb. 11). Überdies wurde mit den Proteinsequenzen der Homologe CLCA1 (NP_001276), CLCA2 (NP_006527), CLCA3 (NP_004912) und mit dem Genprodukt CLCA4 (AAQ88834) ein Alignment erstellt. Dies war Grundlage für die Abschätzung der zu erwartenden Kreuzreaktivität der hCLCA4-Antikörper zu den übrigen Homologen der Genfamilie. Dazu wurde zum einem die Position von Cystein-Resten in den Proteinen verglichen, diese sind bei der Disulfidbrückenbildung beteiligt und damit wichtig für die Stabilisierung von Tertiärstrukturen (siehe Abb. 13). Und zum anderen wurde der Grad der Homologie zwischen den Proteinsequenzen an den Stellen verglichen, welchen den zur Immunisierung empfohlenen Stellen entsprechen. Die Auswahl der zur Immunisierung verwendeten Epitope fiel insbesondere wegen der geringen Homologie zu den übrigen CLCA-Proteinen auf die Peptide I und IV (siehe Tabelle 2).

Alle Untersuchungen der Proteinsequenzen erfolgten unter Anwendung des Programmes Bioedit in der Version 7.0.1 [80]. Als Quelle für die verwendeten Proteinsequenzen wurde die NCBI-Datenbank [87] herangezogen.

Hydrophobizitätsplot des Proteins hCLCA4 – Hinweis für extramembranär gelegene Domänen

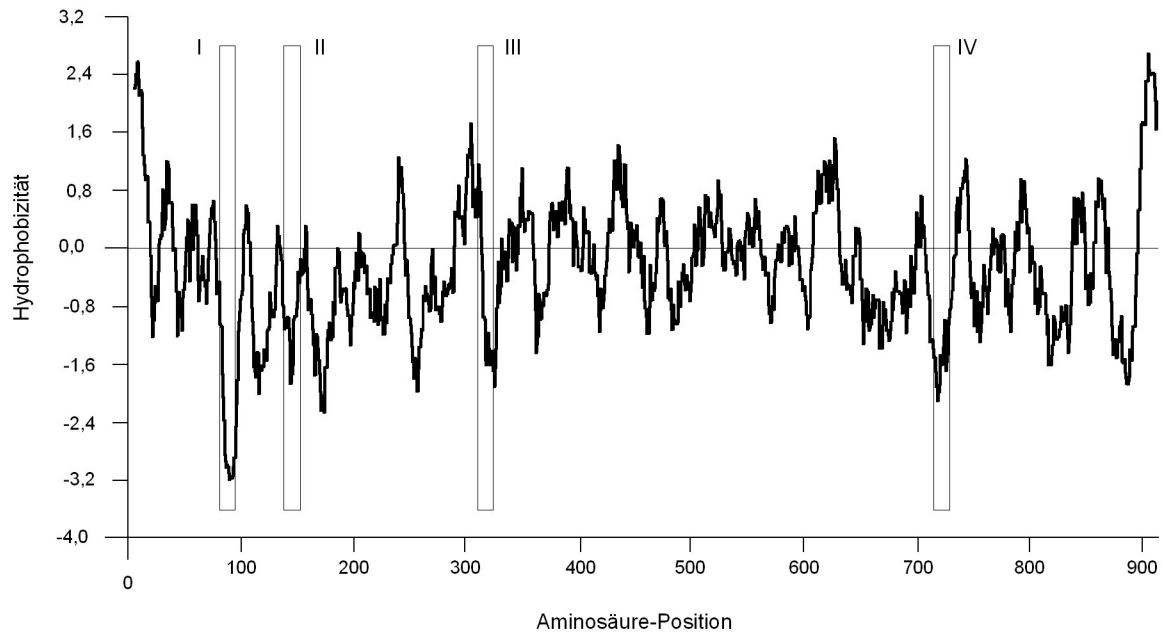


Abb. 11: Es ist die durchschnittliche Hydrophobizität in Abhängigkeit der Aminosäureposition dargestellt. Positive Werte weisen auf eine hydrophobe Domäne hin und negative geben Hinweise auf hydrophile Bereiche des Proteins.

Daneben sind die Positionen der zur Immunisierung vorgeschlagenen Peptide mit den Ziffern I, II, III und IV markiert. Sie liegen alle in hydrophilen Bereichen des Proteins.

Die Berechnung erfolgte nach dem Algorithmus von Kyte&Doolittle mit einer Fenstergröße von 13 Aminosäuren anhand der Proteinsequenz des Homologes hCLCA4 (AAQ88834).

Vergleich der verschiedenen für das Homolog hCLCA4 veröffentlichten Proteinsequenzen – Nachweis von Polymorphismen

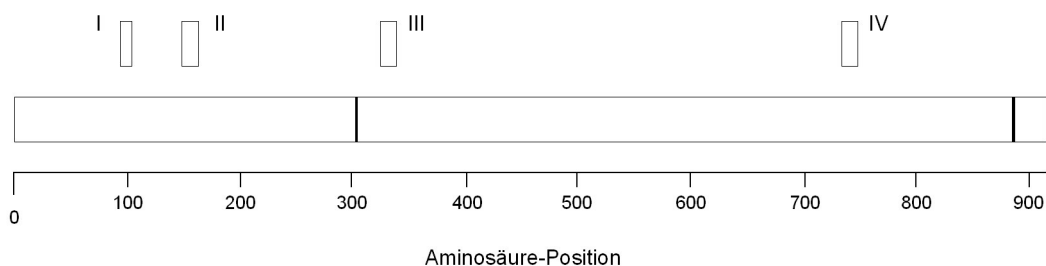


Abb. 12: Grundlage dieser Darstellung bildet ein Vergleich der für das Homolog hCLCA4 veröffentlichten Proteinsequenzen AAQ88834, CAI22170 und NP_036260. Unterschiede sind jeweils durch einen senkrechten Strich kenntlich gemacht. So steht in den ersten beiden Sequenzen an der Position 303 ein Serin, während in der letzteren ein Arginin vorhanden ist. Zudem tragen die ersten beiden Sequenzen an der Stelle 886-887 das Dipeptid Threonin-Prolin, wogegen bei der letzteren eine Deletion vorliegt.

Zudem sind die zur Immunisierung empfohlenen Peptide mit den Ziffern I, II, III und IV angegeben.

Positionen der Aminosäure Cystein in den Proteinsequenzen der vier homologen CLCA-Proteine – Hinweis für die Ausbildung und Lokalisation von Disulfidbrücken

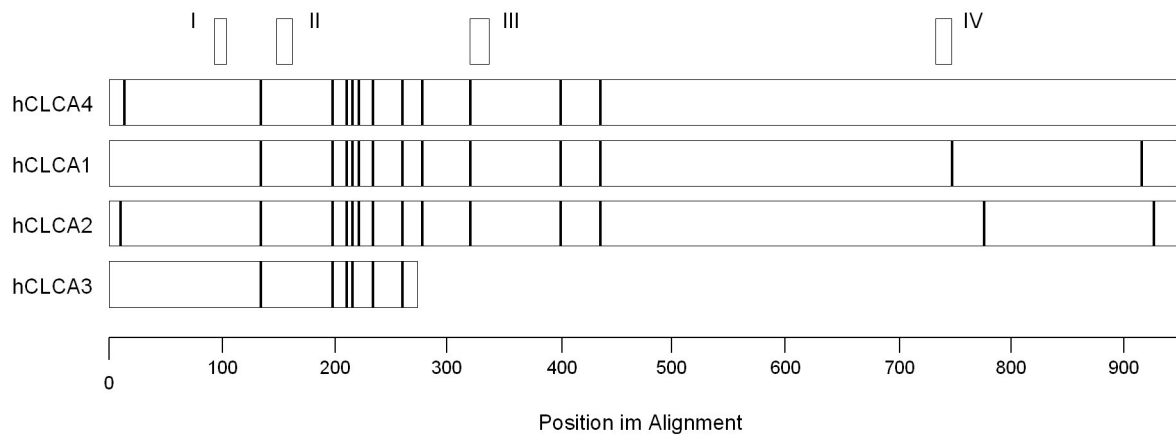


Abb. 13: Grundlage dieser Darstellung ist ein Alignment der Referenzsequenzen für die vier Homologe hCLCA4 (AAQ88834), hCLCA1 (NP_001276), hCLCA2 (NP_006527) und hCLCA3 (NP_004912). Die Position eines jeden Cystein-Restes in der Proteinsequenz ist jeweils durch einen senkrechten Strich markiert. Überdies sind die zur Immunisierung vorgeschlagenen Oligopeptide mit den Ziffern I, II, III und IV bezeichnet.

Das Verteilungsmuster der Aminosäure Cystein des Homologes hCLCA4 entspricht im wesentlichen dem der übrigen hCLCA-Proteine. Unterschiede gibt es am C-terminalen Ende, hier weisen die Varianten hCLCA1 und hCLCA2 zwei zusätzliche Cystein-Reste auf. Des weiteren fehlt beim Homolog hCLCA1 und hCLCA3 ein Cystein-Rest am N-terminalen Ende.

Homologie der zur Immunisierung ausgewählten Peptide I und IV zu den übrigen CLCA-Proteinen

	Aminosäuresequenz	Anteil gleicher Aminosäuren
Peptidsequenz I (84-96)	WKENPQYKRPKHC	
hCLCA4 (AAQ88834)	WKENPQYKRPKH	100%
hCLCA1 (NP_001276)	W KTKAD Y VRPKL	50%
hCLCA2 (NP_006527)	W KANN S KI~KQ	42%
hCLCA3 (NP_004912)	Y KSKSE Y L I PKQ	33%
	Aminosäuresequenz	Anteil gleicher Aminosäuren
PeptidsequenzIV (715-729)	PRPEID~EDTQTTLED C	
hCLCA4 (AAQ88834)	PRPEID~EDTQTTLED	100 %
hCLCA1 (NP_001276)	PRPEI NKDDVQ HKQ VC	47%
hCLCA2 (NP_006527)	PR KSVG~RNEEERKWG	13%
hCLCA3 (NP_004912)	nicht vorhanden	

Tabelle 2: Die Übersicht gibt die Aminosäuresequenzen der zur Immunisierung verwendeten Peptide I und IV wieder. Zusätzlich sind deren genauen Positionen im Homolog hCLCA4 (AAQ88834) vermerkt. Des weiteren sind die Sequenzen der übrigen Homologe in diesen Bereichen notiert. Dadurch werden die identischen Aminosäuren in diesen Abschnitten ersichtlich (fett gedruckt). Der prozentuale Anteil dieser ist in der rechten Spalte dargestellt.

2.15.2 Immunisierung der Kaninchen

Die Immunisierung der Nutz-Tiere inklusive Synthese der Peptid-Epitope (WKENPQYKRPKH und PRPEIDEDTQTILED) wurde bei der Firma Eurogentec (Seraing, Belgien) in Auftrag gegeben.

Die Peptide wurden nach Modifikation der endständigen aminogenen Gruppen über ein zusätzlich der Sequenz hinzugefügtes Cystein kovalent an KLH (keyhole-limpet-hemocyanin) gekoppelt, Carboxyl-Gruppen wurden dabei amidiert und die Aminogruppen acetyliert.

Zwei Kaninchen wurden parallel jeweils mit beiden Peptiden immunisiert.

2.15.3 Immuno-Dotblot – Kontrolle der Immunisierung

Zur Erfolgskontrolle der Immunisierung wurden Immuno-Dotblots durchgeführt.

Dabei wird das von den Tieren gewonnene Antiserum mit dem zur Immunisierung verwendeten Epitop inkubiert. Dieses ist auf einem Trägermaterial fixiert, sodaß spezifisch bindende (primäre) Antikörper an diesem immobilisiert werden.

Mittels eines sekundären Antikörpers, welcher gegen den primären gerichtet ist und einer an diesen gekoppelten Nachweisreaktion kann der spezifisch bindende Antikörper semi-quantitativ abgeschätzt werden.

2.15.4 Auftragen der für die Immunisierung verwendeten Peptide auf eine Membran

Die zur Immunisierung verwendeten Peptide wurden in bidestilliertem Wasser gelöst und in dekadischen Stufen verdünnt, sodaß jeweils mit dem Volumen von 1 µl 100 ng bis 0,001 ng Peptid auf die Nitrocellulosemembran Hybond ECL (Firma Amersham-Biosciences) aufgetragen werden konnte. Nach Trocknung bei Raumtemperatur wurden diese in Folie eingeschweißt und bis zur weiteren Verwendung bei 4°C zwischengelagert.

2.15.5 Durchführung der Immundetektion

Die Durchführung erfolgte nach dem Protokoll des ECL-Western Blot detection Kit der Firma Amersham Biosciences. Dieses Nachweissystem basiert auf einer Chemilumineszenz-Reaktion: Luminol wird in Anwesenheit von H₂O₂ durch Merrettich-Peroxidase (HRP) oxidiert, dabei wird Licht mit der Wellenlänge von 428 nm emittiert. Das Enzym HRP ist an den sekundären Antikörper gekoppelt. Folglich kann man aufgrund der Schwärzung eines Röntgenfilmes auf die Anwesenheit und Lokalisation der primären Antikörper schließen.

Alle nachfolgend beschriebenen Schritte erfolgten bei Raumtemperatur:

Zunächst wurden die Dotblot-Streifen für 1 h in Blocklösung auf einer Schwenkwippe äquilibriert.

Anschließend folgte für 2 Stunden eine Inkubation der Membranen mit den entsprechenden Antiseren. Diese wurden im Verhältnis von 1:750 mit Blocklösung verdünnt.

Nach dem Verwerfen der Antiseren folgten 3 Waschkvorgänge. Dazu wurden die Streifen für jeweils 10 min mit T-TBS inkubiert.

Jetzt wurde der sekundäre Antikörper (Anti-rabbit IgG Horseradish Peroxidase linked F(ab')₂-fragment from donkey, Firma: Amersham Biosciences) im Verhältnis von 1:5000, ebenfalls verdünnt in Blocklösung, hinzugegeben und mit den Streifen für 1 h inkubiert.

Es folgten 5 Waschkvorgänge mit T-TBS (s.o.).

Nachweisreaktion: Die beiden Reagenzien des ECL-Detection Kit wurden in gleichen Volumenverhältnissen miteinander vermischt, die Streifen aus der Waschlösung entnommen, für 50 s unter leichtem Schwenken mit dem Reagenz vollständig benetzt, dann in Folie frei von Luftblasen eingeschlagen und die Lumineszenz mit einem Röntgenfilm detektiert.

3 Ergebnisse

3.1 Charakterisierung der typisierten Marker

In diesem Abschnitt erfolgt eine Beschreibung der typisierten Marker, welche zur Analyse der genetischen Region der CLCA-Proteine verwendet wurden. Beginnend mit einer Darstellung der physikalischen Lokalisation innerhalb des *CLCA*-Clusters (Kapitel 3.1.1), erfolgt in Kapitel 3.1.2 ein Vergleich der aus der Datenbank ermittelten Allelfrequenzen mit den gemessenen Werten. In Kapitel 3.1.3 und 3.1.4 werden qualitative Eigenschaften der Marker wie die Informativität sowie die Vereinbarkeit mit dem Hardy-Weinberg-Gleichgewicht analysiert.

Abschließend sind in Kapitel 3.1.5 die Allelfrequenzen sowie die Informativität der Marker zusammengefaßt, welche die Grundlage für die Abbildungen in den vorangehenden Kapiteln sind.

3.1.1 Lokalisation der Marker im *CLCA*-Gen-Cluster

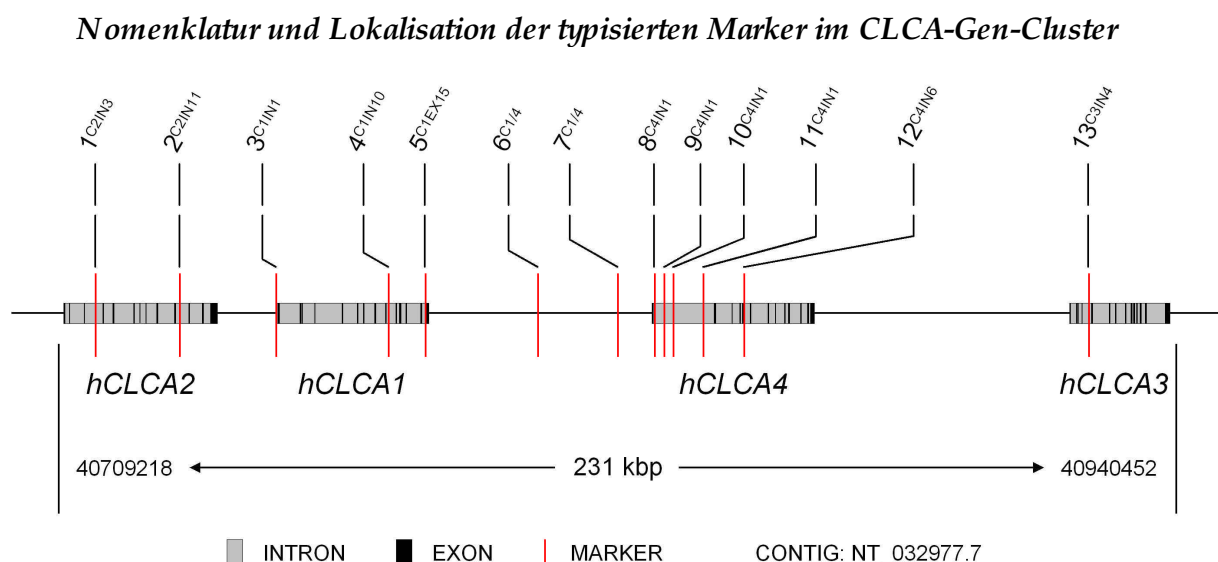


Abb 14: Es sind die Positionen der Marker im *CLCA*-Gen-Cluster eingezeichnet. Die Bezeichnung der Marker setzt sich zusammen aus einer vorangestellten Seriennummer, der dann folgenden Abkürzung des Gens, in welchem sich der Marker befindet und ein hinten angestelltes Kürzel für das Exon oder Intron des jeweiligen Gens.

Die Marker wurden der öffentlichen SNP-Datenbank [67] nach den in Kapitel 2.3.2 des Methodenteiles beschriebenen Auswahlkriterien entnommen. So sind für 19 verschiedene SNPs Primer entworfen und passende Restriktionsenzyme ausgewählt worden.

Durch Optimierung der PCR und Durchführung des diagnostischen Verdaues anhand von unverwandten Kontroll-Proben konnten schließlich fünf dieser Marker für eine Typisierung des Studienkollektives etabliert werden. Die übrigen SNPs waren aufgrund einer geringen Informativität (PIC-Wert kleiner als 0,25), einem schlecht lesbaren Muster des Restriktionsverdaues oder nicht optimierbarer Primerpaare für eine Typisierung der Studienpopulation nicht geeignet.

Gemeinsam mit den bereits vor der Anfertigung dieser Arbeit typisierten und grob analysierten Markern standen so insgesamt 13 verschiedene SNPs zur Verfügung.

Die Positionen der Marker sind in Abb. 14 dargestellt, 11 von diesen sind intronisch, zwei intergenisch und einer exonisch gelegen.

Der in Exon 15 des Genes *bCLCA1* lokalisierte Marker 5^{C1EX15} verändert zwar das kodierende Basentriplett, aber aufgrund des degenerierten genetischen Codes wird kein Aminosäurepolymorphismus verursacht.

3.1.2 Vergleich der veröffentlichten und gemessenen Allel-Frequenzen

In Abb. 15 erfolgt für jeden Marker ein Vergleich der veröffentlichten Allelfrequenz mit der gemessenen, indem jeweils die Häufigkeiten des selteneren Alleles gegeneinander aufgetragen sind.

Während für die Werte aus der Datenbank der Stand zum Zeitpunkt der Etablierung der jeweiligen Marker verwendet (dbSNP build 119 und 126) wurden, sind die gemessenen Werte anhand der Elterngeneration als unverwandte Individuen (n=54) erhoben worden.

Eine relativ geringe Abweichung und damit eine nachvollziehbare Angabe der veröffentlichten Allelfrequenzen ist bei den Markern der Gruppe 2 gegeben. Hier beträgt die Differenz der Frequenzen maximal 0,1.

Bei fünf Markern ist die Abweichung der Frequenzen größer als 0,15 (Gruppe 1). Bei drei von diesen SNPs stimmt zudem die qualitative Angabe der Base für das seltenere Allel zwischen der Datenbank und den durch die Typisierung der Elterngeneration erhobenen Daten nicht überein. Daher wird bei diesen Markern die für das seltenere Allel geltende theoretische Grenze von 0,5 überschritten.

Für den Marker 11^{C4IN1} sind keine Angaben über Allelfrequenzen veröffentlicht. Dieser wurde abweichend von den in Kapitel 2.3.2 beschriebenen Kriterien ausgewählt. Der SNP liegt in

einer Region des *CLCA*-Clusters, für die keine typisierbaren Marker mit einer Heterozygotität größer als 0,3 katalogisiert sind. Daher wurden Rezepte für fünf Marker etabliert und die Informativität anhand von acht unverwandten Proben ermittelt. Von diesen konnte lediglich bei Marker 11^{C4IN1} für das seltene Allel eine hohe Frequenz von 0,43 nach dem Maximum (0,5) für hochinformativ diallelische Marker gemessen werden.

Vergleich der gemessenen mit der veröffentlichten Frequenz des jeweils selteneren Alleles der typisierten Marker

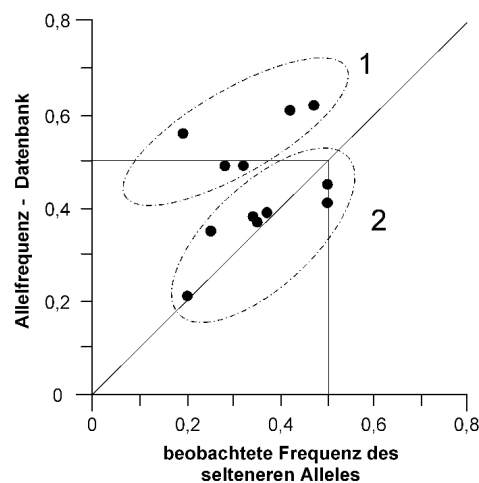


Abb. 15: Es sind die beobachteten Allelfrequenzen der SNPs mit den veröffentlichten verglichen. Punkte, welche oberhalb der Diagonalen liegen, zeigen im Vergleich zur gemessenen Frequenz des selteneren Alleles einen zu großen Betrag an, Punkte unterhalb der Diagonalen geben einen Hinweis für zu kleine Werte in der Datenbank. Die vertikale und horizontale Linie gibt die theoretische Grenze (0,5) für die Frequenz des selteneren Alleles an.

3.1.3 Informativität der Marker

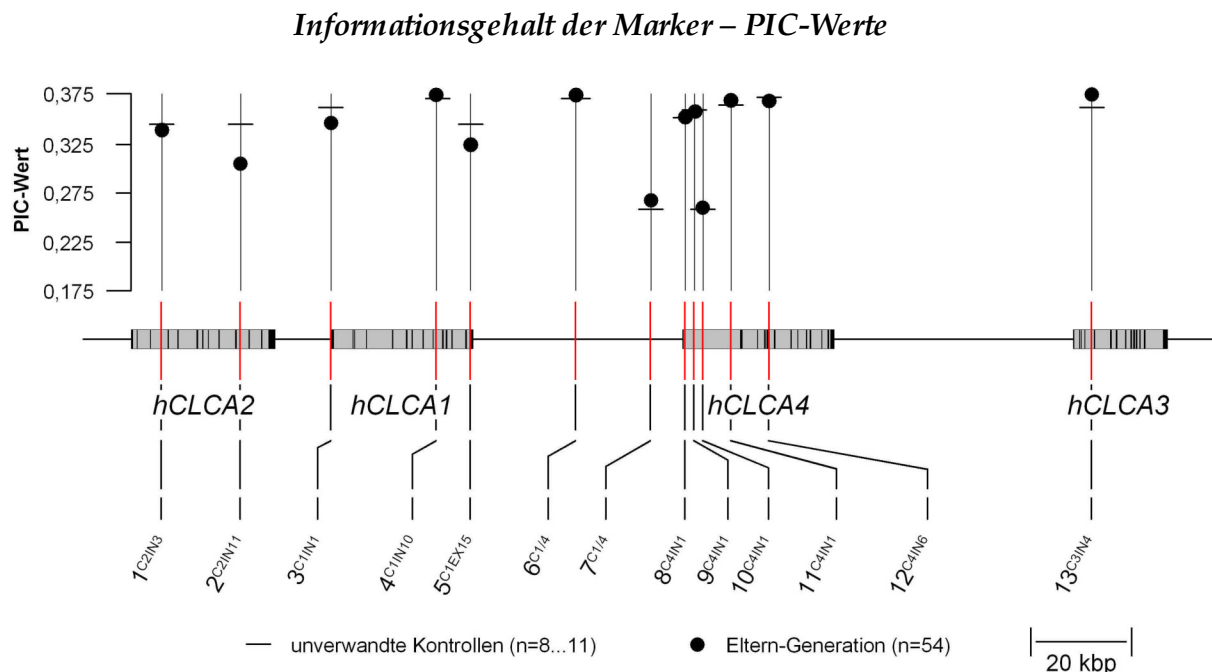


Abb. 16: Darstellung des Polymorphismus-Informationsgehaltes der typisierten Marker im *CLCA*-Gen-Cluster

In Abb. 16 sind die PIC-Werte der typisierten Marker dargestellt. Sie sind anhand der Elterngeneration (n=54) der Studienpopulation erhoben worden. Bei den vier SNPs 4^{C1IN10} , $6^{C1/4}$, 11^{C4IN1} , 12^{C4IN6} und 13^{C3IN4} wird das Maximum von 0,37 für diallelische Marker erreicht. Die SNPs $7^{C1/4}$ und 10^{C4IN1} sind mit Werten von 0,27 und 0,26 weniger informativ und die übrigen acht SNPs haben PIC-Werte größer als 0,3.

Betrachtet man vergleichend die anhand der unverwandten Kontrollen (n=8 ... 11) erhobenen Werte mit denen der Studienpopulation (n=54), so wird nur bei Marker 2^{C2IN11} eine größere Abweichung ersichtlich. In der Mehrzahl der Fälle stimmen die PIC-Werte, erhoben anhand der kleinen Stichprobe, mit denen der Elterngeneration recht gut überein.

Ein Test der Marker an einer kleinen Population zur Einschätzung der Informativität liefert somit valide Ergebnisse.

3.1.4 Frequenzen der Genotypen – Hardy-Weinberg-Gleichgewicht

Genotypfrequenzen und Vereinbarkeit der Marker mit dem Hardy-Weinberg-Gleichgewicht

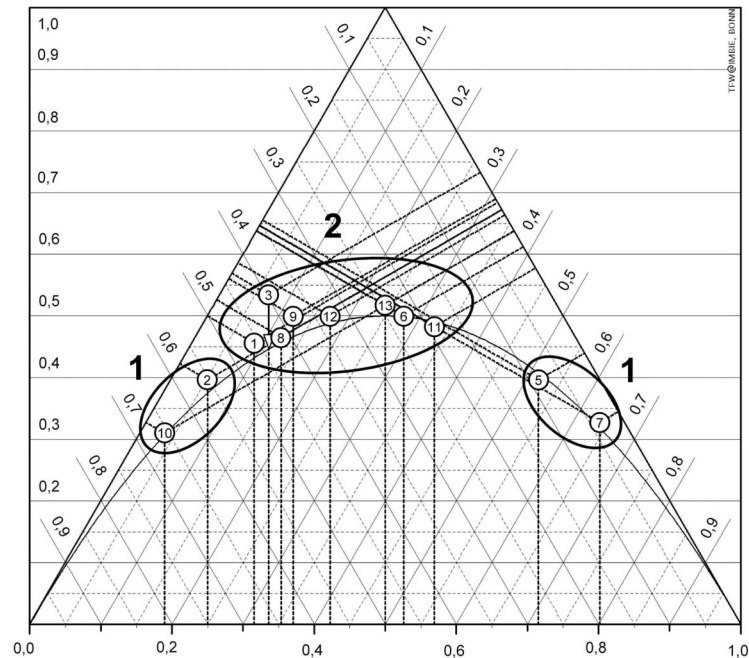


Abb. 17: Jeder Punkt innerhalb des Diagrammes repräsentiert einen diallelischen Marker. Die Lage wird durch drei Koordinaten festgelegt, diese entsprechen den Anteilen der Genotypen. Die Frequenz der heterozygoten Genotypen wird anhand der Ordinate aufgetragen, und für die Häufigkeit der homozygoten Genotypen werden die äußeren Seiten des Dreieckes als Achsen verwendet.

So läßt sich durch Projektion des Lotes auf die Ordinate der Anteil der Heterozygoten bestimmen. Durch Projektion des linken Lotes nach oben auf die rechte Seite des Dreieckes, kann die Frequenz des Genotypes 1-1 abgelesen werden. Analog gilt dies für die Häufigkeit des Genotypes 2-2, diese wird anhand der Skala der linken Seite bestimmt.

Für den Marker 2^{C2IN11} (entspricht Symbol 2) beispielsweise sind die Frequenzen der Genotypen 1-1, 1-2 und 2-2 mit (0,05/0,4/0,55) zu entnehmen.

Anhand der Position des vertikalen Lotes auf der Abszisse kann die Frequenz von Allel 1 bestimmt werden. Und indirekt über die Beziehung $p(a_1) + p(a_2) = 1$ (siehe Kapitel 2.3.1) kann der Anteil von Allel 2 abgeleitet werden.

Die nach unten offene Parabel gibt die Koordinaten von Markern an, die sich aufgrund der Verteilung ihrer Genotypen im Hardy-Weinberg Gleichgewicht befinden. Eine mögliche Abweichung davon ist für die eingetragenen Marker mit einer durchgehend gezeichneten vertikalen Linie kenntlich gemacht.

Diese Form der Darstellung geht auf de Finetti (1926) zurück. Für die Anfertigung wurde das Programm „FINE’TTI“ von Prof. T. Winker (Universität Bonn, IMBIE, Abt. Genetische Epidemiologie) verwendet.

In Abb. 17 ist ein de Finetti Diagramm gezeigt. Jeder Punkt repräsentiert einen dimorphen Marker. Wie in der Legende beschrieben, werden als Koordinaten die beobachteten Genotypfrequenzen verwendet, und Marker, die sich aufgrund der Verteilung ihrer

Genotypen im Hardy-Weinberg-Gleichgewicht befinden, liegen auf der nach unten offenen Parabel.

Die Lage der SNPs ist also durch die Daten, gewonnen anhand der Elterngeneration der Studienpopulation, bestimmt.

Die Punkte lassen sich zu zwei Gruppen zusammenfassen:

Gruppe 2 liegt etwa in der Mitte des Koordinatensystems, die Frequenz der Heterozygoten beträgt ungefähr 0,45 bis 0,53. Diese Marker entsprechen den hoch informativen mit PIC-Werten größer als 0,34 (siehe Abb. 16).

Dann ist noch die weiter außen liegende Gruppe 1 zu sehen. Sie umfaßt die Marker 2^{C2IN11} und 5^{C1EX15} mit einem Heterozygoten-Anteil von etwa 0,4 und die Marker 10^{C4IN1} und $7^{C1/4}$ mit der Heterozygotenfrequenz von ungefähr 0,3. Diese letzteren genannten Marker haben den geringsten Informationsgehalt von 0,26 und 0,27 (siehe Abb. 16).

Setzt man voraus, daß für kein Allel der typisierten Marker ein Selektionseffekt vorliegt und für die untersuchte Population die Bedingung der Panmixie gegeben ist, dann sollten die tatsächlichen Genotypfrequenzen mit dem Hardy-Weinberg-Gleichgewicht vereinbar sein.

Wie Abb. 17 zu entnehmen ist, zeichnen die Punkte aller Marker inklusive des SNP 3^{C1IN1} die Parabel nach, sodaß für keinen Marker eine Divergenz vom Hardy-Weinberg vorliegt. Auch für Marker 3^{C1IN1} liegt keine signifikante Abweichung vor ($p=0,13$).

Damit können systematische Fehler bei der Genotypisierung wie sie etwa durch einen Partial- oder Fehlverdau der PCR-Produkte aufgrund von Aktivitäts- oder Spezifitätsverlusten des eingesetzten Enzymes auftreten können, weitgehend ausgeschlossen werden. Denn in einem solchen Fall würde der Anteil der Proben, bei denen ein heterozygoter Genotyp ermittelt werden würde, erheblich von der erwarteten Frequenz abweichen.

3.1.5 Allelfrequenzen und Informativität der etablierten Marker

Allelfrequenzen und Informativität der etablierten Marker

Marker	Quelle	Position	Allele (1/2)	f(a ₁)	Heterozygotität (erwartet)	Heterozygotität (beobachtet)	PIC
1 ^{C2IN3}	rs3737672	40715934	A / G	A: 0,32	0,43	0,46	0,34
2 ^{C2IN11}	rs3765986	40733565	C / T	C: 0,25	0,38	0,40	0,30
3 ^{C1IN1}	rs2791519	40753569	A / C	A: 0,34	0,45	0,53	0,35
4 ^{C1IN10}	rs2791498	40777017	A / C	A,C: 0,50	0,50	0,52	0,38
5 ^{C1EX15}	rs1882753	40784812	T / C	C: 0,28	0,41	0,40	0,32
6 ^{C1/4}	rs2791473	40808102	T / C	C: 0,47	0,50	0,50	0,37
7 ^{C1/4}	rs1321681	40824795	T / G	G: 0,20	0,32	0,33	0,27
8 ^{C4IN1}	rs1321680	40832538	A / G	A: 0,35	0,46	0,47	0,35
9 ^{C4IN1}	rs974843	40834572	C / T	C: 0,37	0,47	0,50	0,36
10 ^{C4IN1}	rs972764	40836486	A / G	A: 0,19	0,31	0,31	0,26
11 ^{C4IN1}	rs6684219	40842728	T / G	G: 0,43	0,49	0,48	0,37
12 ^{C4IN6}	rs772602	40851323	A / G	A: 0,42	0,49	0,50	0,37
13 ^{C4IN4}	rs1517899	40923251	G / A	A,G: 0,50	0,50	0,52	0,38

Tabelle 3:

- Quelle: Zugriffsnummer des Markers in der SNP-Datenbank [67]
- Position: Die Positionsangaben beziehen sich auf den Contig: NT_032977.7
- Allele (1/2): Die Basen der jeweiligen Allele sind gemäß der dbSNP notiert.
- f(a₁): Frequenz des jeweils selteneren Alleles
- Heterozygotität – erwartet: Anteil der heterozygoten Individuen bei Gültigkeit des Hardy-Weinberg-Gleichgewichtes, berechnet nach Formel (3), siehe Kapitel 2.3.1
- Heterozygotität – beobachtet: Diese Werte beruhen auf einer Auszählung der - Elterngeneration (n=54)
- PIC: Polymorphismus-Informationsgehalt, berechnet gemäß Formel (4), siehe Kapitel 2.3.1

In Tabelle 3 sind für die typisierten Marker die Zugriffsnummern in der SNP-Datenbank [67] aufgelistet. Des Weiteren ist hier für jeden Marker definiert, welche Base dem Allel 1 und 2 entspricht. Zudem sind hier die Allelfrequenz des jeweils selteneren Allels sowie der Polymorphismus Informationsgehalt aufgelistet, welche durch Typisierung der Elterngeneration ermittelt worden sind. Überdies kann eine mögliche Abweichung der Marker vom Hardy-Weinberg-Gleichgewicht anhand der Differenz zwischen der erwarteten und beobachteten Heterozygotität abgeschätzt werden.

3.2 Ergebnisse der Assoziationsstudie: Vergleich von CF-Patienten mit alternativer Chloridionenleitfähigkeit mit Patienten ohne Chloridionenleitfähigkeit im *CLCA*-Gen-Cluster

In der Assoziationsstudie wurden Patienten, bei denen eine alternative Chloridionenleitfähigkeit nachgewiesen werden konnte (Gruppe I), mit Patienten ohne residuelle Chloridionenleitfähigkeit (Gruppen IIa und IIb) im *CLCA*-Gen-Cluster verglichen. Dazu wurden die in Kapitel 3.1 beschriebenen Marker bei den Patienten typisiert und die Verteilung der Allele zwischen den beiden Gruppen verglichen.

Dieser Studie liegt das in Kapitel 2.9.1.1 beschriebene Patientenkollektiv zugrunde, welches mittels der in Kapitel 2.2 beschriebenen Methode der intestinalen Ionenstromfluß-Messung nach dem Phänotyp klassifiziert wurde. Für die Auswertung ist das in Kapitel 2.9.2 beschriebene Verfahren der p -Wert-Schätzung mittels Monte-Carlo-Simulation angewendet worden.

Die Resultate der Allelverteilung bei der Analyse von einzelnen Markern ist in Kapitel 3.2.1 beschrieben und in Kapitel 3.2.2 sind die Ergebnisse der Auswertung bei Kombination der SNPs zu Zwei-Marker-Haplotypen dargelegt.

Überdies sind in Kapitel 3.2.3 die stammesgeschichtlich fixierten Rekombinationsbruchpunkte innerhalb des *CLCA*-Gen-Clusters dargestellt.

Darüberhinaus erfolgte eine Teilresequenzierung des *CLCA*-Gen-Clusters. In Kapitel 3.2.4 wird auf der Grundlage der vorliegenden Ergebnisse aus der Assoziationsstudie die Auswahl des Zielbereiches und die dafür verwendeten DNA-Proben erläutert.

3.2.1 Auswertung der Assoziationsstudie: Vergleich der CF-Patienten an einzelnen Loci

In Abb. 18 und Tabelle 4 sind die Ergebnisse der Assoziationsstudie dargestellt, welche durch die Betrachtung der Marker als einzelne Loci ermittelt wurden. Dabei wurden die nach dem Alter stratifizierte Patientengruppen (I und IIb) miteinander verglichen.

Vergleich der nach dem Alter stratifizierte Gruppen an den einzelnen Loci

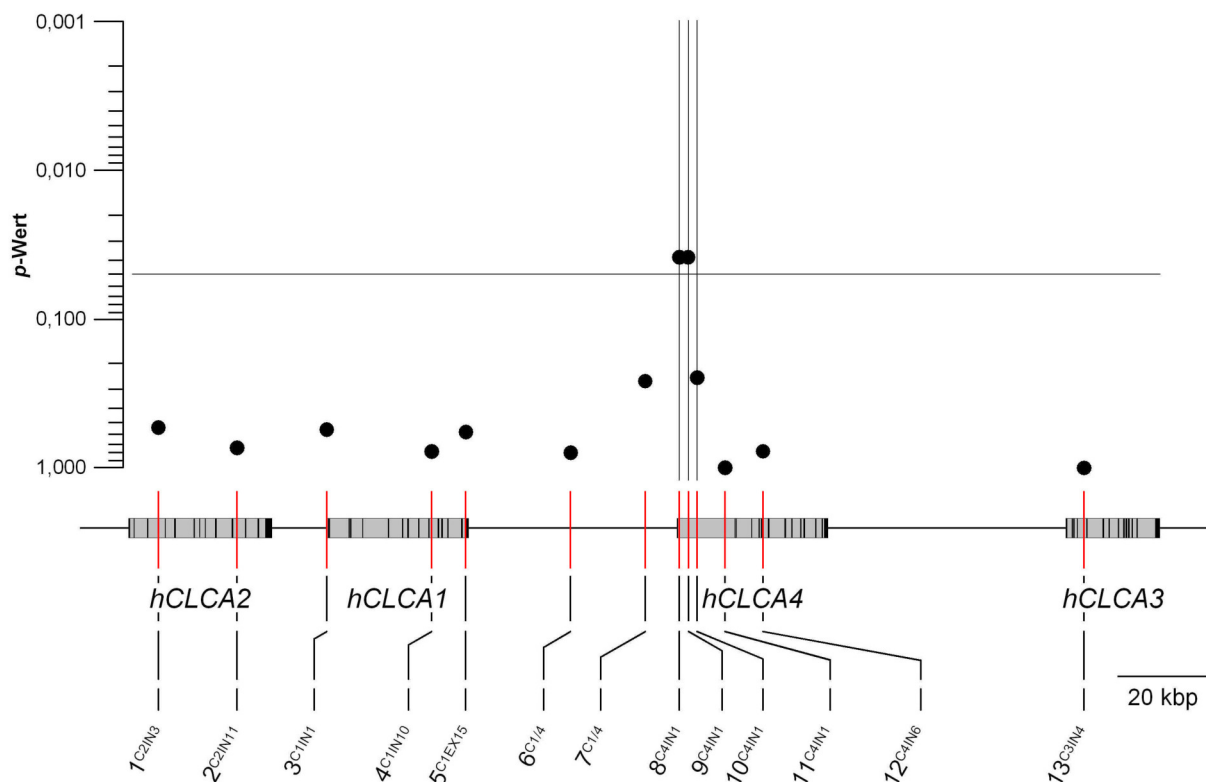


Abb. 18: Es sind die berechneten p -Werte an den Positionen der einzelnen Marker im *CLCA*-Gen-Cluster dargestellt. Die drei vertikalen Linien markieren zur besseren Übersichtlichkeit die Positionen der Marker 8^{C4IN1}, 9^{C4IN1} und 10^{C4IN1}, die horizontale Linie gibt das Signifikanz-Niveau für $\alpha=0,05$ an.

Für die Mehrzahl der Marker läßt sich kein signifikanter p -Wert berechnen, lediglich die SNPs 8^{C4IN1} und 9^{C4IN1} überschreiten mit Werten von $p=0,039$ das Niveau von $\alpha=0,05$.

Für die beiden am nächsten liegenden Marker 7^{C1/4} und 10^{C4IN1} ist der p -Wert mit 0,26 und 0,25 zwar nicht mehr signifikant, gibt aber möglicherweise einen Trend an.

Somit sind lediglich die Allele der Marker 8^{C4IN1} und 9^{C4IN1} zwischen den Studiengruppen signifikant unterschiedlich verteilt. Diese SNPs liegen im ersten Intron des Genes *bCLCA4*, während die beiden benachbarten Marker 7^{C1/4} und 10^{C4IN1} in der Promoter-Region dieses Genes sowie weiter 3' innerhalb des ersten Introns gelegen sind.

Altersstratifizierte Gruppen – beobachtete Allele der typisierten SNPs und berechnete p -Werte

Marker	Patienten-Gruppe	Allel		p -Wert
		1	2	
1 ^{C2IN3}	IIb	6	20	0,54
	I	9	17	
2 ^{C2IN11}	IIb	5	21	0,74
	I	7	19	
3 ^{C1IN1}	IIb	7	19	0,55
	I	10	16	
4 ^{C1IN10}	IIb	12	14	0,78
	I	10	14	
5 ^{C1EX15}	IIb	18	8	0,57
	I	15	11	
6 ^{C1/4}	IIb	15	11	0,78
	I	13	13	
7 ^{C1/4}	IIb	24	2	0,26
	I	20	6	
8 ^{C4IN1}	IIb	5	21	0,039
	I	13	13	
9 ^{C4IN1}	IIb	5	21	0,039
	I	13	13	
10 ^{C4IN1}	IIb	2	24	0,25
	I	6	20	
11 ^{C4IN1}	IIb	17	9	1,0
	I	17	9	
12 ^{C4IN6}	IIb	10	16	0,77
	I	8	16	
13 ^{C3IN4}	IIb	13	11	1,0
	I	15	11	

Tabelle 4: Getrennt nach den beiden Studiengruppen I und IIb (altersstratifiziert) sind für jeden Marker die Anzahl der Allele aufgelistet. In der rechten Spalte sind die daraus berechneten p -Wert notiert, welche Grundlage der Darstellung in Abb. 18 sind.

3.2.2 Auswertung der Assoziationsstudie: Vergleich der Verteilung von Zwei-Marker Haplotypen zwischen den CF-Patienten

In diesem Kapitel erfolgt ein Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den Patienten mit alternativer Chloridionenleitfähigkeit und ohne nachweisbarer Chloridionenleitfähigkeit. Grundlage für die Konstruktion der Haplotypen bilden die in Kapitel 3.1 beschriebene Marker.

Die Auswertung erfolgt bei einer unterschiedlichen Auswahl der Probanden:

So wurden in Kapitel 3.2.2.1 alle verfügbaren Patienten ohne meßbare Chloridionenleitfähigkeit (Gruppe IIa) mit den Patienten mit residueller Chloridionenleitfähigkeit (Gruppe I) verglichen, während in Kapitel 3.2.2.2 lediglich eine Subgruppe der Probanden ohne Chloridionenleitfähigkeit (IIb) zum Vergleich herangezogen wurde. Hier erfolgte ein Matching nach dem Alter, sodaß beide miteinander verglichenen Studiengruppen gleich groß sind und eine Schichtung nach dem Alter vorliegt.

3.2.2.1 Gesamtes Studienkollektiv: Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den Gruppen der CF-Patienten mit residueller Chloridionenleitfähigkeit und ohne Chloridionenleitfähigkeit

Gesamtes Studienkollektiv: Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den CF-Patienten mit alternativer (I) und ohne nachweisbare Chloridionenleitfähigkeit (IIa)

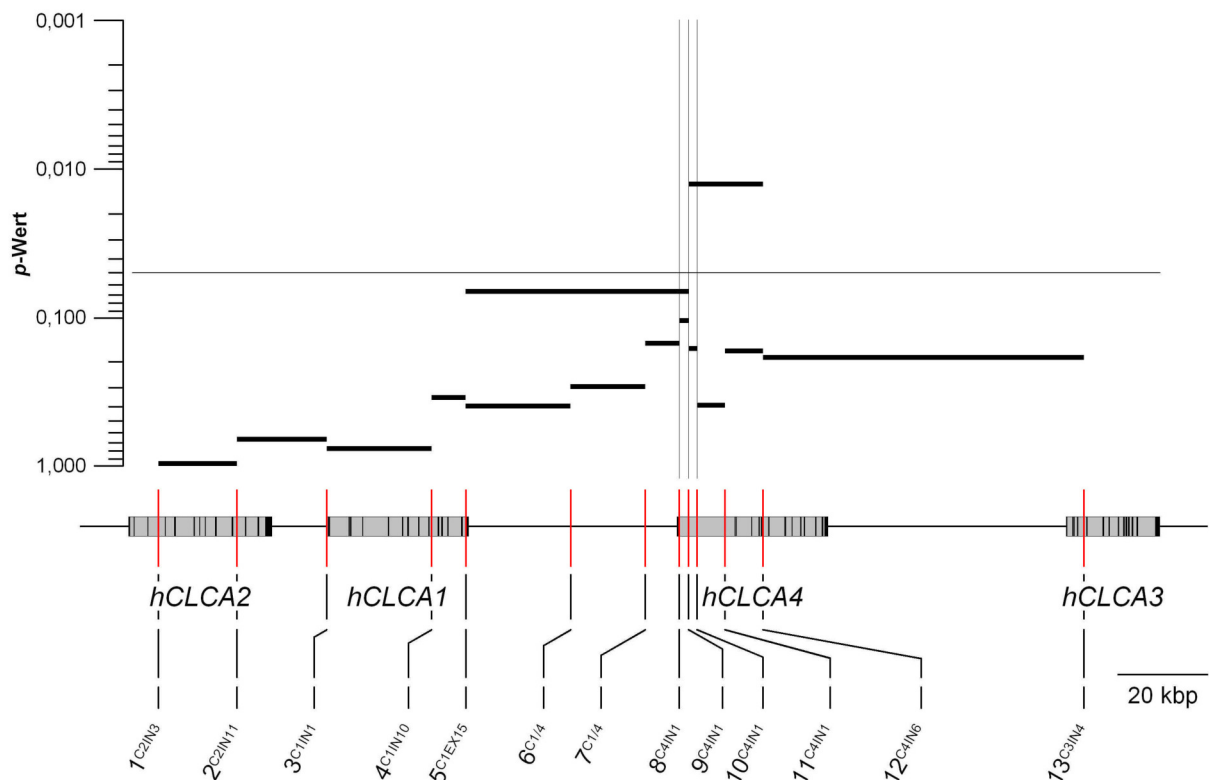


Abb. 19: Es sind die berechneten p -Werte an den Positionen der einzelnen Zwei-Marker-Haplotypen im *CLCA*-Gen-Cluster dargestellt. Die drei vertikalen Linien markieren zur besseren Übersichtlichkeit die Positionen der Marker 8^{C4/N1}, 9^{C4/N1} und 10^{C4/N1}, die horizontale Linie gibt das Signifikanz-Niveau für $\alpha=0,05$ an.

Die Haplotypen lassen sich anhand des Signifikanz-Niveaus in zwei Gruppen einteilen. So ist der Haplotyp 9^{C4/N1}_12^{C4/N6} mit einem p -Wert von 0,013 zwischen den Gruppen statistisch signifikant unterschiedlich verteilt. Beim Haplotyp 5^{C1EX15}_9^{C4/N1} läßt sich ein p -Wert von 0,068 berechnen, damit ist das Niveau von $\alpha=0,05$ zwar nicht unterschritten, es wird jedoch ein Trend für eine unterschiedliche Verteilung zwischen den Patientengruppen angezeigt.

Die übrigen Haplotypen weisen alle keine signifikanten p -Werte auf und lassen sich damit zu einer zweiten Gruppe zusammenfassen.

Der Tabelle 5 ist die Anzahl der nicht konstruierbaren Phasen zu entnehmen. Eine Phase ist dann nicht eindeutig bestimmbar, wenn aufgrund der Allelkonstellation bei den Patienten und dessen Eltern eine manuelle Rekonstruktion der Haplotypen unmöglich ist. Von den

insgesamt 490 zu bestimmenden Phasen konnten 28 nicht eindeutig abgeleitet werden, dies entspricht einem Anteil von 5,7%. Die Auswertung in Abb. 19 beruht damit auf einer relativ vollständigen Datenerhebung.

Gesamtes Studienkollektiv: beobachtete Haplotypen und berechnete p -Werte bei den Patienten mit alternativer (I) und ohne residuelle Chloridionenleitfähigkeit (IIa).

Marker	Abstand [kbp]	Patienten- Gruppe	Haplotyp				nicht konstruierbare Phasen [n]	p -Wert
			1-1	1-2	2-1	2-2		
1 ^{C2IN3} _2 ^{C2IN11}	17,6	IIa	6	4	1	29	2	0,95
		I	5	2	0	15	2	
2 ^{C2IN11} _3 ^{C1IN11}	20,0	IIa	6	0	6	26	3	0,66
		I	5	1	4	14	1	
3 ^{C1IN1} _4 ^{C1IN10}	23,4	IIa	6	8	11	17	1	0,75
		I	5	3	4	10	2	
4 ^{C1IN10} _5 ^{C1EX15}	7,8	IIa	16	0	12	12	2	0,34
		I	8	0	3	9	3	
5 ^{C1EX15} _9 ^{C4IN1}	49,8	IIa	7	19	2	8	4	0,068
		I	9	4	2	7	2	
5 ^{C1EX15} _6 ^{C1/4}	23,3	IIa	16	11	7	4	3	0,40
		I	4	8	6	2	3	
6 ^{C1/4} _7 ^{C1/4}	16,7	IIa	25	0	14	3	1	0,29
		I	12	0	7	5	1	
7 ^{C1/4} _8 ^{C4IN1}	7,7	IIa	9	30	3	0	1	0,15
		I	7	12	5	0	1	
8 ^{C4IN1} _9 ^{C4IN1}	2,0	IIa	10	0	0	28	3	0,11
		I	11	0	0	11	2	
9 ^{C4IN1} _12 ^{C4IN6}	16,7	IIa	9	2	13	16	2	0,013
		I	7	4	0	11	2	
9 ^{C4IN1} _10 ^{C4IN1}	1,9	IIa	3	9	0	30	1	0,16
		I	5	7	0	12	1	
10 ^{C4IN1} _11 ^{C4IN1}	6,2	IIa	3	1	18	22	0	0,39
		I	3	1	12	6	2	
11 ^{C4IN1} _12 ^{C4IN6}	14,8	IIa	1	18	21	0	2	0,17
		I	0	14	6	0	3	
12 ^{C4IN6} _13 ^{C3IN4}	71,9	IIa	16	6	8	10	2	0,19
		I	3	3	8	6	3	

Tabelle 5: Neben der Größe, welche sich aus dem Abstand der Marker ergibt, sind für jede der beiden Patientengruppen I und IIa die beobachteten Haplotypen sowie der daraus berechnete p -Wert notiert. Überdies ist jeweils die Anzahl der Patienten aufgelistet, bei denen aufgrund der Allelkonstellation eine manuelle Rekonstruktion der Haplotypen nicht durchführbar war (nicht konstruierbare Phasen).

3.2.2.2 Altersstratifizierte Gruppen: Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den Patienten mit alternativer Chloridionenleitfähigkeit und ohne meßbare Chloridionenleitfähigkeit

Altersstratifizierte Gruppen: Vergleich der Verteilung von Zwei-Marker-Haplotypen zwischen den CF-Patienten mit alternativer (I) und ohne nachweisbare Chloridionenleitfähigkeit (IIb)

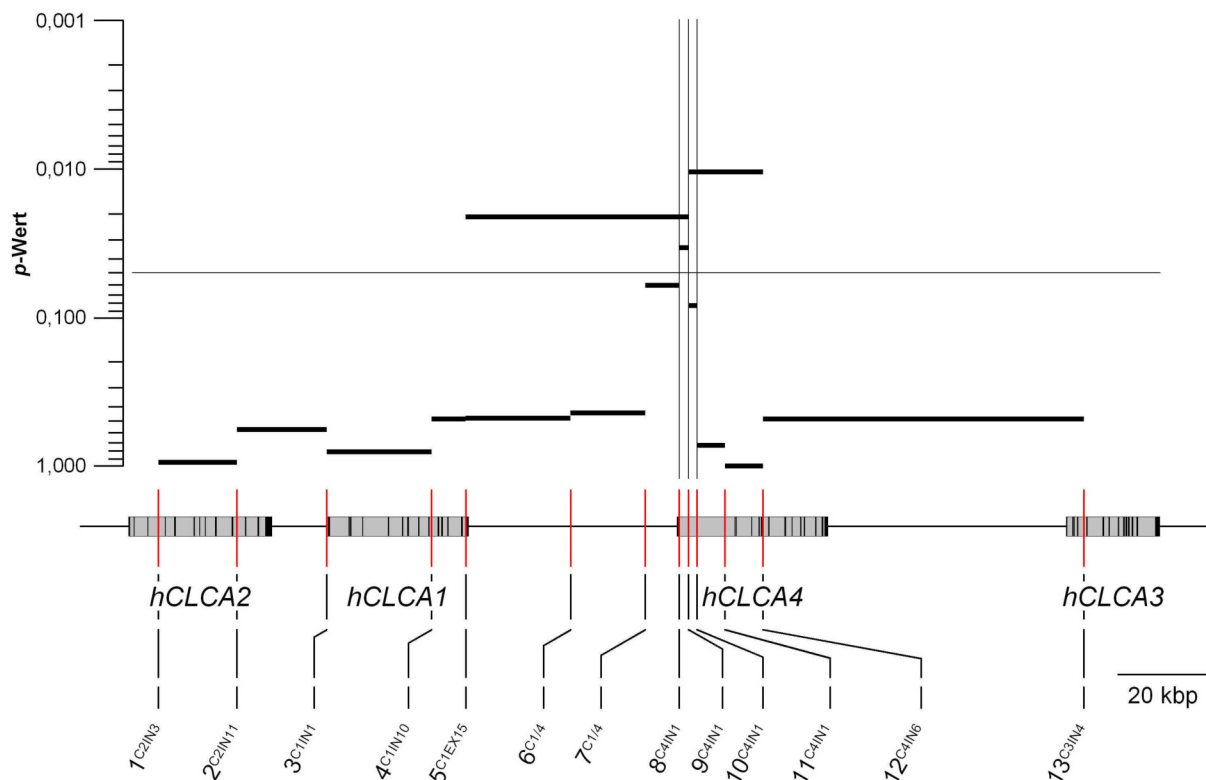


Abb. 20: Es sind die berechneten p -Werte an den Positionen der einzelnen Zwei-Marker-Haplotypen im *CLCA*-Gen-Cluster dargestellt. Die drei vertikalen Linien markieren zur besseren Übersichtlichkeit die Positionen der Marker $8^{C4/IN1}$, $9^{C4/IN1}$ und $10^{C4/IN1}$, die horizontale Linie gibt das Signifikanz-Niveau für $\alpha=0,05$ an.

Die Haplotypen lassen sich anhand des Signifikanz-Niveaus und der Größe in drei Gruppen einteilen.

So ist der Haplotyp 5^{C1EX15} _ $9^{C4/IN1}$ mit einem p -Wert von 0,02 zwischen beiden Studiengruppen signifikant unterschiedlich verteilt. Dieser überspannt ausgehend vom letzten Exon des Genes *bCLCA1* und endend im Intron 1 des Homologes *bCLCA4* den gesamten nicht-kodierenden intergenischen Bereich. Ein weiterer Haplotyp $9^{C4/IN1}$ _ $12^{C4/IN6}$ zeigt mit $p=0,01$ ebenfalls eine statistisch signifikant unterschiedliche Verteilung, dieser erstreckt sich vom Anfang des Introns 1 bis zum Intron 6 des Genes *bCLCA4* und deckt dieses somit etwa zur Hälfte ab. Diese beiden bilden die erste Gruppe.

Eine weitere Gruppe mit weniger signifikanten p -Werten besteht aus den Haplotypen $7^{C1/4}_8^{C4IN1}$ ($p=0,06$), $8^{C4IN1}_9^{C4IN1}$ ($p=0,03$) und $9^{C4IN1}_{10}^{C4IN1}$ ($p=0,08$). Diese erstrecken sich auf eine Distanz von 11,6 kbp und überspannen die Promoter-Region, das erste Exon und einen Teil des ersten Introns des Genes *bCLCA4*.

Die übrigen Haplotypen weisen alle nur sehr große p -Werte auf (z. B. $6^{C1/4}_7^{C1/4}$: $p=0,5$) und sind damit zwischen den miteinander verglichenen Patientengruppen nicht signifikant unterschiedlich verteilt. Diese bilden gemeinsam die dritte Gruppe.

Tabelle 6 ist durch Summierung der Spalte „nicht konstruierbare Phasen“ zu entnehmen, daß in 18 Fällen eine Haplotyprekonstruktion unmöglich war. Dies entspricht von der gesamten Zahl der zu bestimmenden Phasen (364) einem Anteil von 4,9%. Die Auswertung beruht damit auf einem relativ vollständigen Datensatz.

Altersstratifizierte Gruppen: Beobachtete Haplotypen und berechnete p-Werte bei den Patienten mit alternativer (I) und ohne residuelle Chloridionenleitfähigkeit (IIb).

Marker	Abstand [kbp]	Patienten- Gruppe	Haplotyp				nicht konstruierbare Phasen [n]	p-Wert
			1-1	1-2	2-1	2-2		
1 ^{C2IN3} _2 ^{C2IN11}	17,6	IIb	4	2	1	19	0	0,93
		I	5	2	0	15	2	
2 ^{C2IN11} _3 ^{C1IN11}	20,0	IIb	4	0	2	18	1	0,56
		I	5	1	4	14	1	
3 ^{C1IN1} _4 ^{C1IN10}	23,4	IIb	4	3	8	11	0	0,80
		I	5	3	4	10	2	
4 ^{C1IN10} _5 ^{C1EX15}	7,8	IIb	11	0	6	7	1	0,49
		I	8	0	3	9	3	
5 ^{C1EX15} _9 ^{C4IN1}	49,8	IIb	3	<u>14</u>	1	6	1	0,021
		I	<u>9</u>	4	2	<u>7</u>	2	
5 ^{C1EX15} _6 ^{C1/4}	23,3	IIb	9	7	4	2	2	0,47
		I	4	8	6	2	3	
6 ^{C1/4} _7 ^{C1/4}	16,7	IIb	15	0	9	2	0	0,43
		I	12	0	7	5	1	
7 ^{C1/4} _8 ^{C4IN1}	7,7	IIb	3	21	2	0	0	0,06
		I	7	12	5	0	1	
8 ^{C4IN1} _9 ^{C4IN1}	2,0	IIb	4	0	0	20	1	0,034
		I	11	0	0	11	2	
9 ^{C4IN1} _12 ^{C4IN6}	16,7	IIb	2	2	<u>7</u>	13	1	0,010
		I	<u>7</u>	4	0	<u>11</u>	2	
9 ^{C4IN1} _10 ^{C4IN1}	1,9	IIb	2	3	0	21	0	0,083
		I	5	7	0	12	1	
10 ^{C4IN1} _11 ^{C4IN1}	6,2	IIb	2	0	15	9	0	0,72
		I	3	1	12	6	2	
11 ^{C4IN1} _12 ^{C4IN6}	14,8	IIb	1	14	7	0	2	1,0
		I	0	14	6	0	3	
12 ^{C4IN6} _13 ^{C3IN4}	71,9	IIb	7	2	6	9	1	0,48
		I	3	3	8	6	3	

Tabelle 6: Neben der Größe, welche sich aus dem Abstand der Marker ergibt, sind für jede der beiden Patientengruppen I und IIb die beobachteten Haplotypen sowie der daraus berechnete p-Wert notiert. Überdies ist jeweils die Anzahl der Patienten aufgelistet, bei denen aufgrund der Allelkonstellation eine manuelle Rekonstruktion der Haplotypen nicht durchführbar war (nicht konstruierbare Phasen).

3.2.3 Stammesgeschichtlich fixierte Rekombinationsbruchpunkte im *CLCA*-Gen-Cluster

Im diesem Kapitel erfolgt eine Darstellung der im *CLCA*-Gen-Cluster ermittelten stammesgeschichtlich fixierten Rekombinationsbruchpunkte. Die Rekonstruktion erfolgte durch Anwendung des Vier-Gameten-Testes. Dabei wird eine weit in der Vergangenheit zurückliegende Rekombination angenommen, wenn für zwei diallelische Marker mehr als drei verschiedene Haplotypen beobachtet werden (siehe Kapitel 2.8 im Methodenteil).

Die Darstellung in Abb. 21 beruht auf der Analyse aller typisierten Proben, sodaß neben den Patienten mit alternativer Chloridionenleitfähigkeit und ohne meßbare Chloridionenleitfähigkeit auch deren Geschwister berücksichtigt wurden, welche einen gemischten Phänotyp in der intestinalen Ionenstromflußmessung zeigten oder gesund waren.

Stammesgeschichtlich fixierte Rekombinationsbruchpunkte im CLCA-Gen-Cluster, alle typisierten Nachkommen

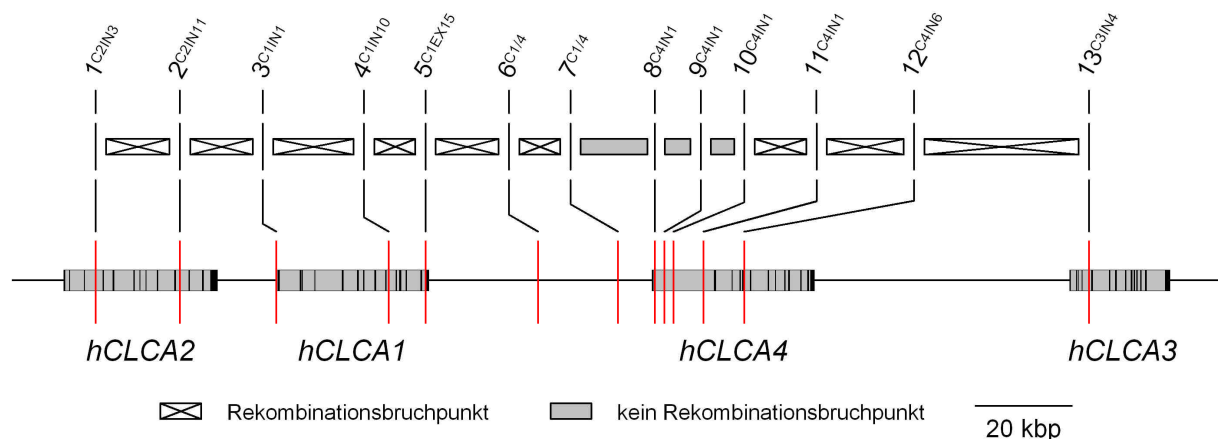


Abb. 21: In der oberen Zeile sind die mit Hilfe des Vier-Gameten-Testes abgeleiteten stammesgeschichtlich fixierten Rekombinationsbruchpunkte dargestellt. Ist ein ausgefülltes Symbol gezeigt, liegt kein Rekombinationsbruchpunkt vor, ist ein Kreuz dargestellt, so weist der Test auf eine Rekombination zwischen den korrespondierenden Markern hin.

Die Längenskalierung der Symbole ist zur Verbesserung der Übersichtlichkeit abweichend von der Anordnung der Marker vorgenommen worden.

Diese Darstellung beruht auf der Analyse von 65 typisierten Patienten-Proben.

Wie Abb. 21 zu entnehmen ist, können für die Mehrzahl der benachbarten Marker Rekombinationsbruchpunkte konstruiert werden. Eine Ausnahme bilden die Abschnitte zwischen den Markern 7^{C1/4} und 8^{C4IN1}, 8^{C4IN1} und 9^{C4IN1} sowie 9^{C4IN1} bis 10^{C4IN1}. Hier können durch Betrachtung der Haplotypen keine stammesgeschichtlich fixierten Rekombinationsbruchpunkte abgeleitet werden.

Diese Marker sind über einen Bereich verteilt, welcher die Promoter-Region, das erste Exon sowie einen Teil des ersten Introns vom Homolog *bCLCA4* überspannt.

3.2.4 Auswahl des Zielbereiches und der DNA-Proben für die Resequenzierung

In der Assoziationsstudie wurde die Verteilung der Allele von dimorphen Markern im *CLCA*-Gen-Cluster zwischen CF-Patienten mit alternativer Chloridionenleitfähigkeit und ohne meßbare Chloridionenleitfähigkeit verglichen. Dabei wurde eine Korrelation zwischen dem Phänotyp und den Haplotypen bestehend aus den Markern $7^{C1/4}$ _ 8^{C4IN1} , 8^{C4IN1} _ 9^{C4IN1} sowie 9^{C4IN1} _ 10^{C4IN1} beobachtet. Diese überspannen in der 5'-Region des Homologes *bCLCA4* einen Bereich mit der Länge von 11,6 kbp (siehe Abb. 22, Markierung β).

Um Hinweise für mögliche weitere DNA-Variationen in diesem durch die Marker $7^{C1/4}$ bis 10^{C4IN1} eingegrenzten Abschnitt zu erhalten, kann eine Resequenzierung erfolgen. Die zu sequenzierenden DNA-Proben sollten anhand derjenigen Haplotypen ausgewählt werden, welche zwischen den beiden miteinander verglichenen Patientengruppen den deutlichsten Verteilungsunterschied aufweisen. Um bei der Resequenzierung eindeutige Signale zu erhalten, sollten zudem die Proben nur von solchen Patienten eingesetzt werden, die für die zur Auswahl verwendeten Haplotypen homozygot sind.

Für die Auswahl der Proben wurden die Ergebnisse der Assoziationsstudie herangezogen, welche anhand der nach dem Alter stratifizierten Gruppen ermittelt worden sind (siehe Kapitel 3.2.2.2). Die Auswertung zeigt, daß für die Haplotypen 5^{C1EX15} _ 9^{C4IN1} und 9^{C4IN1} _ 12^{C4IN6} *p*-Werte mit der höchsten Signifikanz berechnet werden können. Diese überspannen den gesamten intergenischen Bereich zwischen *bCLCA1* und *bCLCA4* und etwa die Hälfte des Genes *bCLCA4* (siehe Abb. 22, Markierung α).

Die folgenden Ausführungen beziehen sich auf die Tabelle 6 auf der Seite 66:

Betrachtet man die Marker 5^{C1EX15} und 9^{C4IN1} , so ist der Haplotyp 1-2 in der Gruppe der Patienten ohne Chloridionenleitfähigkeit (IIB) überrepräsentiert. Dagegen sind die Haplotypen 1-1 und 2-2 bei den Patienten mit nachgewiesener alternativer Chloridionenleitfähigkeit (I) häufiger anzutreffen.

Bei dem Haplotyp bestehend aus den SNPs 9^{C4IN1} und 12^{C4IN6} ist die Variante 2-1 bei den Patienten ohne Chloridionenleitfähigkeit im Vergleich zu den Patienten mit alternativer Chloridionenleitfähigkeit (I) überrepräsentiert. Dagegen ist der Haplotyp 1-1 bei den Patienten mit residueller Chloridionenleitfähigkeit (IIB) vergleichsweise häufiger anzutreffen. Die Variante 2-2 ist in beiden Patienten-Gruppen (I und IIB) etwa gleich häufig typisiert worden.

Die zu sequenzierenden DNA-Proben wurden anhand dieser beiden Haplotypen kombiniert zu Drei-Marker-Haplotypen ausgewählt. In Tabelle 7 sind diese Patienten-Proben und deren Haplotypen bestehend aus den Markern 5^{C1EX15} 9^{C4IN1} 12^{C4IN6} aufgelistet.

Auflistung der zur Sequenzierung ausgewählte Patienten-Proben

	5^{C1EX15}	9^{C4IN1}	12^{C4IN6}
343A	1 1	2 2	1 1
67B *	1 1	1 1	1 2
23A *	1 2	1 2	2 2
465B *	1 2	1 2	2 2

Tabelle 7: Es sind die zur Sequenzierung verwendeten Proben und deren Drei-Marker-Haplotypen bestehende aus den Markern 5^{C1EX15} , 9^{C4IN1} und 12^{C4IN6} aufgelistet. Jedes der beiden Chromosomen eines Patienten ist durch eine Zeile repräsentiert (oben väterlich, unten mütterlich). Die Patienten mit alternativer Chloridionenleitfähigkeit sind mit einem Sternchen (*) markiert und sind gelb hinterlegt. Die Probe 343A stammt dagegen von einem Patienten ohne residuelle Chloridionenleitfähigkeit und ist blau hinterlegt.

Für die Gruppe der Patienten ohne Chloridionenleitfähigkeit wurde der Patient 343A ausgewählt, dieser ist homozygot für den Haplotyp 1-2-1.

In der Gruppe der Patienten mit alternativer Chloridionenleitfähigkeit war kein homozygoter Träger verfügbar. Demnach wurden die Patienten 23A und 465B jeweils mit den gleichen Haplotypen 1-1-2 und 2-2-2 und der Patient 67B, mit den Haplotypen 1-1-1 und 1-1-2 ausgewählt.

Wie bereits oben beschrieben, wurde in der Assoziationsstudie eine Korrelation zwischen dem gastrointestinalen Phänotypen der Patienten und den Zwei-Marker-Haplotypen $7^{C1/4}$ 8^{C4IN1} , 8^{C4IN1} 9^{C4IN1} sowie 9^{C4IN1} 10^{C4IN1} nachgewiesen. Diese überspannen einen 11,6 kbp großen Bereich in der 5'-Region des Genes *bCLCA4*.

Da hier der Promoter und die 5'-UTR des Exon 1 gelegen ist und es sich bei diesen Elementen um wichtige regulatorische Bereiche eines Genes handelt, wurde das Ziel der Sequenzierung auf diese Region festgelegt. So wurde schließlich als initiales Target der

Resequenzierung ein 2965 bp großer Teilbereich dieser Region, 5'-gelegen von Marker 8^{C4IN1} für eine Resequenzierung ausgewählt (siehe Abb. 22 Markierung γ).

Festgelegter Zielbereich der Resequenzierung im CLCA-Gen-Cluster

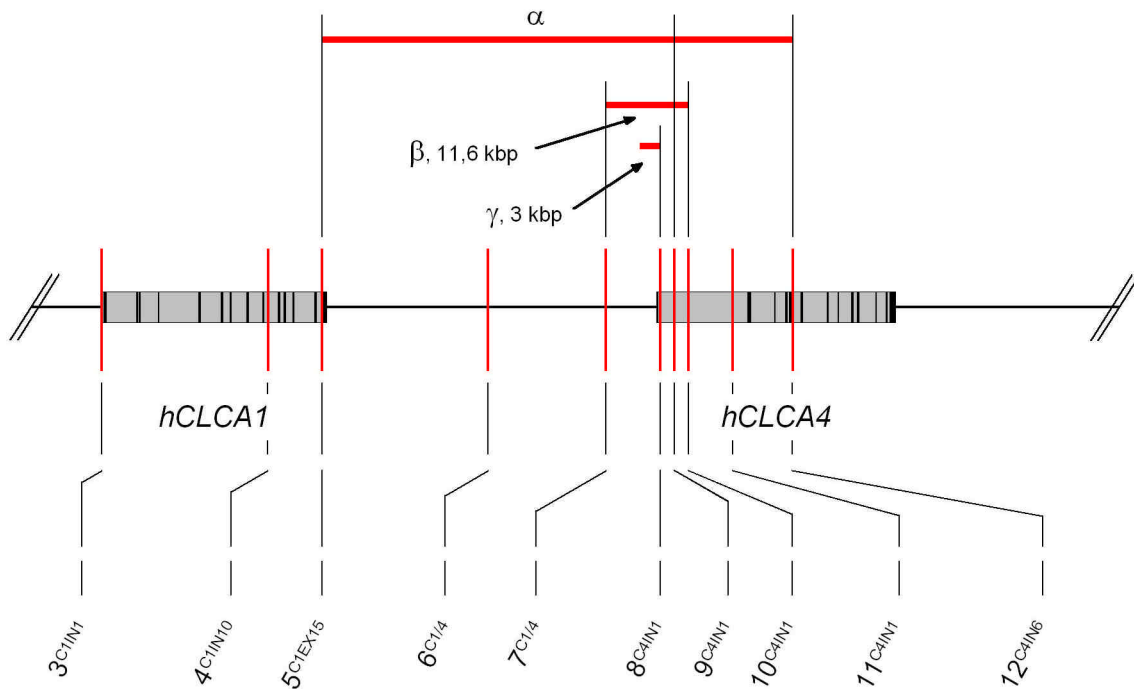


Abb. 22: Die Karte gibt einen Ausschnitt des *CLCA*-Gen-Clusters wieder. Mit α ist der Bereich markiert, welcher von dem Drei-Marker-Haplotypen 5^{C1EX15} _ 9^{C4IN1} _ 12^{C4IN6} überspannt wird, dieser wurde zur Auswahl der Proben herangezogen, welche sequenziert werden sollten. Das durch Auswertung der Assoziations-Studie ermittelte Fragment, welches mit dem gastrointestinalen Phänotyp korreliert, ist mit β gekennzeichnet. Und mit γ ist derjenige Teil bezeichnet, welcher als Ziel der Resequenzierung festgelegt wurde.

3.3 Ergebnisse der Teilresequenzierung des *CLCA*-Gen-Clusters

In diesem Kapitel werden die Ergebnisse der Teilresequenzierung des Promoters und des ersten Exons von dem Homolog *bCLCA4* dargestellt.

Da die Sequenzierung bei den entscheidenden DNA-Variationen nicht vollkommen verlässlich war, erfolgt in Kapitel 3.3.1 zunächst eine Darstellung der Qualität der Sequenzierung. Anschließend werden in Kapitel 3.3.2 die Ergebnisse der Rohdatenauswertung gezeigt und dann erfolgt in Kapitel 3.3.3 eine Darlegung der Resultate von der Nachtypisierung der detektierten Sequenzvariationen.

3.3.1 Qualität der Sequenzierung

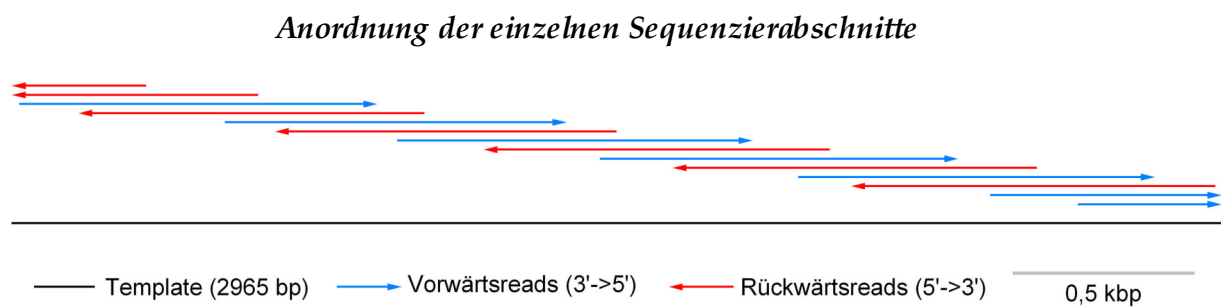


Abb. 23: Im unteren Teil ist das Template dargestellt und im oberen Teil symbolisieren die Pfeile die Länge und Richtung der einzelnen Reads. Die Darstellung basiert auf den Daten von Probe 343A, sie ist repräsentativ für alle Sequenzierungen.

In Abb. 23 sind die in der Sequenzierung erstellten Reads gezeigt, es ist die Länge und Anordnung in Relation des amplifizierten Templates zu entnehmen. Die Zeichnung basiert auf den Daten der DNA-Probe 343A und ist repräsentativ für die übrigen Sequenzierungen.

Das PCR-Produkt ist über seine ganze Länge gleichmäßig und vollständig mit Reads abgedeckt. Es liegt eine mindestens dreifache Überlappung vor und jeder Abschnitt wird durch gegenläufige Reads erfaßt.

Sequenzierqualität – Annotation von Qualitätswerten mit dem Phred-Algorithmus

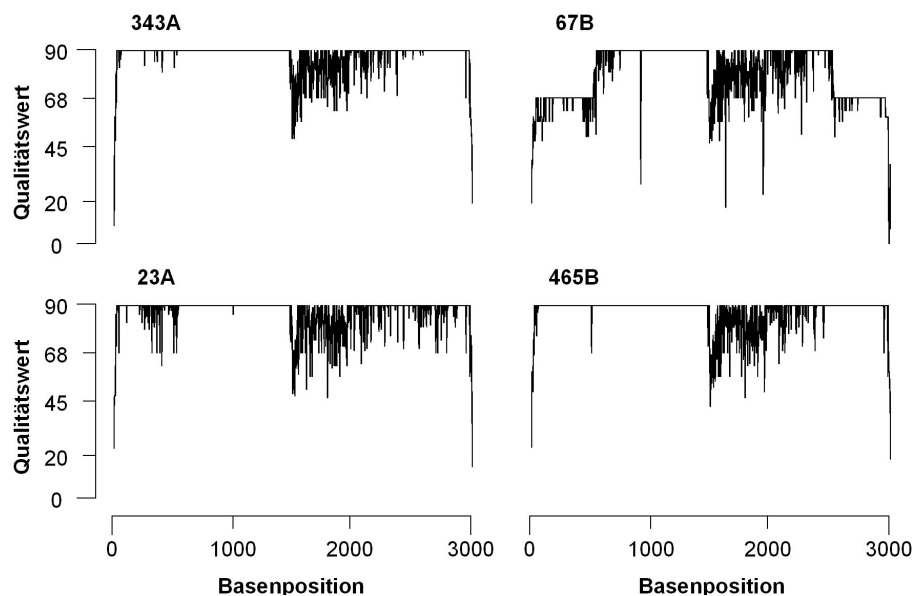


Abb. 24: Es ist für jede Probe die Auslese-Qualität der einzelnen Basen über den gesamten sequenzierten Bereich dargestellt, große Werte entsprechen einer guten Qualität und kleine einer geringen.

Zur Erstellung der Graphen werden die Qualitätswerte von den zwei besten gegenläufigen Reads einer jeden Base addiert. Liegt Information von nur einem Read pro Richtung vor, wird der Gesamtwert auf 68 limitiert. Ist für eine Richtung mehr als ein Sequenzierabschnitt verfügbar, wird der Grenzwert auf 90 gesetzt.

In Abb. 24 ist die Qualität der Basenauslesung der vier sequenzierten Proben graphisch dargestellt.

Für jede Probe wird das Limit von 90 über große Bereiche erreicht. Bei Probe 67B allerdings, ist das Limit für die ersten und letzten 500 bp auf 68 gesetzt.

Alle Proben zeigen einen Abfall der Qualität unter das Limit von 90 etwa über den Bereich von Position 1500 bis 2500. Hier schwanken die Werte zwischen 45 und 90, dennoch überschreiten sie die Grenze von 20 für qualitativ hochwertige Sequenzierungen.

Nebenbefund: exemplarische Darstellung eines (A)₁₁-Repeat-Motives

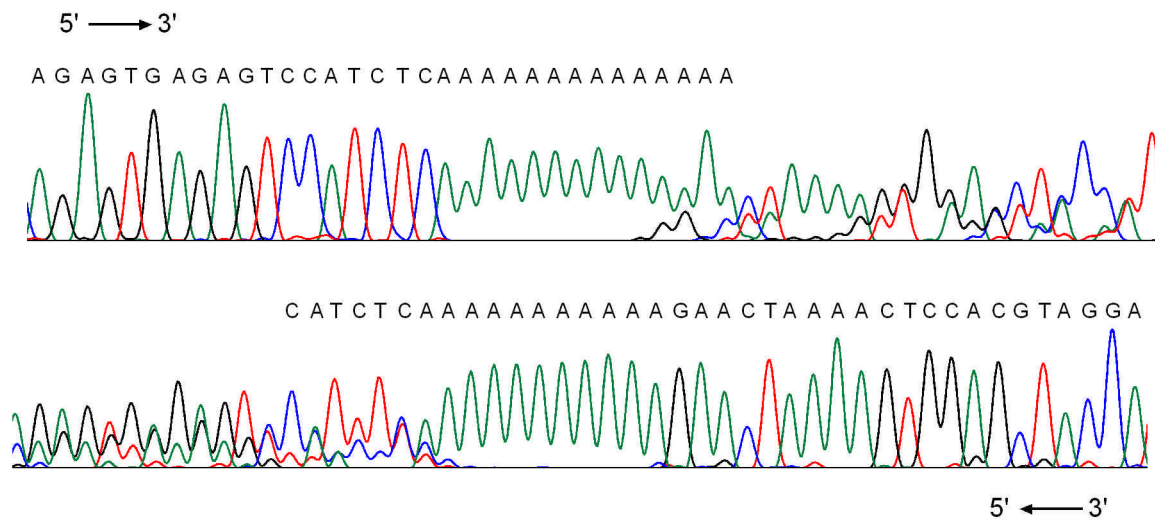


Abb. 25: Die Spurdaten zeigen ein (A)₁₁-Repeat-Motiv und geben den Bereich 79209-79309 der Probe 343A wieder. Die Pfeile zeigen die Richtung der Sequenzierreaktion an. Beachte die unterschiedliche Qualität der Elektropherogramme in der Umgebung des (A)₁₁-Repeats.

Anmerkung:

Auch die Signale der unteren Spur sind durch Sequenzierung in 5'→3' Richtung erstellt worden. Jedoch wurde dafür ein Primer verwendet, der an dem Gegenstrang des Templates bindet. Das Elektropherogramm wurde dann elektronisch in die Daten des reversen und komplementären Plus-Stranges umgewandelt. Auf diese Weise sind alle der in Abb. 23 als Rückwärts-Reads bezeichneten Sequenzierabschnitte konvertiert worden.

In Abb. 25 sind die Elektropherogramme von zwei gegenläufigen Reads wiedergegeben. In der Mitte ist eine Anordnung von 11 nacheinander folgenden Adenin-Nukleotiden zu erkennen, diese befinden sich an an Position 79209-79309 des Templates (siehe Abb. 26, Seite 75).

Betrachtet man die Spuren in der Arbeitsrichtung der Sequenzierpolymerase, so wird ersichtlich, daß die Signale nach dem (A)₁₁-Repeat-Abschnitt unsauber sind.

Ein solcher Abschnitt mit ebenfalls 11 aufeinander folgenden Adenin-Nukleotiden ist an der Position von 79769 bis 79779 vorhanden. Daneben sind 12 verschiedene Mononukleotid-Repeat-Motive mit einer Länge von vier bis sieben Basen über den gesamten sequenzierten Abschnitt verteilt. Diese bestehen zumeist aus Adenin- oder Thymin-Nukleotiden.

Ein Vergleich der Motive zwischen den einzelnen sequenzierten Proben ergab keinen Unterschied bezüglich der Länge.

3.3.2 Detektierte Variationen im Zielbereich der Sequenzierung: Ergebnisse der Rohdatenauswertung der Resequenzierung

Die nachfolgenden Ausführungen beziehen sich auf die Darstellungen der nächsten beiden Seiten.

Abb. 26 gibt die Basensequenz des ausgelesenen DNA-Abschnittes wieder. Sie wurde der NCBI-Datenbank entnommen und stellt somit die Referenz-Sequenz dar. Innerhalb dieser sind die zwischen den Patienten-Proben detektierten Sequenzvariationen hervorgehoben. So konnten insgesamt fünf Einzelnukleotidpolymorphismen gefunden werden. Diese liegen gleichmäßig verteilt mit einem Abstand von etwa 200 bis 300 bp voneinander entfernt in der Mitte des sequenzierten DNA-Auschnittes. Drei von diesen sind C zu T Polymorphismen und bei zweien handelt sich um einen G zu A bzw. A zu G Nukleotidaustausch.

Ein Vergleich mit der öffentlich zugänglichen SNP-Datenbank [67] ergab, daß lediglich drei der Marker bekannt waren (dbSNP-build 126).

In Tabelle 8 sind die detektierten SNPs im Kontext zu den Haplotypen der Patienten aufgelistet, deren DNA-Proben sequenziert worden sind. So ist ersichtlich, daß bei dem Patienten 343A homozygote Genotypen ermittelt wurden. Die Allele entsprechen zudem den Nukleotiden der Referenzsequenz. Von diesem unterscheiden sich die Patienten 23A und 465B durch homozygote Genotypen des Markers V^{C4P} . Dies steht in Kontrast zu den heterozygoten Genotypen der Marker 5^{C1EX15} und 9^{C4IN1} , welche für Auswahl der Sequenzierung verwendet worden sind.

Der Patienten 67B unterscheidet sich von den übrigen an der Position von vier SNPs, drei von diesen haben heterozygote Genotypen und der Marker II^{C4P} ist homozygot für das Allel A.

Sequenzierergebnisse

```

77281 tgtgtcccca acatacttaa ggtaataaaa gccatctatg acaaaccac agccaacatt
77341 atactgaaca gggaaaagt gaaaacatcc cccttgagaa ctagaacaag acaaggatgc
77401 ccactttcac cacttctatt caacataata ctggaagtct tagccagagc aacaagacaa
77461 gaaaaataaa gggcatccag attggtaaa aggaagtcaa actgtcactg tttgctgatg
77521 acatgatcat atatctagaa aaccctaaag actcatctaa aaagctccta gaactgggtca
77581 atgaattcat cagagttttg ggatacgaaa ttaacgtaca caaatcagta gctctcctat
77641 ataccaacag tgaccaagct gagaatcaaa tcaagaagtc aacccatttt acaatagtgc
77701 aaaaaatta aaatacttag gaatataact aaccocaggag gtgacagact tctgcaagga
77761 aaactacaaa acactgctga ataaatcata catgacacaa acaaatgaaa acacatccca
77821 tgctcatgga tggatagaat caatattatg aaaatgatca tactgccaac cacaatctac
77881 aaatccaatg caattcctat caaaaaccac catcattctt cacagaacta ggaaaaaat
77941 ctaaaaattc aaatggaacc caaagaagc ctgcatagcc aaagcaagc taggcaaaaa
78001 gaaaaaacct ggaggcatta cattacctga cttcagactc tactataagg ccatagttac
78061 caccacagca tggtagtagt gtaaaagcag gcatatagac gaatggaaca gaatatagaa
78121 cttagaataa aacaaaaata cttaaagtca actgatcatc acaaaagcaa acaaaaaaa
78181 aagtgaggaa ggacaccata ttcaacaat ggtgctggca taattggcaa taggcaaaaa
78241 gaagaatgaa actgggttct catcttttac cttatacaaa aatcaattca agatagatca
78301 aagacttaaa tctaaaacct gaaccataa aaattctaga agacaacttc agaagaacct
78361 tcctagacat tggetttagc aaaga■ttca tgaccaaaaa cccaaaagca aatgcaacaa
78421 atacaagat aaattggtg gacttagtta aactaaaac cttctgtgca gctaaaagaa
78481 taatcagcag agtaaacaga caaccacag aatgggagaa aatctctgca atctatacag
78541 ctgacaaagg actaatatcc agaactaca acaactcaa acaaatcagc aagaaaaaaa
78601 caaacagtcc catcaaaaag tgggctaagg aatgaatag acaattctca aaagaaggta
78661 tacaacacgc caacaacat atgaaaaaat gctcaacatc actaatgacg atcaaaaatgc
78721 aaatcaaac tacaatgc■ga tacctcctca ctcttgcaag aatggccata atcaaaaaat
78781 caaaaaataa tagatggtg catggatgtg gtgaaaaggg aacactttta cactgctgct
78841 gggaatgtaa actagtacaa cactatgga aacagtatg gagattcctt aaagaactaa
78901 aagtggtgct ggtgtggtg ttcacactg taatcccagc actttgggag gccaaggttg
78961 gtggatcacg agttcaggag ttcaagacca gctgaccaa catggtgaaa acctatctct
79021 actaaaaaac aaacaaaca acaacaat atacaaaaat tagccgggca tgggtggcaca
79081 ■gcctgtaat cccagctact gaggaggctg aggcagaaga attcctttaa ccctggaggc
79141 ggaggttgca gtgagccgat atcgcgcac tgcactccag cctggacaac agagtggagc
79201 tccatctcaa aaaaaaaag aactaaaagt ggagctacca tttgatccag caatcctagc
79261 aataccactc ctgggtatct atccagagga aaagaagtca tatatgaaga agatacttgc
79321 acacacgttt atagcagcac aatgtgcaat tgcaaaaata gaaactagcc caaatgctca
79381 tcaatcaac■ agtgataaa gaaaatgtga tatatatata tacacatata cctacataca
79441 tacacacaca ccaatggaata ttacttagcc ataaaaagaa atgaaataat ggcatttgca
79501 gcagccaatt tcagaaatgg aaaaccaaac atcatatgtt ctactcatg agtgggagct
79561 aagccatgag gatgcaaaag cataagaacg acaca■tgga ctttggggac tccgggaaaa
79621 ggtatggagg gggagggtg ggataaaaaga ctacacattg ggtacagtgt acactgctcg
79681 ggtggttagg tgacaaaaa tctcagaat tataattaa gaacttattc atgtaaccaa
79741 accccacctg ttacccaaa acctattgaa aaaaaaaat cctgttgtc aaggagccct
79801 tagaatatca acaatgacaa caataaagaa aacgtgttcc aacaggaaaa tgtgcaatgc
79861 atttataaa atgatctgat ttattctgac tcttagtcac ttttatgtaa catataatta
79921 agcaatcagg tgaattagtc atgtatttcc tcatgtaatg tacctttata catgcaacat
79981 ccataatca gccaaactg gtaagtttta ggacgtata taaagagaaa gaaaagaaa
80041 gcctcttgaa caaccaaca tttgagccag gaataactag agaggaacaa tggggttatt
80101 cagaggtttt gtttctctct tagttctgtg cctgctgcac cagtcaata cttcctcat
80161 taagtgaaat aataatggct ttgaagatat tgcattggtt atagatccta tgggtccaga
80221 agatgaaaaa ataattgaac aaatagaggt aagaacaaa tggaaatatc ttcattaatt
80281 tttagttttt tactattatt ttaattgac atataattgt acatatttat ggcatacaaa
80341 gtagtatttc tatacatgta tacaattgt aaagatcaa tcagggtaat tagcatattc
80401 agcaactcaa acatttatct ttgcattgtg tccgggaacat tataattatt ctctctagc
80461 tactttgaaa tacacaacag attatcagta actttaattt tcctattgta ctatcaata
80521 ctagaactta ttctgtctac ataactgtgc tttgtacac attgacaac ttctctcat
80581 ttctttctc tcccctatcc tttgcagcct ctaataagta tcattctact ctctatctcc

```

xxx angrenzende Sequenz ■ bekannte SNP
xxx Hybridisierungsposition der Primer ■ unbekannte SNP
xxx Exon1, *hCLCA4*

Position	Variation	Allele	rs-Nummer
78386	SNP I^{C4P}	C / T	Erstbeschreibung in dieser Arbeit
78739	SNP II^{C4P}	G / A	rs2791476
79081	SNP III^{C4P}	C / T	rs11807298
79390	SNP IV^{C4P}	C / T	rs11807638
79596	SNP V^{C4P}	A / G	Erstbeschreibung in dieser Arbeit

Abb. 26: Darstellung der Sequenzierergebnisse. Die Sequenzvariationen sind farbig markiert. (→)

Die Angaben der linken Spalte bezeichnen die Position innerhalb des Klonen AL122002.16, von diesem ist die Referenz-Sequenz entnommen worden (NCBI-Datenbank [87]). Die rs-Nummern bezeichnen die in der dbSNP [67] vorhandenen SNPs.

Durch direkte Sequenzierung erhaltene Genotypen

	5 ^{C1EX15}	I ^{C4P}	II ^{C4P}	III ^{C4P}	IV ^{C4P}	V ^{C4P}	9 ^{C4IN1}	12 ^{C4IN6}
343A	1	C	G	C	C	A	2	1
	1	C	G	C	C	A	2	1
67B *	1	C / T	A	C / T	C / A	A	1	1
	1	C / T	A	C / T	C / A	A	1	2
23A *	1	C	G	C	C	G	1	2
	2	C	G	C	C	G	2	2
465B *	1	C	G	C	C	G	1	2
	2	C	G	C	C	G	2	2

Tabelle 8: Es sind die sequenzierten Patienten-Proben und die entdeckten SNPs I^{C4P} bis V^{C4P} aufgelistet. Daneben sind die zur Auswahl der Proben verwendeten Drei-Marker-Haplotypen 5^{C1EX15}_9^{C4IN1}_12^{C4IN6} dargestellt. Jedes der beiden Chromosomen eines Patienten ist durch eine Zeile repräsentiert (oben väterlich, unten mütterlich). Die Proben von Probanden mit alternativer Chloridionenleitfähigkeit sind mit einem Sternchen (*) markiert und gelb hinterlegt. Die Probe 343A stammt von einem Patienten ohne residuelle Chloridionenleitfähigkeit und ist blau hinterlegt.

3.3.3 Detektierte Variationen im Zielbereich der Sequenzierung: Verifizierung der Ergebnisse durch Nachtypisierung

Im nachfolgenden wurden, wie bei der Typisierung der übrigen Marker, für die durch die Resequenzierung entdeckten SNPs Primer entworfen und Restriktionsenzyme mit passender Erkennungssequenz bestimmt.

Diese Rezepte wurden etabliert, um die Ergebnisse der Sequenzierung zu überprüfen, die Genotypen der Eltern für die Konstruktion der Haplotypen zu bestimmen sowie die Typisierung der Marker bei den Patienten der gesamten Studienpopulation durchzuführen.

In Abb. 27 auf Seite 79 sind die Resultate der Sequenzierung den Ergebnissen der Nachtypisierung gegenübergestellt.

So war es möglich, für die Marker II^{C4P} und III^{C4P} stabile und sehr gut lesbare Rezepte zu etablieren. Für den SNP III^{C4P} zeigt die Sequenzierung als auch der Restriktionsverdau einen heterozygoten Zustand an, sodaß die Allele C/T an diesem Locus für den Patienten 67B bestätigt werden konnten. Anders verhält es sich beim Marker II^{C4P}. Die Spurdaten der Sequenzierung zeigen ein sauberes und ohne jeglichen Hintergrund erscheinendes Signal für

die Base G. Dies spricht demnach für einen homozygoten Genotypen. Die Nachprüfung durch Restriktionsverdau zeigt dagegen einen heterozygoten Zustand mit den Allelen A/G bei dem Patienten 23A an.

Für die Marker I^{C4P} und IV^{C4P} konnten ebenfalls Rezepte etabliert werden. Allerdings erhielt man erst nach Aufreinigung der PCR-Produkte durch Verwendung des Qiagen-Aufreinigungs-Kit's eindeutige Resultate. Zur Interpretation des Restriktionsverdaues der Nachtypisierung von Marker I^{C4P} ist folgendes zu beachten:

In der Abbildung erscheint der Verdau leicht unsauber. So erkennt man beim Vater (F) neben einer deutlichen Bande für das Allel 1 eine schneller laufende mit geringerer Intensität. Diese hat eine Länge entsprechend dem Fragment für Allel 2. Bei der Mutter (M) verhält es sich umgekehrt, so weist hier die Bande für das Allel 1 eine schwächere und die des Allels 2 eine stärkere Intensität auf. In der Spur des Patienten (B) sind dagegen zwei Banden mit gleicher Intensität zu erkennen.

Bei der Probe des Vaters (F) handelt es sich um Zeichen von Star-Aktivität in Kombination mit einem unvollständigen Verdau bei der Probe der Mutter (M). Star-Aktivität bezeichnet eine Eigenschaft von manchen Restriktionsenzymen, DNA an Stellen zu schneiden, welche der eigentlichen Erkennungssequenz ähnlich, jedoch nicht identisch sind. So kann beispielsweise eine Sequenz, die sich in nur einem Nukleotid von der definierten Erkennungsstelle der Endonuklease unterscheidet, bei Abweichungen von den für das Enzym optimalen Reaktionsbedingungen gespalten werden.

Für das Enzym Ahd I, welches bei der Typisierung des SNPs I^{C4P} verwendet wurde, sind gemäß der Produktspezifikation (NEB-Produktkatalog 2005/2006) eine hohe Glycerin-Konzentration (>5%) sowie das organische Lösungsmittel DMSO als begünstigende Faktoren einer Star-Aktivität aufgeführt. Letzteres wurde als Additiv der PCR-Reaktion hinzugesetzt, welche zur Nachtypisierung des Markers I^{C4P} durchgeführt wurde.

Folge einer Abweichungen von den optimalen Reaktionsbedingungen kann zudem eine deutlich verringerte Aktivität des Restriktionsenzymes sein, sodaß der Verdau unvollständig abläuft (siehe Probe der Mutter, I^{C4P}).

Nach einem initialen Versuch gemäß der gewöhnlichen Verfahrensweise, resultierend in einem kaum interpretierbaren Verdau, wurde zunächst eine Aufreinigung mittels Ethanolpräzipitation versucht. Ein akzeptables Ergebnis, wie in der Abb. 27 gezeigt, konnte jedoch erst durch Verwendung des Qiagen-Aufreinigungs-Kit's erreicht werden.

Beurteilt man unter Vorkenntnis dieser Gegebenheiten zur Ableitung der Genotypen des Markers I^{C4P} lediglich die Intensitäten der Banden, so sind die Ergebnisse des Patienten 67B im Vergleich zu dessen Eltern eindeutig heterozygot, so trägt der Vater (F) den Genotyp 1-1, die Mutter 2-2 und der Patient (B) die Allel-Kombination 1-2.

Ähnlich wie für Marker I^{C4P} , konnte für SNP IV^{C4P} ohne Aufreinigung kein eindeutiges Ergebnis erzielt werden, sodaß auch hier das Qiagen-Kit eine Anwendung fand.

Schließlich konnten so beim Patienten 67B die Genotypen der Marker I^{C4P} und IV^{C4P} als heterozygot (1-2) bestimmt werden. Diese Ergebnisse sind konsistent mit den Spurdaten der Sequenzierung, sodaß die Existenz dieser Marker nachgewiesen ist.

Im Gegensatz dazu konnten die Ergebnisse der Resequenzierung für den Marker V^{C4P} nicht bestätigt werden. Zur Nachprüfung der Genotypen wurde hier initial ebenfalls ein Rezept mit Restriktionsverdau entwickelt.

Dieses zeigte jedoch einen unsauberen Verdau, dessen Ergebnis auch durch einen erneuten Versuch mit Aufreinigung der PCR-Produkte durch Ethanolpräzipitation nicht verbessert werden konnte. So erfolgte zwar eine Spaltung der PCR-Produkte, jedoch war das Ergebnis nicht mit den Rohdaten der Sequenzierung vereinbar war, nach welchen eine Spaltung der PCR-Produkte von beiden Chromosomen zu erwarten gewesen wäre.

Daher wurde ein weiteres Rezept mit einem anderen Restriktionsenzym etabliert. Auch hier war der Verdau sowohl ohne als auch mit Aufreinigung der PCR-Produkte durch Ethanolpräzipitation nur unsauber und daher nicht eindeutig interpretierbar. Dieser gab, wie bereits das erste Rezept, Hinweise für einen heterozygoten Genotypen des Patienten.

Aus diesen Grunde wurde als nächstes zur Überprüfung, ob das getestete PCR-Produkt der erwarteten Zielsequenz entspricht, eine Produktkontrolle durchgeführt.

Dazu erfolgte eine Inkubation des Amplifikates mit verschiedenen Restriktionsenzymen, deren Erkennungssequenz innerhalb des DNA-Stranges konserviert sind. Diese Kontrollreaktion wurde mit fünf verschiedenen Enzymen in getrennte Ansätzen durchgeführt. Dabei konnte das erwartete Produkt bestätigt werden.

Schließlich wurde zur Nachprüfung der Rohsequenzdaten bezüglich des SNPs V^{C4P} das Verfahren der allelspezifischen PCR angewendet.

Hierbei konnte mit sorgfältig optimierten Primerpaaren (siehe Methodenteil Kapitel 2.6.3.2) und mehrfacher Durchführung der Reaktion (3 mal) in voneinander unabhängigen Versuchsansätzen ein heterozygoter Genotyp des Patienten 23A für den SNP V^{C4P} ermittelt werden.

Wie der Abb. 27 zu entnehmen ist, differiert das Ergebnis der Sequenzierung somit von dem der Nachprüfung. Hier ist den Spurdaten der Sequenzierung wie bei Marker II^{C4P} ein sauberes Signal für die Base G zu entnehmen, die Verifizierung zeigt jedoch eindeutig ein heterozygotes Resultat beim Patienten 23A für den Marker V^{C4P} an.

Gegenüberstellung der Sequenzierergebnisse mit den Resultaten der Nachtypisierung für die Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P}

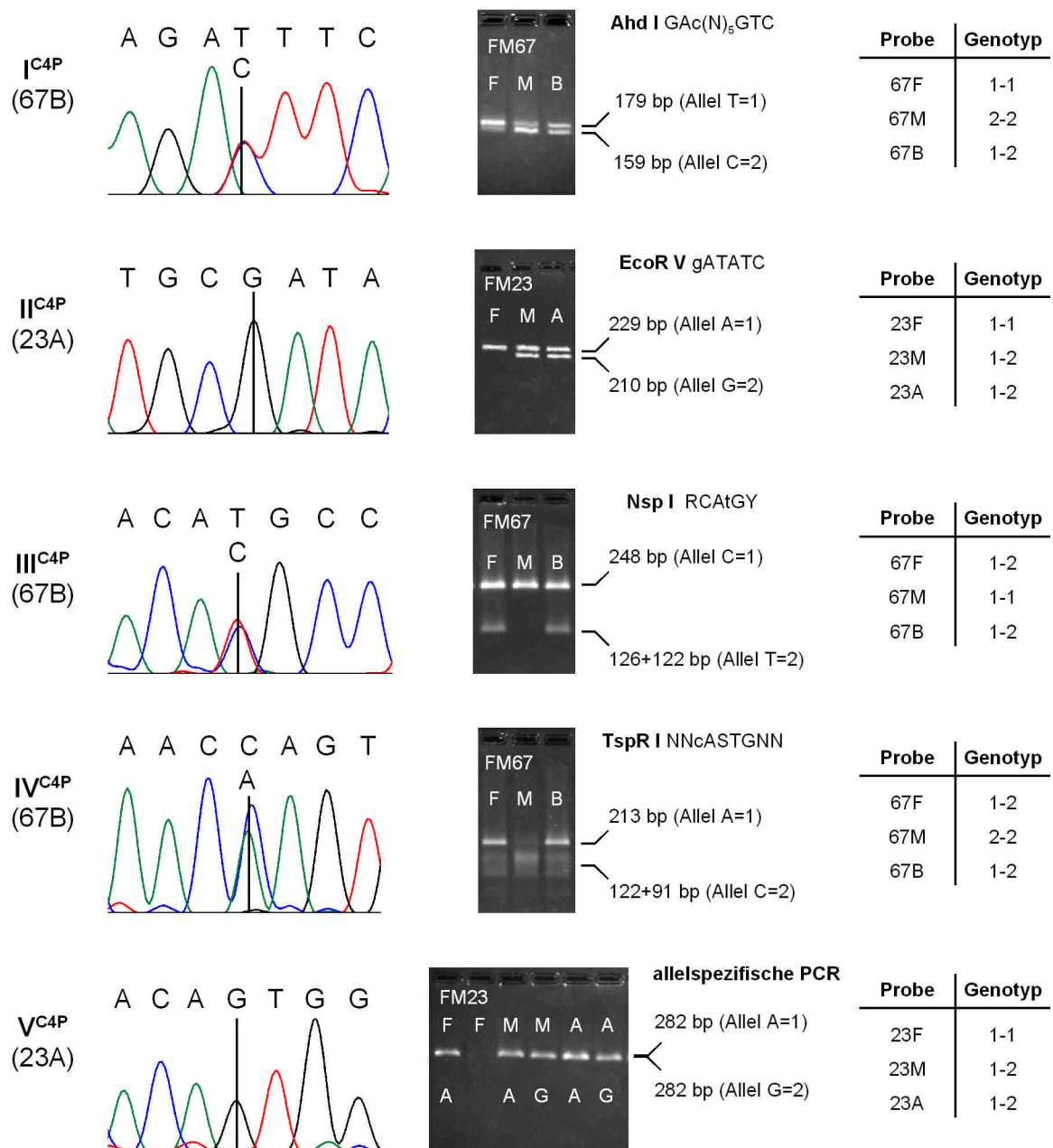


Abb. 27: In der linken Spalte sind die Spurdaten der Sequenzierung dargestellt und in der Mitte die Gelelektrophoresen der Nachtypisierung. In der rechten Spalte sind die Genotypen aufgelistet, wie sich entsprechend der Nachtypisierung erhoben worden sind.

Die Daten der SNPs II^{C4P} und V^{C4P} sind repräsentativ für die Patienten 23A und 465B. Die Resultate der Probe 67B für den Marker II^{C4P} sind nicht gesondert dargestellt.

F: Probe des Vaters, M: Probe der Mutter, A bzw. B: Probe des Patienten

Korrektur und Verifikation der Rohdaten der Sequenzierung

	5 ^{C1EX15}	I ^{C4P}	II ^{C4P}	III ^{C4P}	IV ^{C4P}	V ^{C4P}	9 ^{C4IN1}	12 ^{C4IN6}	Kernhaplotyp
343A	1	C	G	C	C	A	2	1	A
	1	C	G	C	C	A	2	1	A
67B *	1	T	A	T	A	A	1	1	B
	1	C	A	C	C	A	1	2	C
23A *	1	C	A	C	C	A	1	2	C
	2	C	G	C	C	G	2	2	D
465B *	1	C	A	C	C	A	1	2	C
	2	C	G	C	C	G	2	2	D

Tabelle 9: Es sind die sequenzierten Patienten-Proben und die Allele der entdeckten SNPs I^{C4P} bis V^{C4P} nach Verifizierung aufgelistet. Daneben sind die zur Auswahl der Proben verwendeten Drei-Marker-Haplotypen 5^{C1EX15}_9^{C4IN1}_12^{C4IN6} dargestellt.

Daneben sind in der rechten Spalte die Kernhaplotypen aufgelistet, welche sich aus den Markern I^{C4P} bis V^{C4P} konstruieren lassen. Diese sind noch einmal gesondert in der unten folgenden Tabelle 10 gezeigt.

Jedes der beiden Chromosomen eines Patienten ist durch eine Zeile repräsentiert (oben väterlich, unten mütterlich). Die Proben von Probanden mit alternativer Chloridionenleitfähigkeit sind mit einem Sternchen (*) markiert und gelb hinterlegt.

Tabelle 9 gibt eine Übersicht der Resultate von der oben beschriebenen Nachprüfung der Sequenziererergebnisse für die Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} wieder.

Durch die Etablierung der Rezepte für die Nachtypisierung konnte neben einer Verifizierung der Genotypen dieser Marker bei den Patienten auch eine Typisierung der Eltern vorgenommen werden. Dadurch war es möglich, die Phase der Allele von diesen SNPs bei den Patienten zu bestimmen und damit die Kern-Haplotypen bestehend aus den Markern I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} zu konstruieren.

Es lassen sich vier Varianten dieser Kernhaplotypen unterscheiden. Wie Tabelle 9 zu entnehmen ist, trägt der Patient 343A, bei dem keine residuelle Chloridionenleitfähigkeit nachweisbar war, auf beiden Chromosomen den Haplotypen A. Die Sequenz dieses Haplotypen (C_G_C_C_A) unterscheiden sich zudem nicht von der Referenzsequenz der NCBI-Datenbank (siehe Abb. 26). Dagegen weisen die Patienten mit alternativer Chloridionenleitfähigkeit Haplotypen auf, deren Basenreihenfolge nicht mit der Referenzsequenz übereinstimmt. So trägt der Patient 67B den Haplotypen B, welcher sich von dem Kernhaplotypen A durch die Allele der Marker I^{C4P}, II^{C4P}, III^{C4P} und IV^{C4P} abgrenzen läßt (T_A_T_A_A). Bei den Patienten 23A und 465B ist der Kernhaplotyp D jeweils auf einem Chromosom vorhanden, welcher sich von dem Haplotyp A durch das Allel des Markers V^{C4P} unterscheiden läßt (C_G_C_C_G). Alle Patienten mit DIDS-sensitiver

Chloridionenleitfähigkeit tragen den Kernhaplotypen C auf einem Chromosom. Dieser unterscheidet sich vom Haplotypen A durch das Allel von Marker II^{C4P} (C_**A**_C_C_A). In Tabelle 10 sind die beschriebenen Kernhaplotypen noch einmal vergleichend aufgelistet.

Übersicht der vier Kernhaplotypen, welche durch Sequenzierung und Nactypisierung entdeckt worden sind

	I ^{C4P}	II ^{C4P}	III ^{C4P}	IV ^{C4P}	V ^{C4P}
A	C	G	C	C	A
B	T	A	T	A	A
C	C	A	C	C	A
D	C	G	C	C	G

Tabelle 10: Es sind die vier Kernhaplotypen aufgelistet, welche aus der Sequenzierung hervorgegangen sind. Diese bestehen aus den Allelen der Marker I^{C4P} bis V^{C4P}. Diejenigen Allele, durch welche sich die Varianten B, C und D vom Haplotypen A unterscheiden, sind fett gedruckt.

3.4 Ergänzung der Assoziations-Studie: Typisierung der gesamten Studienpopulation an den durch Resequenzierung ermittelten Loci

Nachdem mittels einer Teilresequenzierung des *CLCA*-Gen-Clusters weitere fünf Marker nachgewiesen wurden, wodurch sich die Index-Patienten 343A, 67B, 23A und 465B voneinander unterscheiden, wurde als nächstes eine Typisierung dieser Marker bei der gesamten Studienpopulation vorgenommen. Zunächst werden in Kapitel 3.4.1 diese Marker charakterisiert, dann erfolgt in Kapitel 3.4.2 eine Ergänzung der Darstellung der stammesgeschichtlichen fixierten Rekombinationsbruchpunkte im *CLCA*-Gen-Cluster und schließlich wird in Kapitel 3.4.3 die Assoziationsstudie ergänzt. Bei dieser Studie wurden CF-Patienten mit nachgewiesener alternativer Chloridionenleitfähigkeit (Gruppe I) mit Patienten ohne residuelle Chloridionenleitfähigkeit (Gruppe IIa) in der Allelverteilung der Marker I^{C4P} bis V^{C4P} miteinander verglichen.

3.4.1 Charakterisierung der ermittelten Marker

Im folgenden werden die Eigenschaften der detektierten Marker beschrieben. Eine Übersicht der physikalischen Position ist Abb. 28 zu entnehmen. Die Werte der Allelhäufigkeiten sowie der Informationsgehalt der Marker (PIC-Wert) sind in Tabelle 11 zusammen gefaßt.

*Lokalisation der durch Sequenzierung detektierte Marker im Promoterbereich des Homologes *hCLCA4**

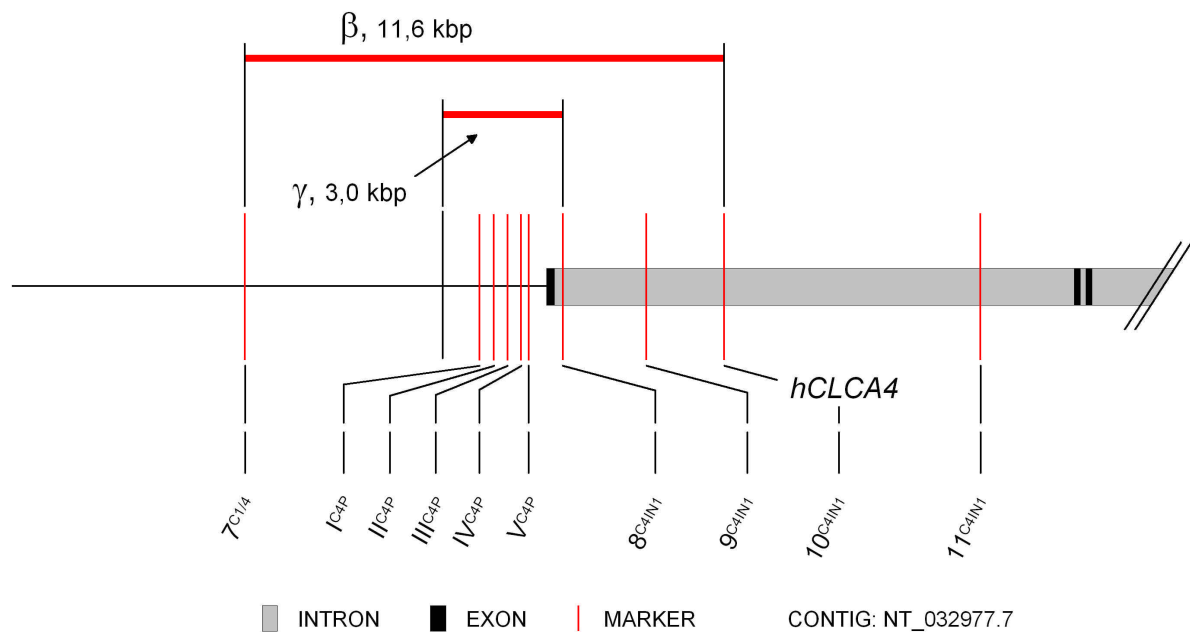


Abb. 28: Es ist die Lokalisation der durch Sequenzierung ermittelten Marker im Promoter-Bereich des Gene *hCLCA4* dargestellt. Der sequenzierte Bereich ist mit γ markiert und mit β ist das 11,6 kbp große Fragment bezeichnet, welches zwischen den CF-Patienten mit alternativer und ohne nachweisbare Chloridionenleitfähigkeit unterschiedlich verteilt ist.

Allelfrequenzen und Informativität der durch Resequenzierung ermittelten Marker

Marker	Quelle	Position	Allele (1/2)	f(a _i)	Heterozygotität (erwartet)	Heterozygotität (beobachtet)	PIC
I ^{C4P}	Erstbeschreibung	40830492	T / C	n.t.	n.t.	n.t.	n.t.
II ^{C4P}	rs2791476	40830845	A / G	A: 0,37	0,47	0,50	0,36
III ^{C4P}	rs11807298	40831187	C / T	T: 0,15	0,25	0,22	0,22
IV ^{C4P}	rs11807638	40831496	A / C	n.t.	n.t.	n.t.	n.t.
V ^{C4P}	Erstbeschreibung	40831702	A / G	G: 0,09	0,17	0,19	0,16

Tabelle 11:

- Quelle: Zugriffsnummer des Markers in der SNP-Datenbank [67]
- Position: Die Positionsangaben beziehen sich auf den Contig: NT_032977.7
- Allele (1/2): Die Basen der jeweiligen Allele sind gemäß der dbSNP notiert.
- f(a_i): Frequenz des jeweils selteneren Alleles
- Heterozygotität – erwartet: Anteil der heterozygoten Individuen bei Gültigkeit des Hardy-Weinberg-Gleichgewichtes, berechnet nach Formel (3), siehe Kapitel 2.3.1
- Heterozygotität – beobachtet: Diese Werte beruhen auf einer Auszählung der Elterngeneration (n=54)
- PIC: Polymorphismus-Informationsgehalt, (berechnet gemäß Formel (4), siehe Kapitel 2.3.1)

Lediglich die Marker II^{C4P}, III^{C4P} und IV^{C4P} waren in der öffentlich zugänglichen SNP-Datenbank [67] (dbSNP-build 126) registriert, für diese waren jedoch keine Informationen bezüglich der Allelhäufigkeiten verfügbar.

Die Allelfrequenzen und die daraus ermittelten Größen der Marker beruhen auf der Typisierung der Elterngeneration der Studienpopulation (n=54).

Die größte Frequenz für das jeweils seltenere Allel wurde mit 0,37 für den SNP II^{C4P} bestimmt. Demnach handelt es sich bei diesem um einen hochinformativen Marker, so liegt der Polymorphismus-Informationsgehalt mit 0,36 nahe dem Maximum von 0,37 für dimorphe Marker.

Dagegen sind die Frequenzen des jeweils selteneren Alleles bei den Markern III^{C4P} und V^{C4P} wesentlich kleiner, diese betragen 0,15 bzw. 0,09. Dementsprechend ist die Informativität dieser Marker geringer, so beträgt der Polymorphismus-Informations-Gehalt für den Marker III^{C4P} 0,22 und für den SNP V^{C4P} 0,16.

Bei allen drei genannten Markern weicht die gemessene Heterozygotität von der erwarteten um keine größeren Beträge ab, sodaß etwaige systematische Typisierungsfehler weitgehend ausgeschlossen werden können.

3.4.2 Stammesgeschichtlich fixierte Rekombinationsbruchpunkte im *CLCA*-Gen-Cluster – Ergänzung

*Untersuchung des Promotor-Bereiches des Gen *hCLCA4* auf stammesgeschichtlich fixierte Rekombinationsbruchpunkte – Ergänzung um die Marker II^{C4P} , III^{C4P} und V^{C4P}*

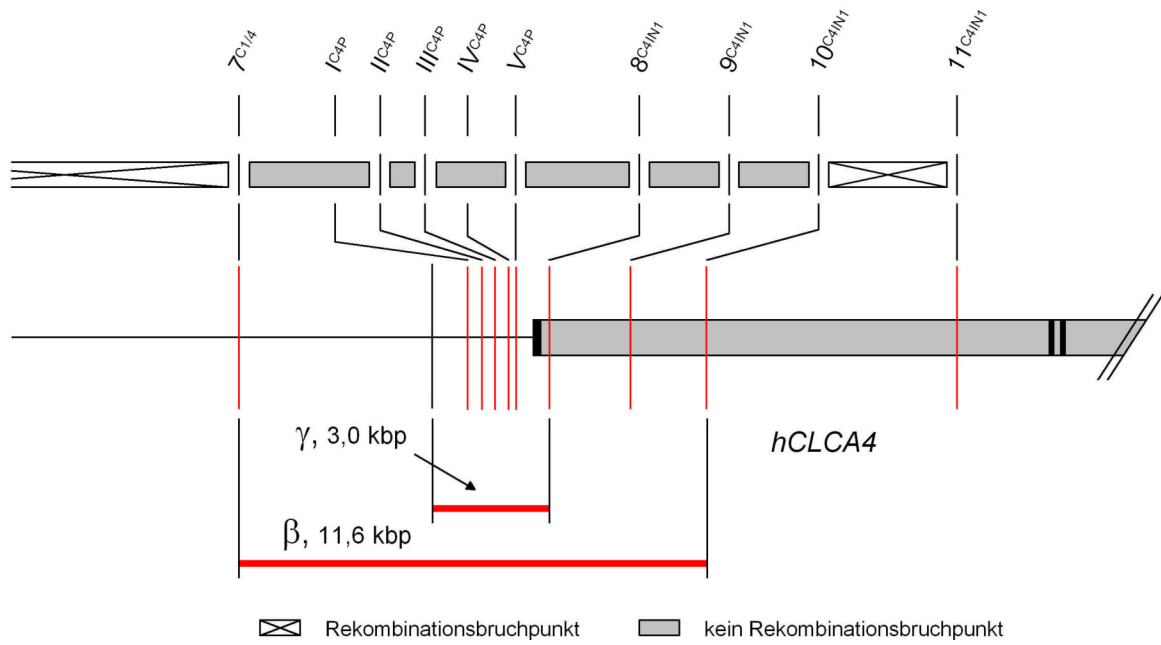


Abb. 29: Im oberen Teil sind die mit dem Vier-Gameten-Testes abgeleiteten stammesgeschichtlich fixierten Rekombinationsbruchpunkte dargestellt. Ist ein ausgefülltes Symbol gezeigt, liegt kein Rekombinationsbruchpunkt vor, ist ein Kreuz dargestellt, so weist der Test auf eine Rekombination zwischen den korrespondierenden Markern hin.

Diese Darstellung beruht auf der Analyse von allen typisierten Patienten-Proben ($n=65$).

Mit β ist im unteren Teil das durch Auswertung der Assoziationsstudie ermittelte Fragment gekennzeichnet, welches mit dem gastrointestinalen Phänotyp korreliert. Mit γ ist der sequenzierte Teil markiert.

Durch Auszählung der verschiedenen Haplotypen, bestehend aus den Markern $7^{C1/4}$ _ II^{C4P} , II^{C4P} _ III^{C4P} , III^{C4P} _ V^{C4P} sowie V^{C4P} _ 8^{C4IN1} konnten keine Hinweise für stammesgeschichtlich fixierte Rekombinationsbruchpunkte gewonnen werden.

Durch Betrachtung der Allelkombinationen zwischen den Markern $7^{C1/4}$, II^{C4P} , III^{C4P} , V^{C4P} , 8^{C4IN1} , 9^{C4IN1} , und 10^{C4IN1} konnte festgestellt werden, daß im Gegensatz zu den $2^7=128$ theoretisch möglichen Kombinationen nur fünf verschiedene bei den 65 typisierten Patienten beobachtet werden konnten.

Beide Befunde sprechen dafür, daß es sich bei der untersuchten 5'-Region des Genes *hCLCA4* um einen hochkonservierten Bereich handelt.

3.4.3 Ergänzung der Assoziations-Studie – Verteilung der detektierten DNA-Variationen innerhalb der gesamten Studienpopulation

In diesem Abschnitt erfolgt eine Auswertung der Typisierung von Marker II^{C4P}, III^{C4P} und V^{C4P} anhand der gesamten Studienpopulation. So werden die SNPs in Abb. 30 zunächst als einzelne Loci betrachtet, dann in Abb. 31 als Kombination der vier Kernhaplotypen A, B, C und D, und am Ende des Kapitels erfolgt schließlich eine Darstellung der beobachteten Diplotypen in der untersuchten Studienpopulation.

Vergleich der Allelfrequenzen von den Markern II^{C4P}, III^{C4P} und V^{C4P} zwischen den Patienten mit alternativer (Gruppe I) und ohne Chloridionenleitfähigkeit (Gruppe IIa)

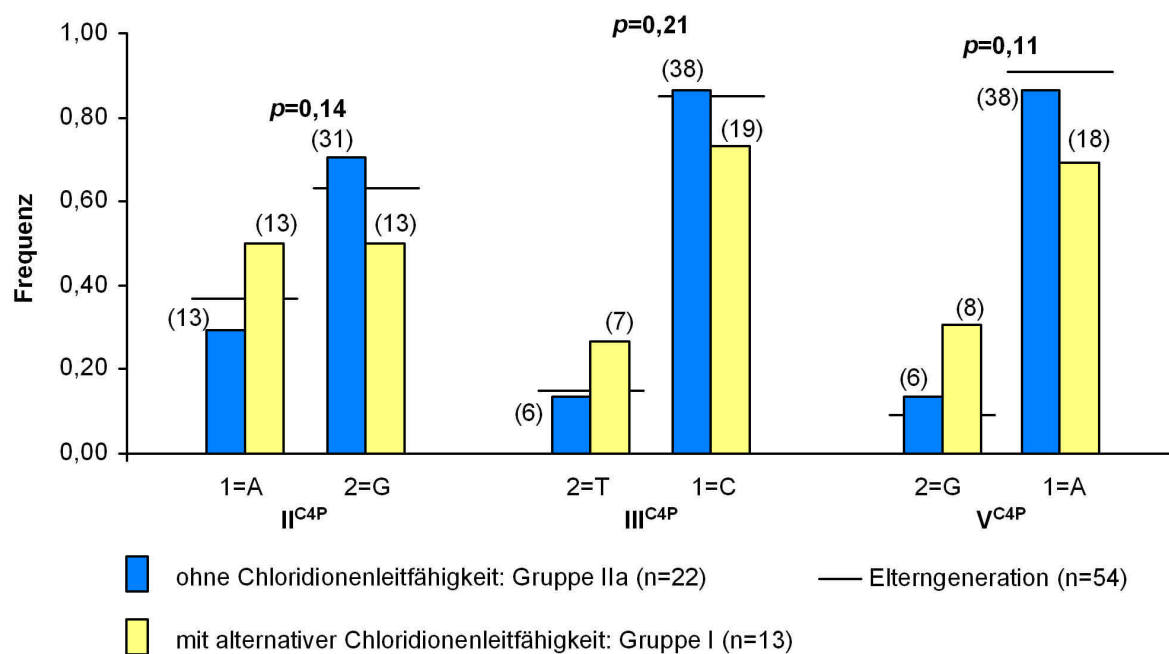


Abb. 30: Es sind die Allelfrequenzen der Markern II^{C4P}, III^{C4P} und V^{C4P} von den Patienten getrennt nach den beiden Phänotypen und den Eltern dargestellt. Überdies geben die Werte in den Klammern die absolute Anzahl der beobachteten Allele bei den Patienten an, die p -Werte wurden auf dieser Grundlage berechnet (siehe Text).

In Abb. 30 sind die Allelfrequenzen der Marker II^{C4P}, III^{C4P} und V^{C4P} für die Patienten mit alternativer Chloridionenleitfähigkeit (Gruppe I) und ohne nachweisbare Chloridionenleitfähigkeit (Gruppe IIa) dargestellt. Überdies sind in der Zeichnung auch die Allelfrequenzen der Eltern gezeigt.

Durch einen direkten Vergleich der absoluten Anzahl der Allele zwischen den Patientengruppen lassen sich keine signifikanten p -Werte berechnen, sie zeigen jedoch einen Trend an:

So fällt auf, daß bei den Patienten mit alternativer Chloridionenleitfähigkeit das jeweils seltenere Allel im Vergleich zu den Patienten ohne Chloridionenleitfähigkeit überrepräsentiert ist. Dies wird insbesondere bei den Marker II^{C4P} und V^{C4P} deutlich.

Verteilung der Kernhaplotypen A, B, C und D zwischen den Patienten mit alternativer Chloridionenleitfähigkeit (Gruppe I) und ohne nachweisbare Chloridionenleitfähigkeit (Gruppe IIa)

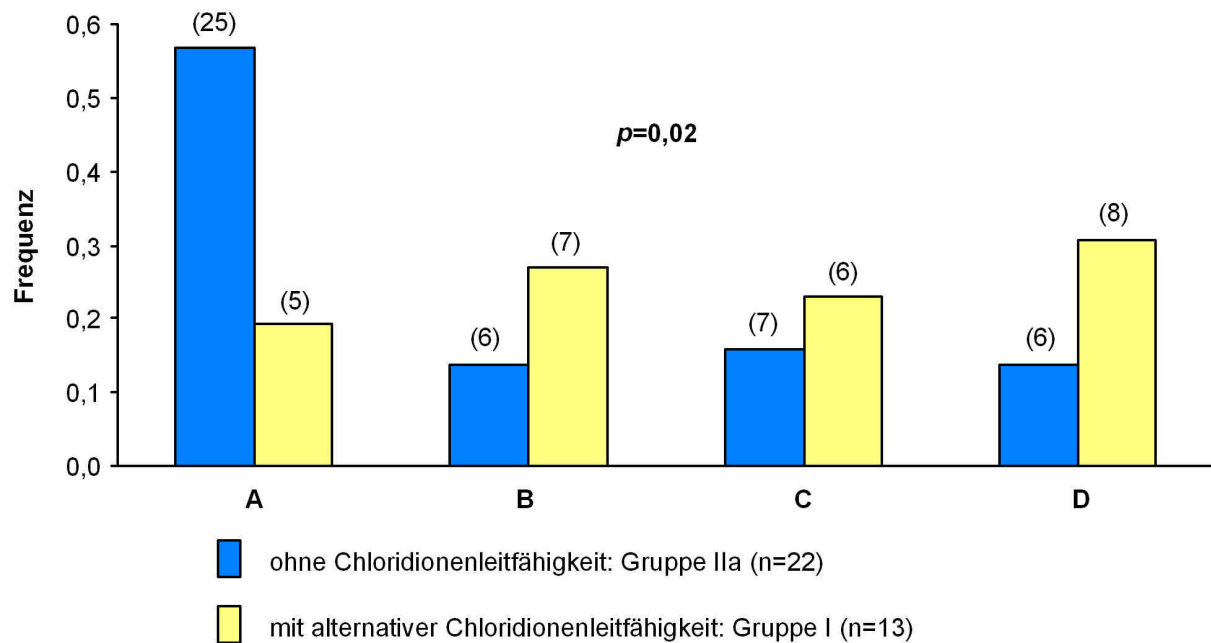


Abb. 31: Es sind die Haplotypfrequenzen der Patienten getrennt nach den beiden gastrointestinalen Phänotypen dargestellt. Die Werte in den Klammern geben die absolute Anzahl der Haplotypen in der jeweiligen Patientengruppe an, das Signifikanz-Niveau wurde anhand dieser Werte berechnet (siehe Text).

Diese Darstellung beruht auf der Typisierung von Marker II^{C4P} , III^{C4P} und V^{C4P} mittels Restriktionsverdau oder allelspezifischer PCR in der gesamten Studienpopulation. Für die Konstruktion der Kernhaplotypen A, B, C und D wurden die durch Typisierung erhaltenden Drei-Marker-Haplotypen unter der Annahme, daß es sich um konservierte Haplotypen handelt, zu den entsprechenden Fünf-Marker-Haplotypen ergänzt.

In Abb. 31 sind die Häufigkeiten der Kernhaplotypen A, B, C und D (Nomenklatur siehe Tabelle 10, Seite 81) bei den Patienten mit alternativer Chloridionenleitfähigkeit (Gruppe I) und ohne nachweisbare Chloridionenleitfähigkeit (Gruppe IIa) dargestellt.

Der p -Wert von 0,02 gibt eine signifikant unterschiedliche Verteilung der Kernhaplotypen zwischen den Gruppen der Patienten an. So ist die Variante A bei den CF-Erkrankten ohne nachweisbare Chloridionenleitfähigkeit im Vergleich zu den übrigen Haplotypen überrepräsentiert (n=25). Bei den Patienten mit alternativer Chloridionenleitfähigkeit sind die

Varianten B und D am häufigsten zu beobachten ($n=7$ bzw. 8), der Haplotyp A ist hier dagegen unterrepräsentiert ($n=5$).

Beobachtete Diplotypen bestehend aus den Haplotypen A, B, C, und D bei den Patienten mit alternativer Chloridionenleitfähigkeit (Gruppe I) und ohne nachweisbare Chloridionenleitfähigkeit (Gruppe IIa)

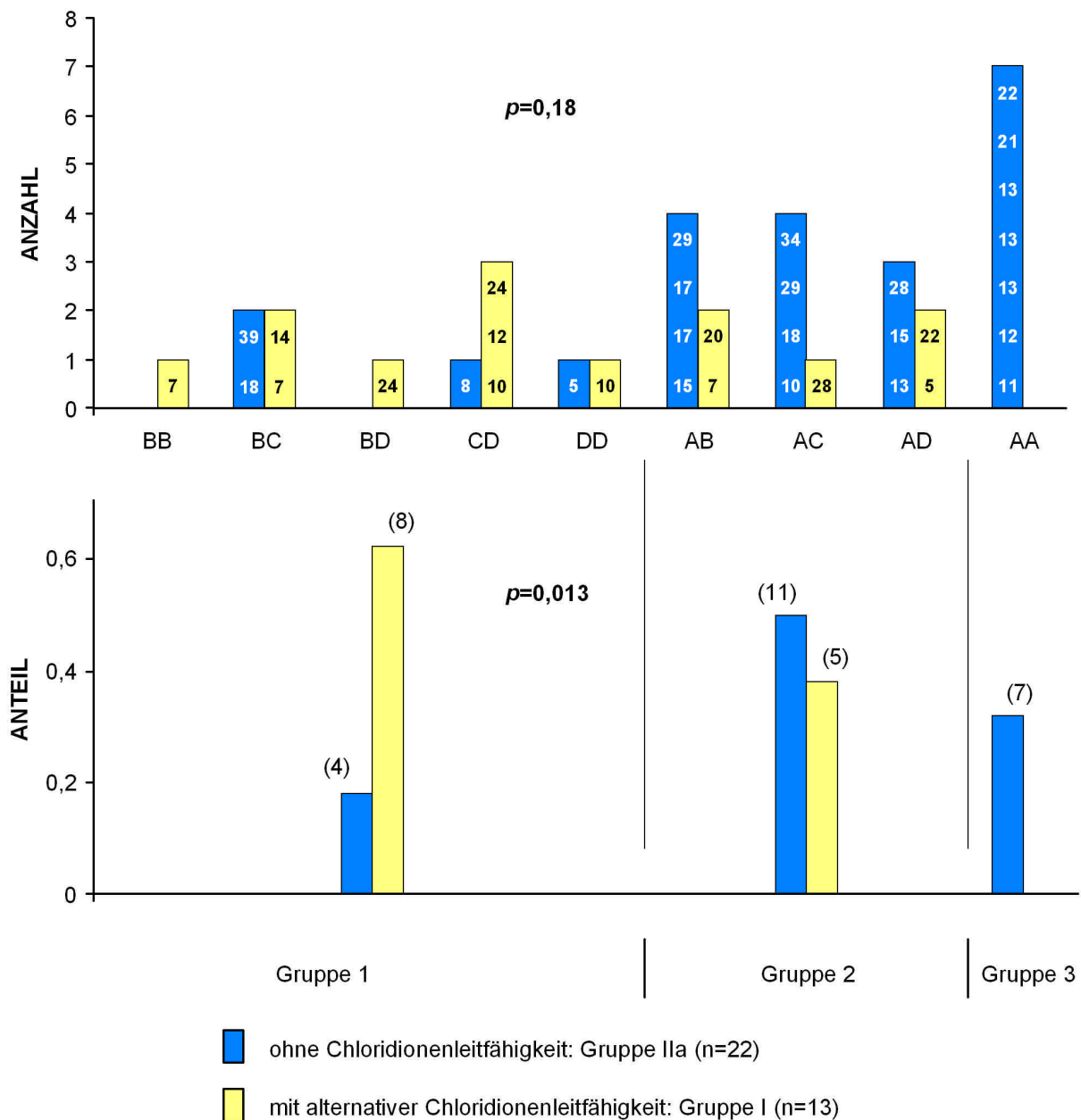


Abb 32: Im oberen Teil ist die absolute Anzahl der beobachteten Diplotypen getrennt nach den beiden gastrointestinalen Phänotypen der Patienten dargestellt. Zusätzlich geben die Ziffern das Alter der einzelnen Probanden an. Im unteren Teil sind die Patienten mit den Diplotypen BB bis DD, AB bis AD und AA zu den Kategorien 1, 2 und 3 zusammengefaßt. Hier ist der relative Anteil der Patienten in diesen drei Gruppen getrennt nach den beiden Phänotypen gezeigt. Die Werte in den Klammern geben die absolute Anzahl der Probanden an. Diese sind Grundlage des gezeigten p -Wertes. Diese Darstellung beruht auf der Typisierung von Marker II^{C4P} , III^{C4P} und VC^{4P} mittels Restriktionsverdau oder allelspezifischer PCR in der gesamten Studienpopulation. Für die

Konstruktion der Kernhaplotypen A, B, C und D wurden die durch Typisierung erhaltenen Drei-Marker-Haplotypen unter der Annahme, daß es sich um konservierte Haplotypen handelt, zu den entsprechenden Fünf-Marker-Haplotypen ergänzt.

In Abb. 32 sind die beobachteten Kombinationen der Kernhaplotypen A, B, C und D dargestellt. Sie sind getrennt nach der Gruppe der Patienten mit alternativer Chloridionenleitfähigkeit (I) und ohne residuelle Chloridionenleitfähigkeit (IIa) gezeigt.

Anhand dieser Kombinationen lassen sich die Patienten in drei Kategorien einteilen:

So gibt es eine Anzahl von Patienten, welche den Haplotyp A auf keinem Chromosom tragen (Gruppe 1), dann eine weitere, welche den Haplotyp A auf nur einem Chromosom tragen (Gruppe 2) und schließlich eine Kategorie von Patienten, die für den Haplotyp A homozygot sind (Gruppe 3).

Vergleicht man die Verteilung der Probanden von der Gruppe IIa und I auf diese drei Kategorien, so fällt auf, daß 62% aller Patienten mit alternativer Chloridionenleitfähigkeit in der Gruppe 1 und 38% in der Kategorie 2 vertreten sind, jedoch keiner dieser Probanden in der Gruppe 3 anzutreffen ist.

Dagegen sind nur 18% der Probanden ohne nachweisbare Chloridionenleitfähigkeit in der Gruppe 1 vertreten und 50% dieser Patienten in der Gruppe 2 sowie 32% in der Kategorie 3 anzutreffen. Diese Verteilung ist mit einem p -Wert von 0,013 als statistisch signifikant zu bewerten.

Ergänzend ist im oberen Teil der Abb. 32 das Alter von den Patienten vermerkt. Es fällt auf, daß mehrere Probanden ohne nachweisbare Chloridionenleitfähigkeit anzutreffen sind, die das zwanzigste Lebensjahr wesentlich überschritten haben (älter als 25 Jahre). Dies ist insbesondere bei den CF-Erkrankten mit den Diplotypen der Kategorie 2 der Fall: So betrifft dies je einen Patienten mit dem Diplotyp AB und AD sowie zwei Probanden mit dem Diplotypen AC. Daneben ist noch ein weiterer Proband mit dem Diplotyp BC in der Gruppe 1 vorhanden, welcher wesentlich älter als 20 Jahre ist.

Unter den Patienten mit alternativer Chloridionenleitfähigkeit ist nur ein Proband vorhanden, welcher das zwanzigste Lebensjahr bereits weit überschritten hat; dieser trägt den Diplotypen AC.

3.5 Ergebnisse der proteinchemischen Untersuchungen

Um das Ergebnis der Assoziationsstudie auf proteinchemischer Ebene zu verifizieren, wurden Antikörper gegen das Homolog hCLCA4 benötigt.

Da für diesen Zweck keine geeigneten kommerziellen Antikörper zur Verfügung standen, wurden in Kooperation mit der Firma Eurogentec (Seragin, Belgien) polyklonale Peptidantikörper synthetisiert (siehe Kapitel 2.15 im Methodenteil).

Erfolgskontrolle der Immunisierung: Nachweis von Antikörpern gegen das Peptid-Epitop I des Proteins hCLCA4

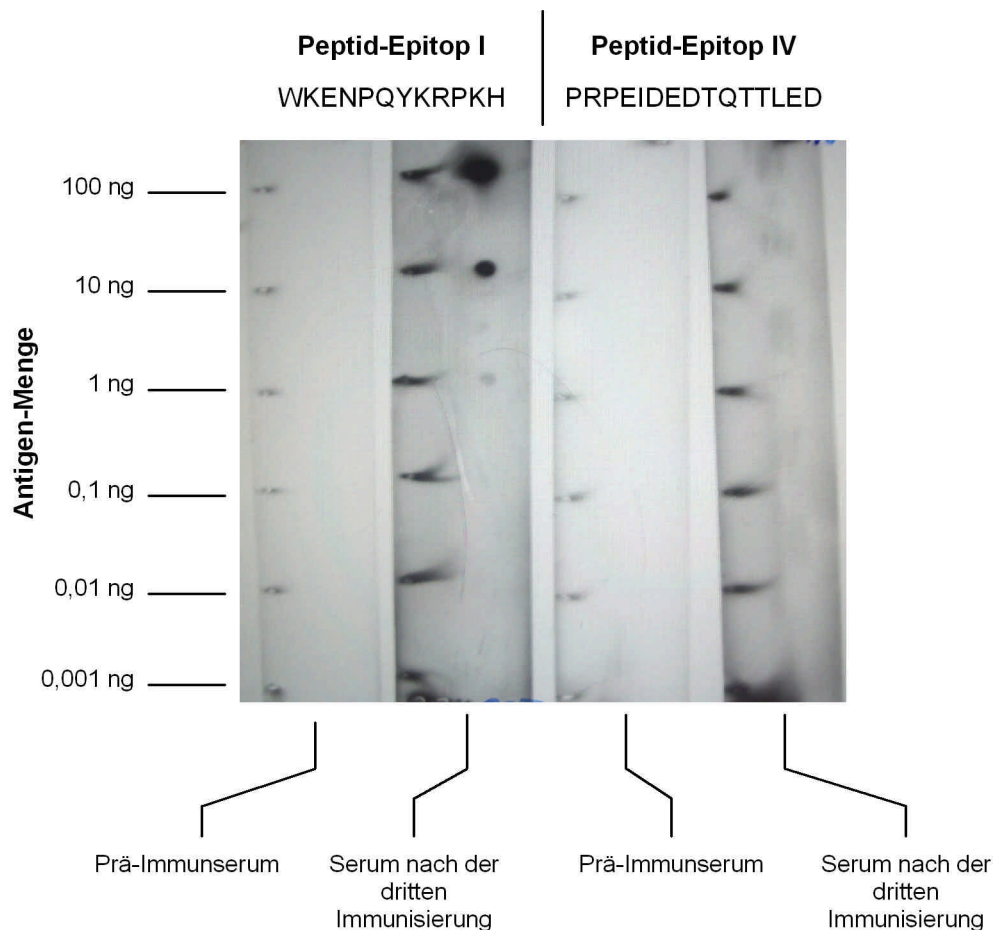


Abb. 33: Immunodotblot: Getrennt nach Peptid-Epitop I und IV erfolgte die Inkubation der Dotblotstreifen mit Präimmun- und Immunsereum des Kaninchens SA5113.

Das Peptid-Epitop I kann durch das Serum bis zu einer Masse von 1 ng nachgewiesen werden, während das Epitop IV nicht sicher durch das Immunsereum erkannt wird. Im Präimmunserum können keine Antikörper in relevanter Quantität oder Affinität gegen die zur Immunisierung verwendeten Epitope nachgewiesen werden.

Versuchsbedingungen: Verdünnung der Seren: 1:750, Verdünnung des sekundären Antikörpers: 1:5000, Belichtungszeit: 180 s

Wie in Abb. 33 dargestellt, ist es gelungen, gegen das Peptidepitop I des Proteins hCLCA4 Antikörper durch Immunisierung von Kaninchen zu gewinnen. Da die Produktion dieser Antiseren jedoch länger als vorgesehen dauerte, konnten keine weiteren Experimente durchgeführt werden.

Neben der erfolgreichen Immunisierung der Kaninchen wurden auch Hühner zur Antikörperproduktion verwendet, sodaß die Voraussetzung für weitere Experimente geschaffen worden sind, um Daten über die quantitative Expression und Lokalisation des Proteins hCLCA4 auf zellulärer Ebene zu gewinnen.

3.6 Analyse der detektierten Marker – welche Einflüsse üben sie aus?

Wie in den vorangehenden Kapiteln beschrieben, ergab die Auswertung der Assoziationsstudie, daß es sich bei dem Homolog *bCLCA4* um das Gen handelt, welches den Modulator des gastrointestinalen Phänotyps kodiert. Weiterhin konnten durch eine Teilresequenzierung des Promoters dieses Genes die vier Kernhaplotypen A, B, C und D ermittelt werden. Sie sind zwischen den miteinander verglichenen Gruppen der Mukoviszidose-Patienten unterschiedlich verteilt.

Nimmt man an, daß die alternative Chloridionenleitfähigkeit mit einer zur Vergleichsgruppe abweichenden quantitativen Menge des Genproduktes von *bCLCA4* korreliert, liegt es nahe Mechanismen zu untersuchen, welche die Expressionsrate dieses Genes in Abhängigkeit der beschriebenen Haplotypen beeinflussen können.

So werden in diesem Abschnitt die Ergebnisse von verschiedenen funktionellen *in silico* Analysen der Kernhaplotypen dargestellt. In Kapitel 3.6.1 wird die Auswirkung der Allele von den Markern I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} auf bestimmte strukturchemische Eigenschaften der DNA im sequenzierten Abschnitt analysiert und in Kapitel 3.6.2 wird der Einfluß dieser Marker auf die Bindung von Transkriptionsfaktoren untersucht.

3.6.1 Auswirkung auf die physikalischen Eigenschaften der DNA – funktionelle Annotation mit der Genome-Atlas-Software

Bei der Transkription von DNA in RNA sind viele verschiedene Proteine beteiligt. Da die Interaktion dieser Proteine mit der DNA neben der Basenzusammensetzung möglicherweise auch von DNA-Konformation abhängig ist, wurde im folgenden in Kooperation mit Prof. David W. Ussery (The Technical University of Denmark, Center for Biological Sequence Analysis, Department of Biotechnology, Lyngby, Denmark) der Einfluß der Allele von Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} auf verschiedene strukturchemische Eigenschaften der DNA untersucht.

Dabei wurde der Frage nachgegangen, ob es Hinweise gibt, weshalb der Kernhaplotyp A im Gegensatz zu den übrigen Varianten B, C und D, vorwiegend bei den Patienten ohne meßbare Chloridionenleitfähigkeit nachgewiesen werden konnte.

So werden in diesem Abschnitt die Ergebnisse der *in silico* Analyse der Kernhaplotypen A, B, C, und D bezüglich der ‚Basenstapelungsenergie‘ (Abb. 34), der ‚Krümmung des DNA-Rückgrates‘ (Abb. 35) und der ‚Streß induzierten DNA-Doppelstrang Öffnungswahrscheinlichkeit‘ (Abb. 36) dargelegt.

Basenstapelungsenergie: Notwendige Dissoziationsenergie für das Auftrennen von intramolekularen Wechselwirkungen zwischen übereinander angeordneten π -Elektronen-Systemen in Einzelstrang-DNA-Molekülen

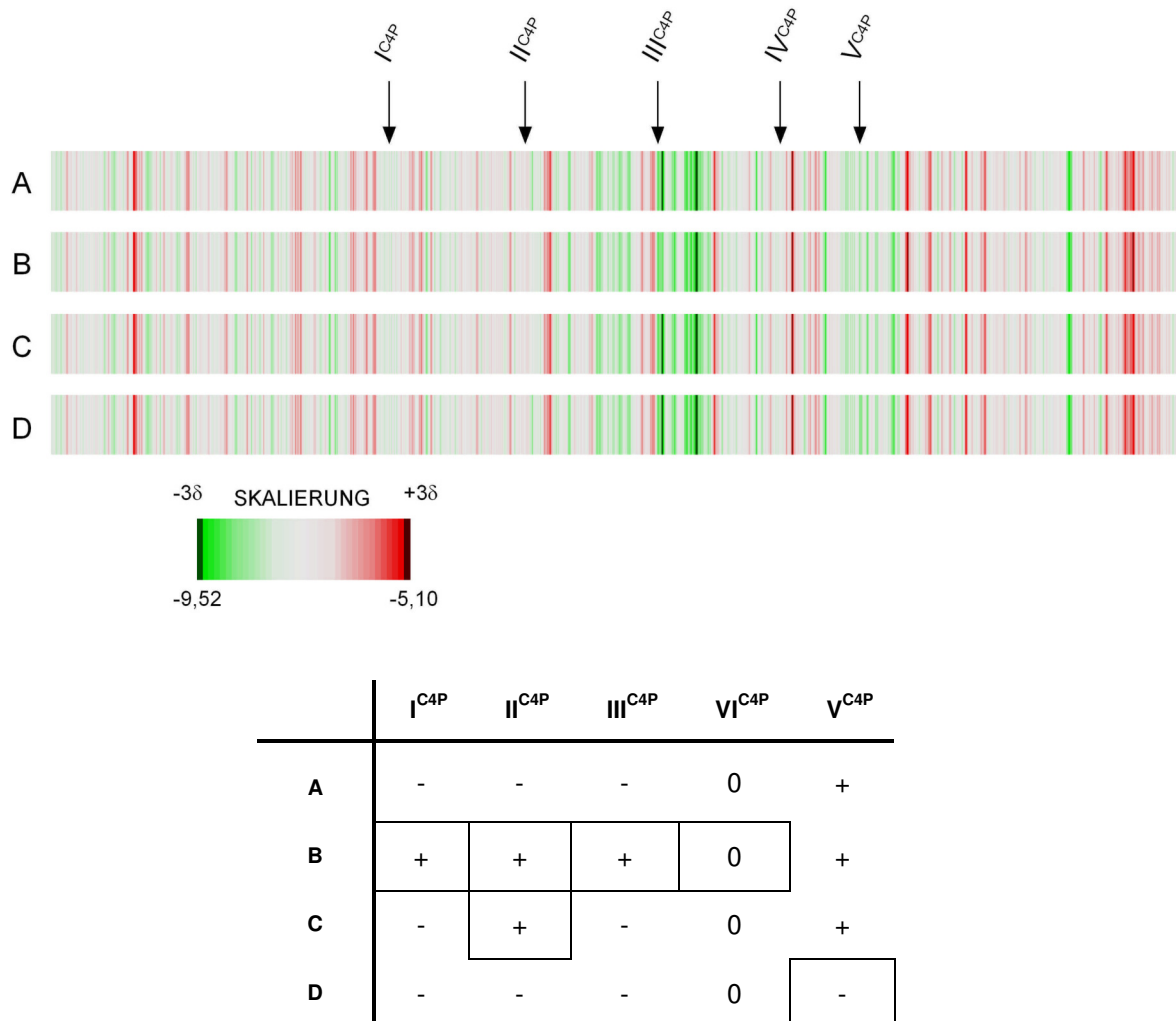


Abb. 34: Im oberen Teil sind die Analysen der Kernhaplotypen A, B, C, und D gezeigt, zusätzlich sind die Positionen der SNPs I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} markiert.

Die berechnete Basenstapelungsenergie ist in Abhängigkeit der Basenposition dargestellt: Kleine Werte repräsentieren stabilere Abschnitte, größere Beträge zeigen dagegen eine geringere Stabilität der DNA an.

Die Skalierung der Darstellung beruht auf einer relativen Betrachtung der ermittelten Werte: So repräsentieren die Endpunkte der Skala jeweils die Beträge, welche der Abweichung vom Mittelwert um drei Standardabweichungen entsprechen.

Die Tabelle faßt den Vergleich der Basenstapelungsenergie zwischen den vier Haplotypen an den Positionen der Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} zusammen: Ein Pluszeichen (+) bedeutet ein im Vergleich zu den übrigen Haplotypen instabileren Abschnitt und ein Minuszeichen (-) zeigt einen in Relation zu den übrigen Haplotypen stabileren Abschnitt der DNA an.

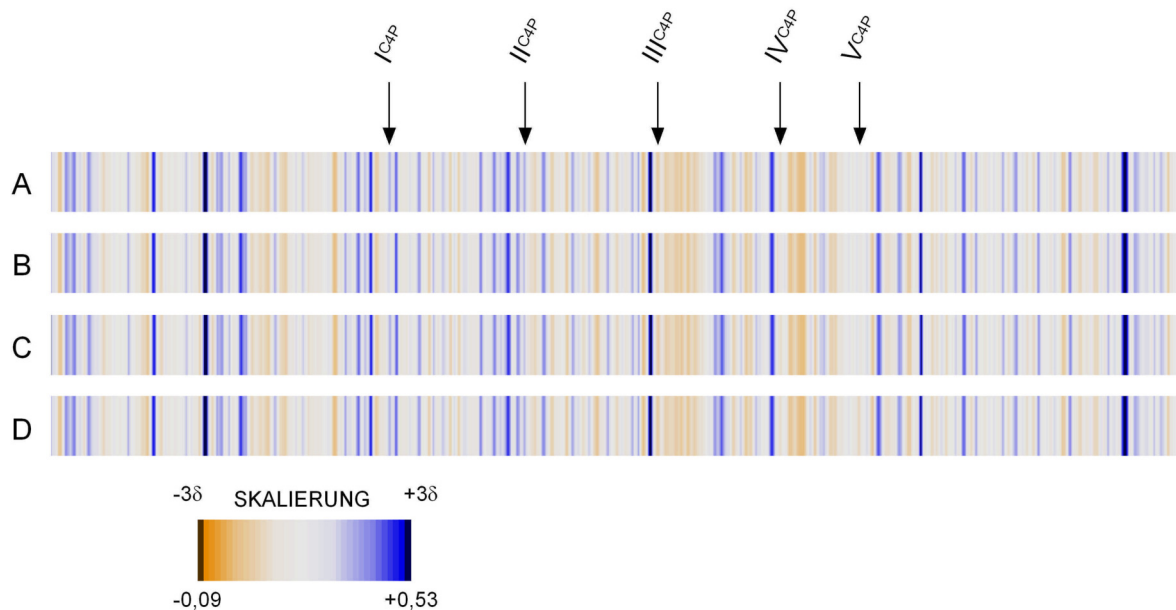
Überdies sind die Marker, anhand welcher sich die Haplotypen B, C und D von der Variante A durch unterschiedliche Allele abgrenzen lassen, mit halbfett gezeichneten Quadraten gekennzeichnet.

Abb. 34 ist zu entnehmen, daß sich die vier Kernhaplotypen im Parameter der ‚Basenstapelungsenergie‘ an den Positionen der Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} unterscheiden. Es wird deutlich, daß inverse Verhältnisse vorliegen: So weist die Analyse bei Haplotyp B an den Positionen der Marker I^{C4P}, II^{C4P}, III^{C4P} und IV^{C4P} auf instabilere

Verhältnisse als bei Variante A hin. Auch der Haplotyp C zeigt bei Marker II^{C4P} einen im Vergleich zum Kernhaplotypen A instabileren Abschnitt.

Der Haplotyp D weist dagegen bei Marker V^{C4P} einen stabileren Bereich als Kernhaplotypen A auf.

Krümmung des DNA-Rückgrates: Ein Maß für den Grad der Bogenbildung von DNA-Molekülen



	I ^{C4P}	II ^{C4P}	III ^{C4P}	IV ^{C4P}	V ^{C4P}
A	+	+	-	-	+
B	-	-	+	+	+
C	+	-	-	-	+
D	+	+	-	-	-

Abb. 35: Im oberen Teil sind die Analysen der Kernhaplotypen A, B, C, und D gezeigt, zusätzlich sind die Positionen der SNPs I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} markiert.

Die berechnete Bogenbildung der DNA ist in Abhängigkeit der Basenposition dargestellt: Kleine Werte repräsentieren eine Bogenbildung mit einem größeren Radius, größere Beträge zeigen dagegen eine Ausbildung von Bögen mit kleinem Radius an.

Die Skalierung der Darstellung beruht auf einer relativen Betrachtung der ermittelten Werte: So repräsentieren die Endpunkte der Skala jeweils die Beträge, welche der Abweichung vom Mittelwert um drei Standardabweichungen entsprechen.

Die Tabelle faßt den Vergleich zwischen den vier Haplotypen an den Positionen der Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} zusammen: Ein Pluszeichen (+) bedeutet die Ausbildung von Bögen mit einem im Vergleich zu den übrigen Haplotypen kleineren Radius und ein Minuszeichen (-) zeigt eine Bogenbildung mit einem in Relation zu den übrigen Haplotypen größeren Radius an. (→)

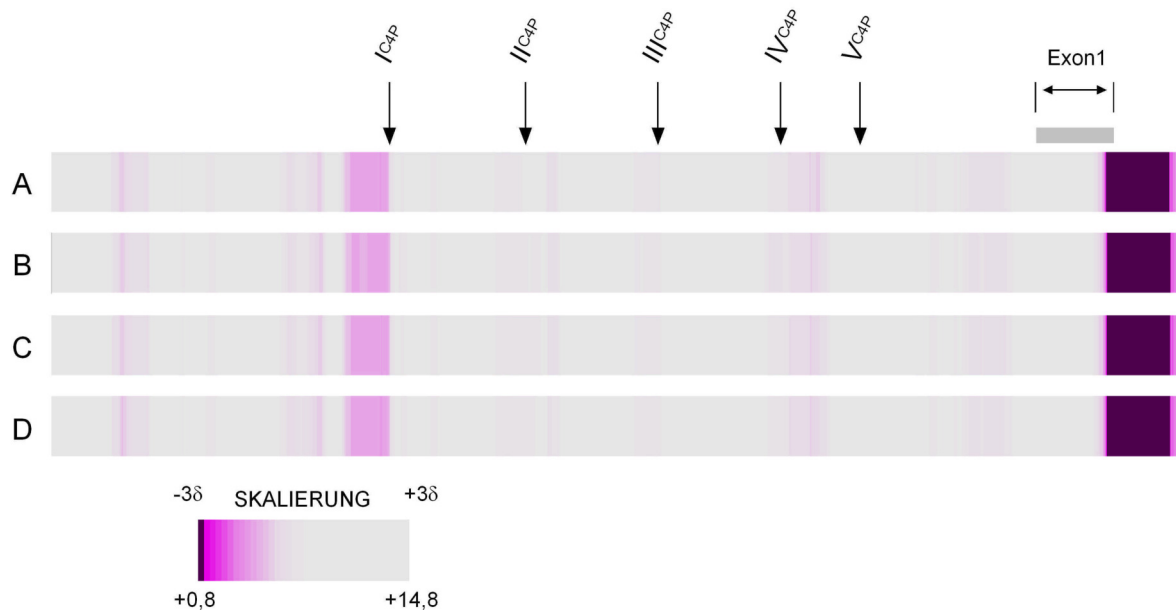
Überdies sind die Marker, anhand welcher sich die Haplotypen B, C und D von der Variante A durch unterschiedliche Allele abgrenzen lassen, mit halbfett gezeichneten Quadraten gekennzeichnet.

Abb. 35 ist zu entnehmen, daß bei den Haplotypen B, C, und D an den Positionen der Marker, welche eine Unterscheidung vom Kernhaplotypen A durch Genotypisierung zulassen, auch ein Unterschied im Ausmaß der Bogenbildung vorliegt. Es wird deutlich, daß inverse Verhältnisse vorliegen:

So weist der Haplotyp B im Vergleich zur Variante A bei Marker I^{C4P} und II^{C4P} die Ausbildung eines weiteren Bogens sowie bei SNP III^{C4P} und IV^{C4P} die eines engeren Bogens auf.

Der Kernhaplotyp C unterscheidet sich vom Haplotyp A an der Position von SNP II^{C4P} durch eine verminderte Bogenbildung, und der Kernhaplotyp D weicht von der Variante A bei SNP V^{C4P} ebenfalls durch die Ausbildung einer verminderten Bogenbildung ab.

Streß induzierten DNA-Doppelstrang Öffnungswahrscheinlichkeit



	I ^{C4P}	II ^{C4P}	III ^{C4P}	IV ^{C4P}	V ^{C4P}
A	0	0	0	0	0
B	0	0	0	0	0
C	0	0	0	0	0
D	0	0	0	0	0

Abb. 36: Im oberen Teil sind die Analysen der Kernhaplotypen A, B, C, und D gezeigt, zusätzlich sind die Positionen der SNPs I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} markiert.

Die Streß induzierte DNA-Doppelstrang Öffnungswahrscheinlichkeit ist in Abhängigkeit der Basenposition dargestellt: Kleine Werte repräsentieren Abschnitte mit großer Öffnungswahrscheinlichkeit, größere Beträge zeigen dagegen Bereiche mit einer geringeren Öffnungswahrscheinlichkeit der DNA an.

Die Skalierung der Darstellung beruht auf einer relativen Betrachtung der ermittelten Werte: So repräsentieren die Endpunkte der Skala jeweils die Beträge, welche der Abweichung vom Mittelwert um drei Standardabweichungen entsprechen.

Die Tabelle faßt den Vergleich der Streß induzierten DNA-Doppelstrang Öffnungswahrscheinlichkeit zwischen den vier Haplotypen an den Positionen der Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} zusammen.

Überdies sind die Marker, anhand welcher sich die Haplotypen B, C und D von der Variante A durch unterschiedliche Allele abgrenzen lassen, mit halbfett gezeichneten Quadraten gekennzeichnet.

Abb. 36 ist zu entnehmen, daß sich die vier Kernhaplotypen A, B, C und D an den Positionen der Marker I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} im Parameter der ‚Streß induzierten DNA-Doppelstrang Öffnungswahrscheinlichkeit‘ nicht unterscheiden. Es weisen jedoch alle

Haplotypen am 3'-Ende des Exon 1 einen Bereich mit sehr hoher Öffnungswahrscheinlichkeit der DNA-Doppelhelix auf, welcher sich bis in das erste Intron fortsetzt.

3.6.2 Auswirkung auf die Bindung von Transkriptionsfaktoren

Die Bindung von Transkriptionsfaktoren an die DNA im Promoterbereich ist ein essentieller Schritt bei der Transkription und Expression von Genen. Nimmt man an, daß die alternative Chloridionenleitfähigkeit mit einer zur Vergleichsgruppe erhöhten Expressionsrate des Genes *bCLCA4* korreliert, liegt es nahe, das Bindungsverhalten von Transkriptionsfaktoren in Abhängigkeit der vier Kernhaplotypen zu untersuchen.

Für diesen Zweck wurde ein möglicher Einfluß der detektierten Sequenzvariationen I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} auf die Bindung von Transkriptionsfaktoren untersucht. Dazu wurden mit Hilfe der Teilfunktion MatInspector und SNPInspector der Datenbank Genomatix Bindungsstellen von bekannten Transkriptionsfaktoren vorhergesagt (siehe Kapitel 2.12).

So wurden für den sequenzierten DNA-Abschnitt, insgesamt 578 verschiedene putative Bindungspositionen für Transkriptionsfaktorfamilien ermittelt. Eine Änderung, verursacht durch die unterschiedlichen Allele der Marker I^{C4P} bis V^{C4P}, ergab sich für 16 Familien. Diese umfassen 93 bekannte Transkriptionsfaktoren.

Für 74 von diesen waren Daten aus der Transkriptomanalyse (siehe Kapitel 2.13) verfügbar, anhand derer eine Abschätzung der Expression im untersuchten Gewebe der Rektum-Schleimhaut erfolgte.

Dazu wurde separat für jeden Transkriptionsfaktor der Median aus den Werten der einzelnen Probanden bestimmt und eine Expression angenommen, wenn dieser größer als 300 war.

So konnten schließlich 18 verschiedene Transkriptionsfaktoren ermittelt werden, die in relevanter Anzahl exprimiert werden und deren Bindung an die DNA durch die Allele der Marker I^{C4P} bis V^{C4P} beeinflusst wird.

Eine Auflistung dieser Transkriptionsfaktoren findet sich in Tabelle 12 und eine Darstellung in Annotation zu den vier Kernhaplotypen A bis D wird in Abb. 37 wiedergegeben.

Für die übrigen 19 Transkriptionsfaktoren, für welche keine Daten aus der Transkriptomanalyse verfügbar waren, wurde eine Expression mittels der UniGene-Datenbank nach dem in Kapitel 2.14 beschriebenen Verfahren abgeschätzt.

Dabei wurde eine relevante Expression im Kolon angenommen, wenn die normierte Häufigkeit des Klonen vom betreffenden Gen größer als 2,5% war. Dieses Kriterium war für vier der 19 Transkriptionsfaktoren erfüllt. Sie sind in Tabelle 13 aufgelistet und in Abb. 38 in Annotation zu den vier Kernhaplotypen dargestellt.

Auflistung der mittels Genomatix-Datenbank ermittelten Transkriptionsfaktoren, deren Expression mit Transkriptomdaten aus Affymetrix-Chips abgeschätzt wurde

Positionsnummer	Symbol	Name (HUGO Gene Nomenclature Committee)
1	ATF3	activating transcription factor 3
2	ATF4	activating transcription factor 4 (tax-responsive enhancer element B67)
3	CREB1	cAMP responsive element binding protein 1
4	CREB3	cAMP responsive element binding protein 3
5	CREBBP	CREB binding protein (Rubinstein-Taybi syndrome)
6	CREM	cAMP responsive element modulator
7	ESRRA	estrogen-related receptor alpha
8	NFIL3	nuclear factor, interleukin 3 regulated
9	GATA2	GATA binding protein 2
10	GATA6	GATA binding protein 6
11	TEAD3	TEA domain family member 3
15	TP53	tumor protein p53 (Li-Fraumeni syndrome)
18	CIC	capicua homolog (Drosophila)
19	HBP1	HMG-box transcription factor 1
20	HMGA1	high mobility group AT-hook 1
21	HMGB3	high-mobility group box 3
22	SOX9	SRY (sex determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal)
23	RREB1	ras responsive element binding protein 1

Tabelle 12

Differenzielle Bindungsspektren der Kernhaplotypen A, B, C und D der Transkriptionsfaktoren, für welche die Expression mittels Transkriptom-Daten aus Affymetrix-Chips abgeschätzt wurde

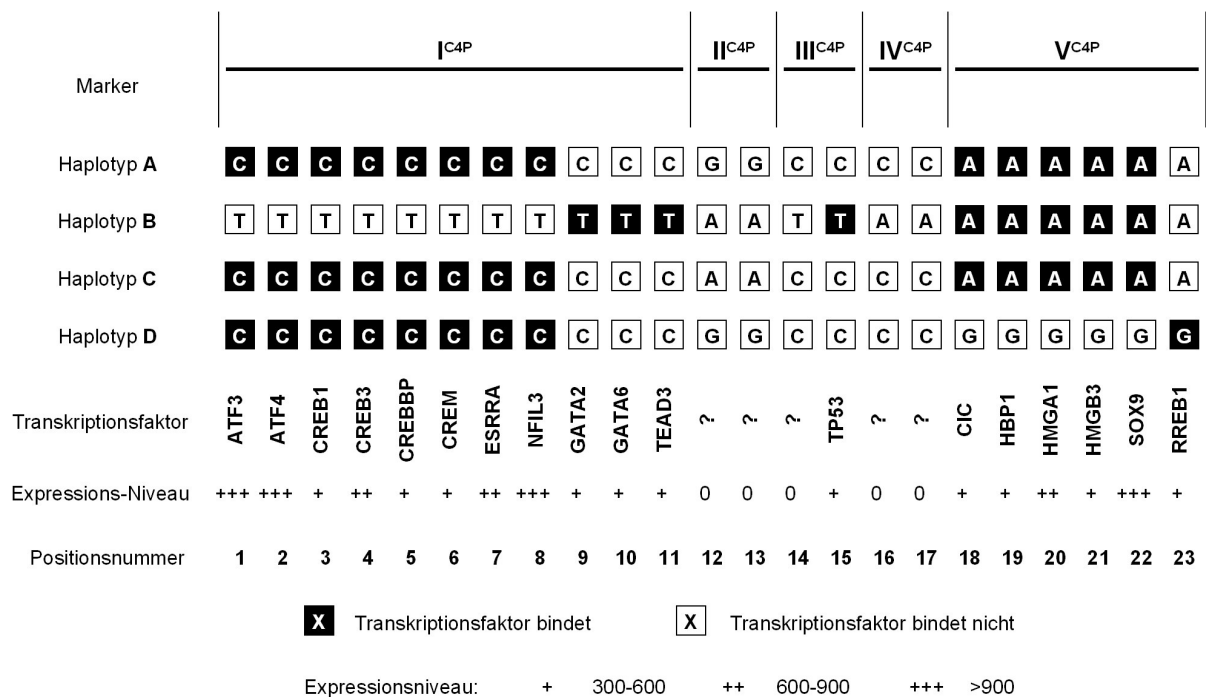


Abb. 37: Es sind die vier Kernhaplotypen A, B, C und D dargestellt. Zur besseren Übersichtlichkeit sind im oberen Bereich der Abbildung die Positionen der SNPs I^{C4P} bis V^{C4P} markiert, aus denen die Haplotypen zusammengesetzt sind.

Der Einfluß der Allele dieser Marker auf die Bindung eines Transkriptionsfaktors ist durch ein Quadrat symbolisiert; ist dieses schwarz ausgefüllt, so zeigt es die Bindung des entsprechenden Transkriptionsfaktors an, ist es in der Farbe Weiß dargestellt, hybridisiert dieser mit der DNA nicht.

So bindet z.B. der Transkriptionsfaktor ATF3 mit der Positionsnummer 1 nur, wenn der Marker I^{C4P} das Allel C trägt, dieses ist für die Kernhaplotypen A, C und D gegeben, jedoch nicht für den Haplotypen B.

Diese Darstellung gibt demnach nur diejenigen Transkriptionsfaktoren wieder, deren Bindung an die DNA durch die Allele der SNPs I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} beeinflusst wird. Damit gibt diese Zeichnung differenzielle Bindungsspektren wieder.

An den Positionsnummern 12, 13, 14, 16 und 17 konnten mit dem beschriebenen Verfahren (siehe Text) keine Transkriptionsfaktoren ermittelt werden, sodaß diese jeweils mit einem Fragezeichen (?) markiert sind.

Im Abb. 37 sind die ermittelten Transkriptionsfaktoren in Annotation der vier Kernhaplotypen A, B, C und D (siehe Tabelle 10, Seite 81) dargestellt. Die Expression dieser Proteine im Rektum wurde mittels Transkriptom-Analyse abgeschätzt.

So binden bei Haplotyp A die Transkriptionsfaktoren mit den Positionsnummern 1-8 sowie die Faktoren mit den Nummern 18-22. Die gleichen binden auf dem Haplotypen C, sodaß sich beide Varianten im Spektrum der Transkriptionsfaktoren nicht unterscheiden.

Der Haplotyp D grenzt sich von diesen beiden durch die fehlenden Transkriptionsfaktoren mit den Positionsnummern 18-22 ab; bei diesem bindet jedoch zusätzlich der Faktor RREB1 mit der Positionsnummer 23.

Der Kernhaplotyp B weicht von den Varianten A und C durch die fehlenden Transkriptionsfaktoren mit den Nummern 1-8 ab. Bei diesem Haplotyp binden jedoch zusätzlich die Transkriptionsfaktoren GATA2, GATA6 und TEAD3 mit der Positionsnummer 9-11 sowie der Faktor TP53 mit der Positionsnummer 15.

Auflistung der mittels Genomatix-Datenbank ermittelten Transkriptionsfaktoren, deren Expression mit der UniGene-Datenbank abgeschätzt wurde.

Positionsnummer	Symbol	Name (HUGO Gene Nomenclature Committee)
2	TEAD2	TEA domain family member 2
3	DDIT3	DNA-damage-inducible transcript 3
4	POU6F2	POU domain, class 6, transcription factor 2
8	NKX2-3	NK2 transcription factor related, locus 3 (Drosophila)

Tabelle 13

Differenzielle Bindungsspektren der Kernhaplotypen A, B, C und D der Transkriptionsfaktoren für welche die Expression mittels UniGene-Datenbank abgeschätzt wurde.

Marker	I ^{C4P}	II ^{C4P}	III ^{C4P}	IV ^{C4P}	V ^{C4P}					
Haplotyp A	C	C	G	G	C	C	C	C	A	A
Haplotyp B	T	T	A	A	T	T	A	A	A	A
Haplotyp C	C	C	A	A	C	C	C	C	A	A
Haplotyp D	C	C	G	G	C	C	C	C	C	C
Transkriptionsfaktor	?	TEAD2	DDIT3	POU6F2	?	?	?	NKX2-3	?	?
Positionsnummer	1	2	3	4	5	6	7	8	9	10

Transkriptionsfaktor bindet
 Transkriptionsfaktor bindet nicht

Abb. 38: Es sind die vier Kernhaplotypen A, B, C und D dargestellt. Zur besseren Übersichtlichkeit sind im oberen Bereich der Abbildung die Positionen der SNPs I^{C4P} bis V^{C4P} markiert, aus denen die Haplotypen zusammengesetzt sind.

Der Einfluß der Allele dieser Marker auf die Bindung eines Transkriptionsfaktors ist durch ein Quadrat symbolisiert; ist dieses schwarz ausgefüllt, so zeigt es die Bindung des entsprechenden Transkriptionsfaktors an, ist es in der Farbe Weiß dargestellt, hybridisiert dieser mit der DNA nicht.

So bindet z.B. der Transkriptionsfaktor TEAD2 mit der Positionsnummer 2 nur, wenn der Marker I^{C4P} das Allel T trägt, dieses ist für den Kernhaplotypen B gegeben, jedoch nicht für die Haplotypen A, C oder D.

Diese Darstellung gibt demnach nur diejenigen Transkriptionsfaktoren wieder, deren Bindung an die DNA durch die Allele der SNPs I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P} beeinflusst wird. Damit gibt diese Zeichnung differenzielle Bindungsspektren wieder.

An den Positionsnummern 1, 5, 6, 7, 9 und 10 konnten mit dem beschriebenen Verfahren (siehe Text) keine Transkriptionsfaktoren ermittelt werden, sodaß diese jeweils mit einem Fragezeichen (?) markiert sind.

In Abb. 38 sind die Transkriptionsfaktoren dargestellt, deren Expression im Kolon mittels UniGene-Datenbank abgeschätzt wurde. Es zeigt sich, daß im Gegensatz zur vorhergehenden Darstellung (Abb. 37) hier Faktoren ermittelt werden konnten, deren Hybridisierung an die DNA durch die Allele von Marker II^{C4P} beeinflußt wird. Damit kann jetzt der Kernhaplotyp C von der Variante A differenziert werden. Denn an diesen binden die Faktoren DDIT3 und POU6F2. Überdies binden an Kernhaplotyp B neben den genannten zusätzlich die Faktoren TEAD2 und NKX2-3.

4 Diskussion

4.1 Kartierung von Sequenzvariationen in der Kandidatengenregion mittels Allelvergleich von genetischen Markern zwischen Studienpopulationen mit kontrastierenden Phänotypen

Das *CLCA*-Gen-Cluster wurde in Vorarbeiten als ein möglicher Modulator des gastrointestinalen Phänotypes bei Mukoviszidose-Patienten identifiziert [62]. Diese Genfamilie besteht aus 4 Homologen und ist auf dem kurzen Arm des Chromosomes 1 lokalisiert. Ziel dieser Arbeit war, die Frage zu beantworten, ob es möglich ist, den Modulator weiter einzugrenzen und gegebenenfalls einem der vier *CLCA*-Gene zuzuordnen.

Dazu wurden anhand von Datenbankeinträgen SNPs im *CLCA*-Gen-Cluster bestimmt und diese bei CF-Patienten mit alternativer Chloridionenleitfähigkeit und ohne nachweisbare Chloridionenleitfähigkeit typisiert. Aus den gewonnenen Daten wurden zwei-Marker-Haplotypen rekonstruiert und die Verteilung dieser zwischen den beiden Studiengruppen verglichen.

Dabei zeigte sich, daß drei benachbarte Haplotypen, bestehend aus den Markern $7^{C1/4}$ _ 8^{C4IN1} , 8^{C4IN1} _ 9^{C4IN1} und 9^{C4IN1} _ 10^{C4IN1} signifikant unterschiedlich verteilt sind oder einen Trend für eine unterschiedliche Verteilung zwischen den miteinander verglichenen Patientengruppen angeben (siehe Kapitel 3.2.2.2). Die Marker $7^{C1/4}$ bis 10^{C4IN1} überspannen einen 11,6 kbp großen Bereich in der 5'-Region von *bCLCA4* und decken damit den nahen Promoter, das Exon 1 sowie einen Teil des ersten Introns des Genes ab (Abb. 20, Seite 64 und Abb. 22, Seite 70), sodaß dieses Homolog vermutlich den Modulator des gastrointestinalen Phänotypes kodiert.

Um falsch positive Ergebnisse der Assoziationsstudie zu vermeiden, wurde die bekannte Einflußgröße, des Alters der Probanden auf die Signalstärke des Chloridionenstromflusses [75], durch Stratifizierung der Kontroll-Gruppe minimiert und (Methodenteil Kapitel 2.9.1.1). Die Auswertung der Studie zeigte, daß durch die Stratifizierung *p*-Werte mit höherer Signifikanz erreicht werden, als bei Vergleich der gesamten Kontrollpopulation mit den CF-Patienten mit alternativer Chloridionenleitfähigkeit. Damit konnte hier der Einfluß des Alters auf das Chloridionensignal bestätigt und die Beeinträchtigung der Assoziationsstudie durch diese Störgröße vermindert werden.

Eine weitere Störgröße kann auftreten, wenn Patienten mit verschiedenem ethnischen Hintergrund (engl. admixture) verglichen werden, denn hier würden per se unterschiedliche

Allelfrequenzen der Marker vorliegen [88, 89]. In der Studie sind mehrere Loci typisiert worden, bei denen keine signifikanten Unterschiede in der Allelhäufigkeiten detektierbar waren, sodaß ein falsch positives Ergebnis durch einen Vergleich von Patienten mit verschiedenem ethnischen Hintergrund weitgehend ausgeschlossen werden kann.

In einer Expressionsstudie konnte die Transkription des Genes *bCLCA4* in 12 verschiedenen Geweben durch Northernblotting-Analysen nachgewiesen werden. Das stärkste Signal wurde dabei im Gewebe vom Kolon gemessen [48]. Des weiteren ergab eine Datenbankanfrage (SOURCE-Datenbank, Stand: August 2006) [86], daß das Gen *bCLCA4* in 17 untersuchten Geweben im Kolon am siebtstärksten exprimiert wird.

Damit ist das ermittelte Ergebnis, *bCLCA4* kodiert den Modulator des gastrointestinalen Phänotyps, konsistent mit den Befunden der Literatur.

Weiterhin wurde eine Analyse der Haplotypblockstruktur im *CLCA*-Gen-Cluster durch Anwendung des Vier-Gameten-Testes [59, 60] durchgeführt. Dabei konnte ein Bereich detektiert werden, für welchen sich keine stammesgeschichtlich fixierten Rekombinationsbruchpunkte ableiten lassen. Dieser entspricht dem durch SNP 7^{C1/4} bis 10^{C4IN1} markierten Abschnitt, welcher zwischen den Studienpopulationen unterschiedlich verteilt ist.

Dies kann zum einen durch eine geringe Mutations- oder Rekombinationsrate im betreffenden Abschnitt gedeutet werden und zum anderen durch populationsgeschichtliche Aspekte, wie Selektionsnachteile von Individuen mit bestimmten Haplotypen.

In beiden Fällen sind nur einige wenige Haplotypen in der untersuchten Population vorhanden, dies spricht dafür, daß es sich hier um eine für die Individuen der untersuchten Studienpopulation relevante genetische Region handelt.

Um die These zu prüfen, ob in dieser Region weitere bisher nicht bekannte DNA-Variationen vorliegen, wurde eine Teil-Resequenzierung der Region durchgeführt. So konnten bei insgesamt vier Indexpatienten mit kontrastierenden Haplotypen aus beiden Studiengruppen fünf Sequenzvariationen (SNPs) in der Kombination der Kernhaplotypen A, B, C und D ermittelt werden.

Der hier verfolgte Ansatz zur Feinkartierung des genetischen Modulators im *CLCA*-Gen-Cluster für den gastrointestinalen Phänotyp bei CF-Patienten wurde in der aktuelleren Literatur bereits vereinzelt in ähnlicher Weise beschrieben. So wurde diese Vorgehensweise zur Identifikation von Sequenzvariation im *GATA2*-Genlocus bei der Untersuchung einer familiären Form der Koronararterien-Krankheit [90] angewendet oder bei der Ermittlung von DNA-Variationen im *MET*-Genlocus bei einer Studie zur Untersuchung der psychischen Erkrankung Autismus [91] angewandt.

4.2 Bewertung der Sequenzierergergebnisse – Fehlerdiskussion

Wie in Kapitel 3.3 des Ergebnisteiles dargelegt ist, wurden von dem 11,6 kbp großen Fragment, welches mit dem gastrointestinalen Phänotyp der CF-Patienten korreliert, lediglich etwa 3 kbp resequenziert. Als initiales Target wurde dabei ein Abschnitt ausgewählt, von dem angenommen werden kann, daß DNA-Sequenzvariationen hier einen großen Einfluß auf die Expression des Genes oder die Funktion des synthetisierten Proteins haben. So umfaßt der sequenzierte Bereich das erste Exon, in welchem Variationen des genetischen Codes zu Aminosäurepolymorphismen führen können und so die Struktur und Funktion des synthetisierten Proteins verändern können. Auch wurde durch die Resequenzierung der nahe Promoter des Genes erfaßt, in welchem Sequenzvariationen Einfluß auf die Bindung von Transkriptionsfaktoren haben können, die die Expressionsrate des Genes modulieren.

Mit der Resequenzierung dieses 3 kbp großen Abschnittes ist jedoch nicht das gesamte Fragment erfaßt worden, welches zwischen den Patienten der beiden miteinander verglichenen Studiengruppen unterschiedlich verteilt ist. Da es sich bei diesem um einen hoch konservierten genetischen Bereich handelt (siehe Kapitel 3.4.2), ist für eine erschöpfende Beantwortung der Frage nach Sequenzvariationen eine vollständige Sequenzierung notwendig.

Durch eine Verifizierung der Sequenzierergergebnisse konnten die Genotypen der SNPs II^{C4P} und V^{C4P} von homozygot nach heterozygot korrigiert werden. Es ist anzumerken, daß die Spurdaten dieser SNPs keine doppelten Maxima für zwei verschiedene Basen aufweisen, wie es bei heterozygoten Genotypen zu erwarten ist, sondern sie zeigen sehr saubere Signale ohne jegliches Hintergrundrauschen (Abb 27, Seite 79). Dies zeigt, daß die automatisierten Sequenzierungsverfahren, welche möglicherweise für die Analyse von haploiden bakteriellen Genomen und einen hohen Durchsatz ausgelegt sind, problematisch bei der Bestimmung der Sequenz von diploiden Genomen mit heterozygoten Genotypen sind.

Abschließend ist zu bemerken, daß die durch Nachtypisierung ermittelten Allele der detektierten Marker in den Kontext der Haplotypen passen, welche zur Auswahl der Proben für die Sequenzierung verwendet worden sind (siehe Tabelle 9, Seite 80).

So wurde der Drei-Marker-Haplotyp aus den Markern 5^{C1EX15}, 9^{C4IN1} und 12^{C4IN6} konstruiert und die Kernhaplotypen aus den SNPs I^{C4P}, II^{C4P}, III^{C4P}, IV^{C4P} und V^{C4P}. Der Kernhaplotyp D war dabei jeweils Teil des Drei-Marker-Haplotypen 2-2-2, der Kernhaplotyp C Teil des Drei-Marker-Haplotypen 1-1-2 und die Variante B Teil des Haplotypen 1-1-1. Der Kernhaplotyp A wurde nur als Teil des Drei-Marker-Haplotypen 1-2-1 beobachtet.

Die Ergebnisse sind damit in sich schlüssig und es kann von deren Richtigkeit ausgegangen werden.

4.3 Mögliche Rolle der Sequenzvariationen bei der Kodierung des Modulators der alternativen Chloridionenleitfähigkeit

Durch die Typisierung der detektierten Sequenzvariationen in der gesamten Studienpopulation konnten die Diplotypen der Patienten, bestehend aus den Haplotypen A, B, C und D, rekonstruiert werden.

Wie in Abb. 32 auf Seite 88 zusammenfassend dargestellt, zeigte sich, daß zwei Drittel der Patienten mit alternativer Chloridionenleitfähigkeit den Haplotypen A auf keinem Chromosom tragen und ein Drittel der Patienten mit diesem Phänotyp die Variante A in Kombination mit den Kernhaplotypen B, C oder D aufweisen.

Dagegen sind etwa ein Drittel der Patienten ohne nachweisbare Chloridionenleitfähigkeit homozygot für den Haplotyp A und ca. die Hälfte dieser Patienten trägt die Variante A neben einer der Haplotypen B, C oder D nur auf einem Chromosom. Lediglich ein Fünftel der Probanden ohne residuelle Chloridionenleitfähigkeit tragen die Varianten B, C oder D auf beiden Chromosomen.

Während bei denjenigen Patienten, bei welchen die Haplotypen B, C oder D auf beiden Chromosomen kodiert sind, vorwiegend eine alternative Chloridionenleitfähigkeit nachweisbar ist, kann bei allen Probanden, die homozygot für den Kernhaplotyp A sind, kein residueller Chloridionenstromfluß gemessen werden.

In der Gruppe der Probanden, welche den Haplotyp A neben einer der Varianten B, C oder D nur auf einem Chromosom tragen, finden sich sowohl Patienten mit alternativer Chloridionenleitfähigkeit als auch Probanden ohne nachweisbaren Chloridionenstromfluß.

Aus diesen Befunden läßt sich schlußfolgern, daß der Haplotyp A die alternative Chloridionenleitfähigkeit nicht vermitteln kann und dagegen die Varianten B, C oder D notwendige Voraussetzung für die Expression der DIDS-sensitiven Chloridionenleitfähigkeit sind. Da die Haplotypen B, C oder D jedoch nicht obligat zur Ausbildung der alternativen Chloridionenleitfähigkeit führen, muß es weitere Faktoren geben, die Einfluß auf die Ausprägung der Chloridionenleitfähigkeit haben.

Wie im Ergebnisteil dargestellt wurde, weisen die drei Kernhaplotypen B, C und D keine gemeinsamen Sequenzvariationen auf, anhand derer sie sich von der Variante A differenzieren lassen.

Folglich kann aus den gewonnenen Daten nicht unmittelbar abgeleitet werden, welche der Sequenzvariationen kausal für die Expression des hier untersuchten Phänotyps der alternativen Chloridionenleitfähigkeit ist. Nimmt man an, daß jede der einzelnen Variationen zur Ausbildung der alternativen Chloridionenleitfähigkeit führen kann, muß es zu den Sequenzunterschieden der DNA übergeordnete Mechanismen geben, die die Expression der

DIDS-sensitiven Chloridionenleitfähigkeit bei den Trägern der Kernhaplotypen B, C oder D erlauben. Um Hinweise für diese kausalen Mechanismen zu bekommen, wurden verschiedene *in silico* Analysen durchgeführt. In den folgenden beiden Kapiteln werden die Resultate dieser Untersuchungen diskutiert.

4.3.1 Differenzierung der Kernhaplotypen nach DNA-Struktur und makromolekularen Verhalten

In Kapitel 3.6.1 wurden die Ergebnisse der Vorhersage verschiedener strukturemischer Eigenschaften der DNA im sequenzierten Abschnitt gezeigt. So konnten bei den Parametern der ‚Basenstapelungsenergie‘ [81] und der ‚Krümmung des DNA-Rückgrates‘ [82] Unterschiede zwischen dem Kernhaplotypen A und den Varianten B, C sowie D vorausgesagt werden.

Im Parameter der Basenstapelungsenergie unterscheiden sich die Varianten B und C an den diskriminierenden Markern I^{C4IN1} , II^{C4IN1} , III^{C4IN1} und IV^{C4IN1} vom Kernhaplotypen A durch größere Werte, sodaß hier energetisch instabilere Abschnitte vorliegen. Dagegen wird bei Haplotyp D am Locus V^{C4IN1} ein kleinerer Wert und damit stabilerer Bereich vorhergesagt als beim Referenzhaplotypen A.

Die Berechnungen der DNA-Krümmung zeigen ein komplexeres Muster. So weist der Haplotyp B in der Analyse bei den Loci I^{C4IN1} und II^{C4IN1} eine schwächere Krümmung und bei den Loci III^{C4IN1} und IV^{C4IN1} eine stärkere Krümmung als der Referenzhaplotyp A bei diesen Markern auf. Die Variante C zeigt im Locus II^{C4IN1} und der Haplotyp D im Locus V^{C4IN1} jeweils eine geringere Krümmung des DNA-Rückgrates als der Kernhaplotyp A.

Aus diesen Beobachtungen lassen sich keine regelhaften Zusammenhänge zwischen der Abweichung der Haplotypen B, C und D in den berechneten Parametern vom Referenzhaplotypen A und der beobachteten Assoziation der Varianten B, C und D mit der alternativen Chloridionenleitfähigkeit ableiten.

Für den Parameter der DNA-Krümmung sind in der Literatur verschiedene experimentelle Nachweise beschrieben, daß er einen Einfluß auf die Transkriptionsrate ausübt [92, 93]. Auch für die Basenstapelungsenergie wird in der Literatur über den Einfluß auf die Transkriptionsrate spekuliert [94]. Es wurde jedoch bisher keine Methode zur Vorhersage des Einflusses dieser Parameter auf die Transkriptionsrate beschrieben, sodaß sich hier auch anhand der Literatur keine eindeutigen Zusammenhänge ableiten lassen.

Damit sind die Parameter DNA-Krümmung und Basenstapelungsenergie als kausale Mechanismen zur Erklärung für die beobachtete Assoziation zwischen Haplotyp und Phänotyp nicht ausgeschlossen, sondern sie bilden möglicherweise wichtige Einflußfaktoren auf das komplexe System der Transkriptionsmaschinerie, deren Wirkungsgrad von der

sterischen Anordnung der miteinander in Wechselwirkung stehenden DNA-bindenden Faktoren abhängig ist.

Die Berechnung der Streß induzierten DNA-Doppelstrang Öffnungswahrscheinlichkeit [83] zeigte keine wesentlichen Unterschiede zwischen den Kernhaplotypen. Bei allen vier Varianten konnte im Anfangsbereich des Genes *bCLCA4* ein Abschnitt mit hoher Öffnungswahrscheinlichkeit des DNA-Doppelstranges vorhergesagt werden.

Dieser befindet sich im Intron 1. Das Homolog *bCLCA4* scheint demnach zu der immer größer werdenden Gruppe von Genen zu gehören, bei denen sich der Doppelstrang nicht im Promoter, sondern im ersten Intron am leichtesten öffnet. Ein solcher labiler Bereich des DNA-Doppelstranges erlaubt die Interaktion mit regulatorischen Bindeproteinen [95], welche möglicherweise die Promoteraktivität des Genes steuern können. Beim humanen *CFTR*-Gen wurde eine solche Region im Intron 1 bereits im Detail experimentell charakterisiert [96]. Vergleichbare Daten zu dem Homolog *bCLCA4* sind bisher nicht beschrieben worden. Zusammenfassend kann angemerkt werden, daß es sich bei der untersuchten Region um einen interessanten genetischen Abschnitt handelt, welcher bei der Regulation der Transkriptionsrate des Genes *bCLCA4* eine wichtige Rolle zu spielen scheint.

4.3.2 Differenzierung der Kernhaplotypen nach potentiellen Transkriptionsfaktor-Bindungsstellen

Wie in Kapitel 3.6.2 dargelegt, wurden auf der Grundlage von *in silico* Analysen und der Auswertung von Mikroarray-Expressionsdaten 22 Transkriptionsfaktoren ermittelt, deren potentielle Bindung an die Promotersequenz des Genes *bCLCA4* durch die Allele der Marker I^{C4IN1} bis V^{C4IN1} beeinflusst wird und die in relevanter Kopienzahl im Kolon exprimiert werden.

Wie oben näher ausgeführt ist, wurde die alternative Chloridionenleitfähigkeit ausschließlich bei denjenigen Patienten beobachtet, welche die Kernhaplotypen B, C oder D auf mindestens einem Chromosom tragen. Daher wird vermutet, daß insbesondere diejenigen Transkriptionsfaktoren bei der differenziellen Rolle von hCLCA4 bei der Expression der alternativen Chloridionenleitfähigkeit relevant sind, welche an die Sequenz der Haplotypen B, C oder D aber nicht an A binden können. Laut *in silico* Vorhersage gilt dies für 9 der 22 ermittelten Transkriptionsfaktoren.

Wegen der potentiell zentralen Rolle dieser regulatorischen Proteine bei der Ausbildung des gastrointestinalen Phänotyps in dem untersuchten Kollektiv der Mukovizidose Patienten soll nachfolgend ein Überblick über die Funktion dieser Transkriptionsfaktoren gegeben werden.

NKX2-3:

Anhand von Untersuchungen an einem Mausmodell, konnte dem Transkriptionsfaktor NKX2-3 eine wichtige Funktion bei der Proliferation der Stammzellen in den Krypten des gastrointestinalen Epithels zugewiesen werden [97].

Da dieses Protein jedoch nicht im epithelialen sondern im mesodermalen Kompartiment des Darmes exprimiert wird [98, 99], können für seine Aufgabe bei der hCLCA4 vermittelten alternativen Chloridionenleitfähigkeit in den epithelialen Zellen keine Hypothesen abgeleitet werden.

TEAD2 und TEAD3:

Die beiden ubiquitären Proteine TEAD2 und TEAD3 werden schwerpunktmäßig im Kolon gebildet. Anhand von detaillierten Expressionsstudien mit Maus-Embryonen wurde gezeigt, daß TEAD2 vorwiegend in neurogenen Geweben, im mesenchymalen Kompartiment des Darmes und der nephrogenen Region gebildet wird [100].

Eine epitheliale Expression von TEAD2 im Darm kann nicht ausgeschlossen werden. Deshalb übt dieser Transkriptionsfaktor möglicherweise einen Einfluß auf die hCLCA4-vermittelte Expression der alternativen Chloridionenleitfähigkeit aus.

Der Faktor TEAD3 wurde aufgrund seines Hauptexpressionsortes, die Plazenta, in seiner Funktion als transkriptioneller Aktivator des humanen Chorionsomatotrophin-Genes untersucht [101, 102]. Dieser Transkriptionsfaktor scheint somit wie auch TEAD2 in der Organogenese und der embryonalen Entwicklung eine wichtige Rolle zu spielen. Damit sind zentrale Vorgänge, wie Zellproliferation, Zelldifferenzierung und Apoptose verbunden.

GATA2 und GATA6:

Die beiden Proteine GATA2 und GATA6 werden aufgrund einer konservierten DNA-Bindungsdomäne zu einer Gruppe von Transkriptionsfaktoren gezählt, die an das gemeinsame DNA-Sequenzmotiv HGATAR binden können [103].

Der Faktor GATA2 wird hauptsächlich in hämatopoetischen Zellen exprimiert [104] und GATA6 wird vornehmlich während der Embryogenese des Darmes in den proliferierenden Stammzellen transkribiert und nach terminaler Differenzierung dieser Zellen herunterreguliert [105].

Damit scheint GATA6 bei der Expression von hCLCA4 gegenüber GATA2 die relevantere Rolle zu spielen. Dieser ist vor allem bei der Zellproliferation aktiv.

TP53:

Der Transkriptionsfaktor TP53 ist ein zentrales Tumorsuppressor-Gen, welches essentieller Faktor zur Regulation des Zell-Zyklus ist. Es steuert insbesondere den Übergang von der G0 in die G1 Phase. Es wird postuliert, daß durch das regulatorische Protein TP53 Gene transkribiert werden, die das Zellwachstum und die Zellinvasion negativ regulieren [87].

POU6F2:

Der Transkriptionsfaktor POU6F2 wird nur in wenigen Geweben exprimiert, so konnte er lediglich in 14 von 46 untersuchten Geweben nachgewiesen werden. Detaillierte Expressionstudien sind bisher nur in der Retina und im zentralen Nervensystem [106] sowie in der embryonalen und adulten Niere [107] erfolgt.

Aufgrund der Ergebnisse von genetischen Untersuchungen („Verlust der Heterozygotie“) bei Patienten mit Wilms-Tumor, wird angenommen, daß es sich bei dem Protein POU6F2 um einen Tumorsuppressor handelt [108]. Es wird vermutet, daß dieser Faktor Aufgaben bei der Differenzierung der embryonalen sowie bei der Erhaltung von adulten Zellen übernimmt.

DDIT3:

Das Protein DDIT3 (DNA damage inducible transcript 3) wird zur Familie der C/EBP-Transkriptionsfaktoren (CCAAT/enhancer binding protein) gezählt [109]. Dies sind stimulatorische Faktoren, welche aufgrund einer hochkonservierten Domäne miteinander Homo- als auch Heterodimere ausbilden können, um dann an das gemeinsame DNA-Sequenzmotiv CCAAT zu binden. DDIT3 weist im Gegensatz zu den übrigen Faktoren dieser Familie eine veränderte DNA-Bindungsdomäne auf. Dadurch ist die Affinität der DDIT3-C/EBP-Heterodimere zu dem gewöhnlichen DNA-Sequenzmotiv vermindert und für eine alternative Oligonukleotid-Sequenz erhöht.

Damit verfügt das regulatorische Protein DDIT3 sowohl über die Funktion eines dominant negativen Inhibitors, als auch über die Eigenschaft eines Stimulators der Transkription [110]. Das Protein DDIT3 wird ubiquitär auf niedrigem Niveau schwerpunktmäßig im Kolon gebildet und bei zellulären Streß wird seine Transkriptionsrate heraufreguliert [110]. Dies kann zur Induktion der Apoptose führen [111].

Insbesondere stellen Einflüsse, welche die Funktion des endoplasmatischen Retikulums beeinträchtigen, Stimuli für die gesteigerte Expression von DDIT3 dar [112].

Dies kann beispielsweise durch das aberrante CFTR-Protein bei unseren Mukoviszidose-Patienten bedingt sein. So führt die Mutation F508del, die bei einem Großteil der Patienten vorliegt [25], zu einer Fehlfaltung sowie einer unvollständigen Reifung des Proteins, sodaß dieses im endoplasmatischen Retikulum akkumuliert wird [22, 25]. Dadurch erfolgt eine Überladung des endoplasmatischen Retikulums, welches zur Funktionsbeeinträchtigung führt, wodurch schließlich das regulatorische Protein DDIT3 vermehrt gebildet wird.

Diese These konnte bereits anhand von Mikroarray-Daten im heterologen Tiermodell bestätigt werden. So wurde bei Mäusen, die das humane F508del-CFTR-Protein exprimieren, eine signifikant höhere Kopienzahl der RNA des DDIT3-Genes gemessen als bei Kontroll-Mäusen, welche das humane Wildtyp-CFTR-Protein exprimieren [113].

Damit ist DDIT3 ein interessanter Transkriptionsfaktor, dessen Beteiligung bei der Expression der DIDS-sensitiven Chloridionenleitfähigkeit aufgrund der gezeigten Zusammenhänge bei den CF-Patienten plausibel erscheint.

RREB-1:

Das ras-Protein (ras: rat sarcoma) zählt zu den kleinen G-Proteinen [114] und ist in der Signaltransduktionskette zwischen stromaufwärts gelegenen Zellrezeptoren mit Tyrosinkinase-Aktivität [115] und stromabwärts gelegenen Proteinkinasen wie die Mitogen-aktivierte Proteinkinase MAPK [116] oder den Streß-aktivierten Proteinkinasen SAPK/JNK [117] geschaltet.

Die ras-Gene sind seit langem bekannte zentrale Proto-Onkogene, welche bei Mutation zu einer dauerhaften Überstimulation des beschriebenen Signalweges führen können. Dadurch erfolgt eine Immortalisierung und maligne Transformation der Zellen [118, 119, 120]. Diese Gene steuern also wichtige Prozesse wie die Mitose und Apoptose. Darüberhinaus wurde beobachtet, daß die Aktivierung des ras-Signalweges zu einer vermehrten Differenzierung der Zellen führen kann [121, 122, 123, 124].

Verschiedenen Arbeitsgruppen konnten zeigen, daß das Protein RREB-1 die Transkriptionsrate von Genen in Abhängigkeit des ras-Signalweges stimuliert. Man vermutet, daß die Aktivierung des Faktors RREB-1 über eine Phosphorylierung durch die Mitogen-aktivierte Proteinkinase MAPK erfolgt [125, 126]. Der Transkriptionsfaktor RREB-1 wurde dabei insbesondere in seiner Funktion bei der Differenzierung von Zellen untersucht [125, 127].

Da dieses regulatorische Protein ubiquitär mit Schwerpunkt im Kolon nachweisbar ist, kann angenommen werden, daß es bei der hCLCA4-kodierten alternativen Chloridionenleitfähigkeit eine relevante Rolle spielt.

Zusammenfassend kann angemerkt werden, daß die Transkriptionsfaktoren RREB-1, DDIT3 und POU6F2 wahrscheinlich die wichtigste Rolle bei der hCLCA4 kodierten alternativen Chloridionenleitfähigkeit spielen. So differenziert RREB-1 als einziger potentiell hybridisierender Faktor den Kernhaplotypen D von der Referenzvariante A. Dieser Transkriptionsfaktor hat sein Expressionsmaximum zudem im Kolon.

Die beiden regulatorischen Proteine POU6F2 und DDIT3 sind ebenfalls die einzigen hybridisierenden Faktoren, welche zwischen den Haplotypen C von A differenzieren.

Im Gegensatz zu POU6F2 wird DDIT3 einerseits im Kolon relativ stark gebildet und andererseits wurde für diesen in der Literatur bereits beschrieben, daß er bei der Mukoviszidose vermehrt exprimiert wird. Damit kann angenommen werden, daß DDIT3 gegenüber POU6F2 wahrscheinlich den Ausschlag gebenden Unterschied von Haplotyp C zu A ausmacht.

Diese beiden Proteine hybridisieren laut Vorhersage auch an die Sequenz von Haplotyp B. Da die Träger von Haplotyp C bereits eine alternative Chloridionenleitfähigkeit ausbilden können, wird angenommen, daß die übrigen an Haplotyp B hybridisierenden Faktoren GATA6,

TEAD2, TEAD3 und TP53 nur fakultativ zur Expression der alternativen Chloridionenleitfähigkeit notwendig sind oder aber diese Expression noch verstärken können.

Abschließend ist zu bemerken, daß alle für die Ausprägung der alternativen Chloridionenleitfähigkeit relevanten Transkriptionsfaktoren mit potentieller Bindung an die Sequenz der Kernhaplotypen B, C oder D bei zellulären Prozessen wie Proliferation, Differenzierung oder Apoptose beteiligt sind.

4.4. Mögliche Rolle des Homologes hCLCA4 als Modulator der Mukoviszidose und dessen molekulare Funktion bei der Kodierung der alternativen Chloridionenleitfähigkeit

Für das Homolog hCLCA4 sind bisher keine proteinchemischen Untersuchungen bezüglich der Funktion, der zellulären Lokalisation oder der Membrantopologie beschrieben worden. Da dieses jedoch zu einer Familie von verwandten Genen zählt, die in verschiedenen Säugetieren nachgewiesen wurde, wird unter Betrachtung dieser Homologen versucht, auf die Funktion von hCLCA4 bei der Expression der alternativen Chloridionenleitfähigkeit bei unseren CF-Patienten zurückzuschließen.

In den folgenden beiden Abschnitten wird das Ergebnis der vorliegenden Arbeit mit den in der Literatur beschriebenen Resultaten in Beziehung gesetzt. So wird in Kapitel 4.4.1 die Aufgabe des Homologes hCLCA4 als genetischer Modulator der Mukoviszidose diskutiert und anschließend in Kapitel 4.4.2 seine mögliche molekulare Funktion bei der Kodierung der alternativen Chloridionenleitfähigkeit erörtert.

4.4.1 Die Rolle des Homologes hCLCA4 als genetischer Modulator der Mukoviszidose

Das intestinale Epithel unterliegt als regeneratives Gewebe einer stetigen Erneuerung. Ausgehend von basalen Stammzellen entstehen durch Mitose und Ausdifferenzierung frische epitheliale adulte Zellen, die nach Durchlaufen des Lebenszyklus durch Apoptose und Abschilferung aus dem Epithelverband herausgelöst werden.

Dieser Zellumsatz ist gesteigert, wenn bestimmte Stressoren die Lebensdauer der einzelnen Zellen verkürzen. Ein solcher Streß kann im Falle der Mukoviszidose durch eine gesteigerte Proteinbiosynthese bedingt sein, wie sie zur Kompensation des fehlgefalteten CFTR-Proteins

auftritt. Auch kann dies durch fehlende Redox-Äquivalente in der Zellumgebung verursacht sein, da daß mutierte CFTR-Protein nicht mehr genügend Glutathion (GSH) transportieren kann [9] und so eine elektrochemische Neutralisierung von Radikalen nicht mehr ausreichend erfolgt.

Genügt dieser Mechanismus zur Adaptierung des Epithels nicht, kann der Zellverband sich im Sinne einer Dedifferenzierung weiter anpassen. So findet sich bei Mukoviszidose-Patienten häufig eine Schrumpf- oder Mikrogallenblase, die durch eine Wandfibrose, schleimgefüllte submuköse Zysten sowie durch eine Metaplasie des Epithels gekennzeichnet ist [36, 128].

Bei elektrophysiologischen Untersuchungen des Gallenblasenepithels von Mukoviszidose-Patienten konnte eine alternative Chloridionenleitfähigkeit beobachtet werden. Diese war im Gegensatz zur CFTR vermittelten Leitfähigkeit nicht über die cAMP-Signalkaskade sondern durch Erhöhung der intrazellulären Kalzium-Konzentration stimulierbar. Interessanterweise war diese alternative Leitfähigkeit im Gegensatz zur CFTR-vermittelten nicht im ausdifferenzierten normalen, sondern nur im dedifferenzierten Gallenblasen-Epithel nachweisbar [36].

Auch wurde bereits für das murine Homolog mCLCA2 beschrieben, daß es vor allem in Geweben exprimiert wird, welches einem hohen Zellumsatz unterliegt. Dies konnte durch Bestimmung der mRNA-Kopienzahl mittels RT-PCR-Techniken festgestellt werden. Weiterhin hat sich gezeigt, daß dieses Homolog in der laktierenden oder sich in der Involution befindlichen Glandula mammaria exprimiert wird, während es in dieser Drüse bei adulten Tieren nicht nachweisbar war [54]. Das bedeutet, das mCLCA2 bei Prozessen des Gewebsumbaues eine wichtige Rolle zu spielen scheint.

Dies steht in direktem Zusammenhang mit den Ergebnissen der vorliegenden Arbeit. So wird zum einen für die CLCA-Proteine berichtet, daß sie über eine Erhöhung der intrazellulären Kalzium-Konzentration zu einer vermehrten Chloridionenleitfähigkeit führen können [40, 45, 46, 129]. Zum anderen sind die Transkriptionsfaktoren mit potentieller Bindung an die Kernhaplotypen B, C oder D vorwiegend an Prozessen, wie Zellproliferation, Differenzierung und Apoptose beteiligt. Das bedeutet, hCLCA4 vermittelt eine alternative Chloridionenleitfähigkeit, die ausgebildet wird, wenn das Gewebe sich rasch erneuert oder sich sogar bereits dedifferenziert hat. Damit wird deutlich, warum die Träger der Haplotypen B, C oder D nicht obligat eine alternative Chloridionenleitfähigkeit ausbilden, sondern nur, wenn die entsprechenden Transkriptionsfaktoren bei einem hohen Zellumsatz im Rahmen einer Gewebs-Dedifferenzierung exprimiert werden.

Diese Entdifferenzierung des Gewebes und die damit verbundene Expression der alternativen Chloridionenleitfähigkeit tritt möglicherweise nur bei Patienten auf, die an einer schweren Form der Mukoviszidose erkrankt sind. Denn es wird beschrieben, das eine DIDS-sensitive

alternative Chloridionenleitfähigkeit vorwiegend bei schwer, nicht jedoch bei leicht erkrankten Mukoviszidose-Patienten beobachtet wird [30].

Mehrere Arbeitsgruppen haben zudem eine alternative Chloridionenleitfähigkeit im Lungengewebe von CF-Patienten beschrieben [34, 35]. Da für hCLCA4 in den bisherigen Untersuchungen [48, 85] jedoch keine Expression in der Lunge nachgewiesen werden konnte, handelt es sich hier wahrscheinlich um Proteine mit einer anderer Identität.

Zudem werden die Transkriptionsfaktoren mit vorhergesagter Bindung an die Haplotypen B, C oder D zum großen Teil schwerpunktmäßig im Kolon exprimiert, sodaß angenommen werden kann, daß es sich bei hCLCA4 wahrscheinlich um einen rein gastrointestinalen Modulator handelt.

4.4.2 Molekulare Funktion von hCLCA4 bei der Expression der alternativen Chloridionenleitfähigkeit

Die CLCA-Proteine konnten in verschiedenen Säugetieren nachgewiesen werden. Diese membranständigen Proteine sind beteiligt an der Vermittlung der Chloridionenleitfähigkeit, der Zell-Zell-Adhäsion und Metastasierung sowie bei der Tumorsuppression [130].

Für hCLCA4 sind bisher keine proteinchemischen Untersuchungen bezüglich der Funktion, der zellulären Lokalisation oder Membrantopologie beschrieben worden. Im folgenden wird daher unter Betrachtung der Homologe dieser Familie versucht, auf die molekulare Funktion von hCLCA4 bei der Kodierung der alternativen Chloridionenleitfähigkeit zurück zuschließen.

Wie aus den Eigenschaften der oben beschriebenen Transkriptionsfaktoren mit potentieller Bindung an die Kernhaplotypen B, C oder D abgeleitet werden kann, wird hCLCA4 vorwiegend bei Prozessen wie Wachstum und Differenzierung exprimiert.

Interessanterweise wurde beobachtet, daß CLCA-Proteine das Zellwachstum stimulieren können. So ist eine Interaktion zwischen hCLCA2 und β_4 -Integrin beschrieben worden [51]. Diese spielt bei der Adhäsion und der Vermittlung von mitogenen Signalen bei der hämatogenen Metastasierung von Karzinomzellen eine Rolle. Anhand von Studien mit dem murinen Homolog mCLCA1 konnte nachgewiesen werden, daß diese Interaktion über eine Aktivierung der FAK (focal adhesion kinase) zur Stimulation der ERK (extracellular regulated kinase) führt und somit das Wachstumssignal ins Zellinnere leitet [52]. Diese Wechselwirkung konnte für verschiedene CLCA-Homologe nachgewiesen werden. Als gemeinsames Bindungsmotiv wurde die Peptidsequenz F(S/N)R(I/L/V)(S/T)S ermittelt [53]. Das hier als Modulator des gastrointestinalen Phänotypen identifizierte Homolog hCLCA4 verfügt jedoch nicht über dieses Sequenzmotiv der β_4 -Integrin-Bindungsdomäne. Das bedeutet, hCLCA4 vermittelt seine Wirkung bei der Kodierung der alternativen Chloridionenleitfähigkeit nicht

über die Stimulation des Zellwachstums, sondern es wird lediglich sekundär im Rahmen dieses Prozesses exprimiert.

Durch einen Vergleich der DNA- und Aminosäuresequenzen konnte mCLCA6 [129] als eines der beiden postulierten murine Homologe für hCLCA4 [131] identifiziert werden. Dieses wird wie hCLCA4 im Darm exprimiert und verfügt ebenfalls über keine Bindungsdomäne für β_4 -Integrin [129]. Anhand von elektrophysiologischen Untersuchungen von transient transfizierten Zellen (HEK 293, human embryonic kidney) konnte gezeigt werden, daß mCLCA6 die Chloridionenleitfähigkeit in Abhängigkeit von Ca^{2+} erhöht [129]. Damit kann angenommen werden, daß auch hCLCA4 bei der Vermittlung einer Kalzium-abhängige Chloridionenleitfähigkeit beteiligt ist.

Protein-Strukturvorhersagen auf der Grundlage von aktuellen Verfahren wie der Sequenzvergleich mittels der SMART-Datenbank haben gezeigt, daß hCLCA4 und andere Homologe dieser Protein-Familie im Gegensatz zu früheren Annahmen lediglich über eine Transmembrandomäne im carboxyterminalen Bereich verfügen [49, 50], über welche sie in der Lipid-Doppelschicht verankert sind. Experimentell konnte dieser Befund für hCLCA2 bereits bestätigt werden [132]. Immunhistologische Untersuchungen haben zudem gezeigt, daß mCLCA6 in der Plasmamembran lokalisiert ist [129]. Damit ist anzunehmen, daß auch hCLCA4 nicht etwa in der Membran von inneren Organellen, sondern in der äußeren Zellmembran über eine Transmembrandomäne verankert ist.

Um die Pore eines Ionenkanales auszubilden, müssen mehrere Transmembrandomänen zu einem großen Verband zusammengeschlossen sein und gemeinsam die Lipid-Doppelschicht durchspannen. Das bedeutet, ein einzelnes CLCA-Molekül ist dafür nicht ausreichend. Für hCLCA4 und die übrigen Homologe wurde jedoch eine zentrale von-Willebrand-Domäne vorhergesagt [49, 50]. Solche Domänen werden häufig bei Proteinen beobachtet, welche Teil eines großen Multiprotein-Komplexes sind. Die Aufgabe dieser Domänen besteht dabei in der Vermittlung von Protein-Protein-Interaktionen

Es kann also angenommen werden, daß auch hCLCA4 sich entweder zu Homo- oder Heteromeren zusammenschließt, um die Kalzium-abhängige Chloridionenleitfähigkeit auszubilden.

So wurde bereits eine physikalische Interaktion von mCLCA1 mit der beta-Untereinheit von Kalzium-abhängigen Kaliumkanälen beschrieben [133]. Dabei konnte nachgewiesen werden, daß die Koexpression von mCLCA1 und der bezeichneten beta-Untereinheit im Vergleich zur alleinigen Expression von mCLCA1 zu einer signifikant höheren Chloridionenleitfähigkeit führt, als die alleinige Expression von mCLCA1 in dem untersuchten heterologen System.

Zusammenfassend kann demnach angemerkt werden, daß die Funktion von hCLCA4 bei der Kodierung der alternativen Chloridionenleitfähigkeit bei unseren Mukoviszidose-Patienten wahrscheinlich nicht in der Vermittlung des Zellwachstums besteht, sondern es kann angenommen werden, daß hCLCA4 ein Ionenkanal assoziiertes Protein darstellt.

*Vergleich der Promoter-Sequenz von hCLCA4 mit derjenigen
von anderen Spezies in der Umgebung der Sequenzvariationen I^{CAP} bis V^{CAP}*

Spezies und Homolog (GeneID)	I ^{CAP}	II ^{CAP}	III ^{CAP}	IV ^{CAP}	V ^{CAP}
Pan troglodytes: CLCA4 (469372)	AAAGA C TTTCATG AAAGACTTTCATG	AATGC G ATATACC AATGCAATACC	GCACA G GCCTG GCACACGCCTG	TCAAC G AGTGG TCAACCAGTGG	ACACA A TGGAC ACACAATGGAC
Macaca mulatta: CLCA4 (712123)	AAAGA C TTTCATG AAAGACTTTCATG	AATGC G ATATACC AATGCAATACC	GCACA G GCCTG GCATGCACCTG	TCAAC G AGTGG TCAACCAGTGG	ACACA A TGGAC ACACAATGGAC
Bos taurus: CLCA4 (507504)	AA-AGA C -TT---CA-T ATGAGACCTTGCTCAGT	AATGC G ATAC--CTCC AGTCGGACACGACTGA	GCACA G -CCTG --ATGAGTCTG	TCAAC G AGTGG T-AAGTACTGA	ACACA A TGGAC ACACAATGGAC
Canis familiaris: CLCA4 (612508)	AAAGA C TTTCATG TTAGTTTGGTTTG	AATGC G ATATACC ATAGTTAT-CC	GCACA G GCCTG CCTCAA---TT	TCAAC G AGTGG T-AAAGGGTGC	ACACA A TGGAC ATAAAAATAAT
Mus musculus: AI747448 (99709)	AAAGA- C TTTCATG AAATAACTAAACA	AATGC G ATATACC --TGTGA---C	GCACA G GCCTG CTAATAATCTA	TCAAC G AGTGG GCA--CAGGGT	ACACA A TGGAC A-AC--TGAAC
Mus musculus: CLCA6 (99663)	AAAGA C TTTCATG AAATTGCITTTTTTA	AATGC G ATATAC--C CAGGAGATAGGGC	GCACA C -GCCTG -CACAGGGCCCC	TCAAC G AGTGG TAAA--AGAAA	ACAC-----A--- A TGGAC ATTCTATATAGTAGGGATGGTC

Tabelle 14: Vergleichende Darstellung der Sequenzen von verschiedenen Spezies im Bereich der Marker I^{CAP} bis V^{CAP}. Die obere Zeile einer jeden Spalte entspricht der humanen Sequenz, die untere repräsentiert die Sequenz der jeweiligen Spezies in diesem Bereich. Es wurden Daten folgender ‚Contigs‘ verwendet [86]: Homo sapiens: NT_032977.8, Pan troglodytes: NW_001228844.1, Macaca mulatta: NW_001108704, Bos taurus: NW_001494779.1, Canis familiaris: NW_876321.4, Mus musculus: NT_039240

4.5 Schlußfolgerung und Ausblick

Ein wichtiges Ergebnis der vorliegenden Arbeit ist die Eingrenzung des Modulators des gastrointestinalen Phänotypes bei unseren Mukoviszidose-Patienten auf den nahen Promoter des Genes *bCLCA4*. Daher wird vermutet, daß die alternative Chloridionenleitfähigkeit durch dieses Homolog der *CLCA*-Genfamilie kodiert wird.

Durch eine Teilresequenzierung dieser Promoter-Region wurden vier Kernhaplotypen ermittelt, welche mit dem gastrointestinalen Phänotypen in Beziehung gesetzt werden können. Es wurde beobachtet, daß nur die Träger Haplotypen B, C oder D eine alternative Chloridionenleitfähigkeit nicht obligat aber fakultativ ausbilden.

Bioinformatische Analysen der Kernhaplotypen geben den Hinweis, daß ein differentes Spektrum an Transkriptionsfaktoren an die haplotypdefinierten Sequenzvariationen binden kann. Die Modulation des gastrointestinalen Phänotyps würde demnach auf einer unterschiedlichen Regulation der hCLCA4 mRNA Expression beruhen.

Um diese Hypothese zu verifizieren, ist es notwendig, zum einen die Hybridisierung der Transkriptionsfaktoren an die DNA-Sequenz und zum anderen deren Einfluß auf die Transkriptionsrate nachzuweisen. Ersteres könnte etwa mittels eines EMSA (electric mobility shift assay) erfolgen und für letzteres können Transfektionsversuche von humanen Zellen mit einem Promoter-Reporter-Gen-Konstrukt vorgenommen werden [134].

Die Träger der oben bezeichneten Kernhaplotypen bilden die DIDS-sensitive Chloridionenleitfähigkeit nicht obligat, sondern nur fakultativ aus. Dies wurde darauf zurückgeführt, daß die zur Expression von hCLCA4 notwendigen Transkriptionsfaktoren nur bei schwer erkrankten CF-Patienten vorliegen.

Tiermodelle der Mukoviszidose-Erkrankung sind für eine Untersuchung dieser Fragestellung leider ungeeignet.

Wie Tabelle 14 zu entnehmen ist, sind die Abschnitte um die Positionen der Sequenzvariationen I^{C4IN1} bis V^{C4IN1} nur bei den Primaten konserviert. Eine Ausnahme von dieser Beobachtung bildet lediglich SNP III^{C4IN1} bei den Makaken. Beim Rind, dem Hund oder der Maus weicht die Sequenz um mehr als eine Base von der Referenzsequenz des Menschen ab.

Dies zeigt, daß die *CLCA*-Genfamilie, wie bereits beschrieben wurde [131], einer schnellen Evolution unterliegt und es sich bei den Sequenzabschnitten, welche die Positionen der Marker I^{C4IN1} bis V^{C4IN1} umgeben, um für Primaten spezifische Bereiche handelt. Das bedeutet, Paarhufer, Raubtiere oder Nager eignen sich als Tiermodell für die Untersuchung der oben genannten Fragestellung nicht. Vielmehr ist es notwendig, Mukoviszidose-Patienten mit

unterschiedlichen Schweregrad der Erkrankung zu rekrutieren, die homozygot für die einzelnen Kernhaplotypen sind und anhand von *ex vivo* Untersuchungen die Expressionsrate von hCLCA4 mit den Haplotypen in Abhängigkeit des Schweregrades in Beziehung zu setzen. Anhand von rektalen Schleimhautbiopsien könnten die Kopienzahl des hCLCA4-Transkriptes mittels quantitativer RT-PCR oder die Protein-Menge mit dem bereits gewonnenen Peptid-Antikörper (siehe Kapitel 3.5) im Immunoblot bestimmt werden.

5 Zusammenfassung

Die Mukoviszidose (syn. Cystische Fibrose, CF), das häufigste schwere Erbleiden europäischer Populationen mit autosomal rezessivem Erbgang, wird durch Mutationen im *CFTR* Gen ausgelöst. Störungen oder Ausfall der *CFTR* Funktion führen zu einer verminderten Chloridpermeabilität in der Apikalmembran des sekretorischen Epithels.

Der klinische Verlauf der Mukoviszidose ist selbst bei Patienten mit gleichem *CFTR*-Mutationsgenotyp stark variabel. Alternative Chloridkanäle, die den Ausfall der *CFTR* Funktion kompensieren, können neben Umgebungseinflüssen und weiteren genetischen Faktoren den Erkrankungsschweregrad modulieren. In Messungen des intestinalen Ionenstroms an Rektumschleimhautbiopsien von CF Zwillingen und Geschwistern ließ sich bei einigen Patienten eine *CFTR*-unabhängige, mittels 4,4'-Diisothiocyanatostilben-2,2'-disulfonsäure (DIDS) hemmbare Chloridionenleitfähigkeit nachweisen.

Ein oder mehrere der vier homologen *CLCA*-Gene (engl: Cl⁻ Channel, Ca²⁺ activated) könnten an der Ausbildung dieser alternativen Chloridkanäle beteiligt sein. Die multifunktionalen *CLCA* Proteine sind neben der Vermittlung einer DIDS-inhibierbaren Kalzium-abhängigen Chloridionenleitfähigkeit auch an der Zelladhäsion und Steuerung des Zell-Zyklus beteiligt.

Nachdem das *CLCA*-Gen-Cluster in Vorarbeiten bereits als Modulator des gastrointestinalen Basisdefektes bei CF identifiziert worden war, sollte in dieser Dissertation ermittelt werden, welches der vier Homologe an der alternativen Chloridionenleitfähigkeit beteiligt ist.

Dazu wurden 22 CF-Patienten ohne nachweisbare Chloridionenleitfähigkeit und 13 CF-Patienten mit alternativer Chloridionenleitfähigkeit in 13 Einzelnukleotid-Polymorphismen (SNPs) des *CLCA*-Gen-Clusters vergleichend genotypisiert.

Die Patienten gehören größtenteils zum Kollektiv der Europäischen CF Zwillings- und Geschwisterstudie und sind überwiegend homozygot für die *CFTR*-Mutation F508del. Die Sequenzen der DNA-Marker wurden einer öffentlich zugänglichen Datenbank (dbSNP) entnommen. Während die Allel-Bestimmung der meisten Marker mittels PCR und Restriktionsverdau erfolgte, wurde ein SNP mittels allelspezifischer PCR typisiert.

Anhand der elterlichen Genotypen wurden die SNPs zu Zwei-Marker-Haplotypen kombiniert und anschließend die Verteilung zwischen den Studiengruppen verglichen. Dabei konnte ein signifikanter Unterschied von drei benachbarten Haplotypen in einem 11,6 kbp großen Bereich beobachtet werden. Dieser überspannt den Promoter und das erste Exon von *bCLCA4*, sodaß angenommen wird, daß dieses Homolog den Modulator des

gastrointestinalen Phänotypes kodiert. Daneben wurde mit dem Vier-Gameten-Test im *CLCA*-Gen-Cluster ein hoch konservierter Bereich nachgewiesen. Dieser ist mit dem 11,6 kbp großen Fragment identisch und weist damit auf die Relevanz dieser Region für die Expression der alternativen Chloridionenleitfähigkeit hin.

Eine vergleichende Sequenzierung (Fa. Qiagen, Hilden) kontrastierender Haplotypen von Patienten aus beiden Studiengruppen in einem 3 kbp großen Abschnitt innerhalb der bezeichneten Region ergab Unterschiede in Form von fünf SNPs. Eine Typisierung am gesamten Studienkollektiv zeigte, daß diese SNPs zu vier Kernhaplotypen A, B, C und D segregierten. Während alle Patienten mit dem Diplotyp AA keine Chloridionenleitfähigkeit zeigten, war bei den Probanden, welche die Kernhaplotypen B, C oder D auf mindestens einem Chromosom tragen, eine alternative Chloridionenleitfähigkeit fakultativ nachweisbar.

Datenbanken wurden auf potentielle Transkriptionsfaktoren durchsucht, die different an die Sequenz der Kernhaplotypen binden und im Kolon exprimiert werden. Es wurden neun Transkriptionsfaktoren identifiziert, die an die Sequenz der Kernhaplotypen B, C oder D, aber nicht an A binden können und gleichzeitig im Kolon in relevanter Kopienzahl vorkommen.

Alle diese Transkriptionsfaktoren werden bei der Entdifferenzierung von Geweben vermehrt exprimiert. In der Literatur wurde eine alternative Chloridionenleitfähigkeit immer dann in der Gallenblase beobachtet, wenn sich das Epithel entdifferenziert hatte. Analog könnten die Transkriptionsfaktoren in der Rektumschleimhaut im Rahmen einer Epithelzell-Dedifferenzierung über eine gesteigerte Expression von hCLCA4 zur Ausprägung der alternativen Chloridionenleitfähigkeit führen. Diese Folgerung, daß Remodelling und Entdifferenzierung die Expression der alternativen hCLCA4-vermittelten Chloridleitfähigkeit triggern, steht in Einklang mit der Beobachtung, daß nur bei Patienten mit einer schweren Verlaufsform der CF eine DIDS-sensitive Chloridionenleitfähigkeit nachgewiesen werden konnte. Da hCLCA4 und die bezeichneten Transkriptionsfaktoren schwerpunktmäßig im Kolon exprimiert werden, handelt es sich bei diesem Gen wahrscheinlich um einen rein gastrointestinalen Modulator.

Die Vorhersage zur Rolle von hCLCA4 und alternativen Chloridkanälen bei CF ließen sich in weiterführenden Arbeiten anhand von *ex vivo* Untersuchungen mit CF-Patienten überprüfen. Beispielsweise kann mit dem bereits gewonnenen Peptid-Antikörper die Expression des Proteins hCLCA4 in Abhängigkeit von Kernhaplotyp und Schweregrad der Erkrankung untersucht werden.

7 Literaturverzeichnis

1. Welsh MJ, Tsui LC, Boat TF, Beaudet AL: Cystic fibrosis. In: Scriver CR, Beaudet AL, Sly WS, Valle D (Hrsg.): The metabolic and molecular bases of inherited disease. 7. Auflage, Band III, McGraw Hill, New York 1995: 3799-3876
2. Davis PB: Cystic fibrosis since 1938. *Am J Respir Crit Care Med.* 2006 Mar 1;173(5):475-82.
3. Knowlton RG, Cohen-Haguenauer O, Van Cong N, Frezal J, Brown VA, Barker D, Braman JC, Schumm JW, Tsui LC, Buchwald M, et al: A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. *Nature.* 1985 Nov 28-Dec 4;318(6044):380-2.
4. Rommens JM, Iannuzzi MC, Kerem B, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, et al: Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science.* 1989 Sep 8;245(4922):1059-65.
5. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, et al: Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.* 1989 Sep 8;245(4922):1066-73. Erratum in: *Science* 1989 Sep 29;245(4925):1437.
6. Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC: Identification of the cystic fibrosis gene: genetic analysis. *Science.* 1989 Sep 8;245(4922):1073-80.
7. Sheppard DN, Welsh MJ: Structure and function of the CFTR chloride channel. *Physiol Rev.* 1999 Jan;79(1 Suppl):S23-45.
8. Bear CE, Li CH, Kartner N, Bridges RJ, Jensen TJ, Ramjeesingh M, Riordan JR: Purification and functional reconstitution of the cystic fibrosis transmembrane conductance regulator (CFTR). *Cell.* 1992 Feb 21;68(4):809-18.
9. Linsdell P, Hanrahan JW: Glutathione permeability of CFTR. *Am J Physiol.* 1998 Jul;275(1 Pt 1):C323-6.
10. Kunzelmann K: CFTR: interacting with everything? *News Physiol Sci.* 2001 Aug;16:167-70.
11. Moyer BD, Denton J, Karlson KH, Reynolds D, Wang S, Mickle JE, Milewski M, Cutting GR, Guggino WB, Li M, Stanton BA: An apical PDZ protein anchors the cystic fibrosis transmembrane conductance regulator to the cytoskeleton. *J Biol Chem.* 1998 Jul 31;273(31):19797-801.

12. Short DB, Trotter KW, Reczek D, Kreda SM, Bretscher A, Boucher RC, Stutts MJ, Milgram SL: An apical PDZ protein anchors the cystic fibrosis transmembrane conductance regulator to the cytoskeleton. *J Biol Chem.* 1998 Jul 31;273(31):19797-801.
13. Vankeerberghen A, Cuppens H, Cassiman JJ: The cystic fibrosis transmembrane conductance regulator: an intriguing protein with pleiotropic functions. *J Cyst Fibros.* 2002 Mar;1(1):13-29.
14. Schwiebert EM, Benos DJ, Egan ME, Stutts MJ, Guggino WB: CFTR is a conductance regulator as well as a chloride channel. *Physiol Rev.* 1999 Jan;79(1 Suppl):S145-66.
15. Kunzelmann, Mall M, Briel M, Hipper A, Nitschke R, Ricken S, Greger R: The cystic fibrosis transmembrane conductance regulator attenuates the endogenous Ca²⁺-activated Cl⁻ Conductance of *Xenopus* oocytes. *Pflügers Arch - Eur J Physiol* (1997)435:178-181
16. Wei L, Vankeerberghen A, Cuppens H, Eggermont J, Cassiman JJ, Droogmans G, Nilius B: Interaction between calcium-activated chloride channels and the cystic fibrosis transmembrane conductance regulator. *Pflügers Arch - Eur J Physiol* (1999) 438:635-641
17. Wei L, Vankeerberghen A, Cuppens H, Cassiman JJ, Droogmans G, Nilius B: The C-terminal part of the R-domain, but not the PDZ binding motif, of the CFTR is involved in the interaction with Ca²⁺-activated Cl⁻channels. *Pflügers Arch - Euro J Physiol* (2001)442:280-285
18. Hartzell C, Putzier I, Arreola J: Calcium-activated chloride channels. *Annu Rev Physiol.* 2005;67:719-58.
19. Tümmler B, Stanke F: Schweißtest und elektrophysiologische Methoden. In: Reinhard D, Götz M, Kraemer R, Schöni MH (Hrsg.): *Cystische Fibrose*. 1. Auflage. Springer, Berlin, Heidelberg 2001: 205 – 212
20. Veeze HJ, Sinaasappel M, Bijman J, Bouquet J, De Jonge HR: Ion Transport Abnormalities in Rectal Suction Biopsies From Children With Cystic Fibrosis. *Gastroenterology* 1991; 101:389-403
21. CF genetic analysis consortium. Cystic Fibrosis Mutation Database. The Cystic Fibrosis Genetic Analysis Consortium 2007;<http://www.genet.sickkids.on.ca>
22. Welsh MJ, Smith AE: Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell.* 1993 Jul 2;73(7):1251-4.
23. Kulczycki LL, Kostuch M, Bellanti JA: A clinical perspective of cystic fibrosis and new genetic findings: relationship of CFTR mutations to genotype-phenotype manifestations. *Am J Med Genet A.* 2003 Jan 30;116(3):262-7.
24. European Working Group on Cystic Fibrosis Genetics: Gradient of distribution in Europe of the major CF mutation and of its associated haplotype. European Working Group on CF Genetics (EWGCFG). *Hum Genet.* 1990 Sep;85(4):436-45.

25. Cheng SH, Gregory RJ, Marshall J, Paul S, Souza DW, White GA, O'Riordan CR, Smith AE: Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell*. 1990 Nov 16;63(4):827-34.
26. McKone EF, Emerson SS, Edwards KL, Aitken ML: Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet*. 2003 May 17;361(9370):1671-6.
27. McKone EF, Goss CH, Aitken ML: CFTR genotype as a predictor of prognosis in cystic fibrosis. *Chest*. 2006 Nov;130(5):1441-7.
28. Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H, Tummler B: Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. *Twin Res*. 2000 Dec;3(4):277-93.
29. Bronsveld I, Mekus F, Bijman J, Ballmann M, Greipel J, Hundrieser J, Halley DJ, Laabs U, Busche R, De Jonge HR, Tummler B, Veeze HJ: Residual chloride secretion in intestinal tissue of deltaF508 homozygous twins and siblings with cystic fibrosis. The European CF Twin and Sibling Study Consortium. *Gastroenterology*. 2000 Jul;119(1):32-40.
30. Bronsveld I, Mekus F, Bijman J, Ballmann M, de Jonge HR, Laabs U, Halley DJ, Ellemunter H, Mastella G, Thomas S, Veeze HJ, Tummler B: Chloride conductance and genetic background modulate the cystic fibrosis phenotype of Delta F508 homozygous twins and siblings. *J Clin Invest*. 2001 Dec;108(11):1705-15.
31. Mekus F, Laabs U, Veeze H, Tummler B: Genes in the vicinity of CFTR modulate the cystic fibrosis phenotype in highly concordant or discordant F508del homozygous sib pairs. *Hum Genet*. 2003 Jan;112(1):1-11.
32. Stanke F, Becker T, Cuppens H, Kumar V, Cassiman JJ, Jansen S, Radojkovic D, Siebert B, Yarden J, Ussery DW, Wienker TF, Tummler B: The TNFalpha receptor TNFRSF1A and genes encoding the amiloride-sensitive sodium channel ENaC as modulators in cystic fibrosis. *Hum Genet*. 2006 Apr;119(3):331-43.
33. Schultz BD, Singh AK, Devor DC, Bridges RJ: Pharmacology of CFTR chloride channel activity. *Physiol Rev*. 1999 Jan;79(1 Suppl):S109-44.
34. Anderson MP, Welsh MJ: Calcium and cAMP activate different chloride channels in the apical membrane of normal and cystic fibrosis epithelia. *Proc Natl Acad Sci U S A*. 1991 Jul 15;88(14):6003-7.
35. Boucher RC, Cheng EH, Paradiso AM, Stutts MJ, Knowles MR, Earp HS: Chloride Secretory Response of Cystic Fibrosis Human Airway Epithelia. *J. Clin. Invest*. 1989 Nov; 84: 1424-1431
36. Dray-Charier N, Paul A, Scoazec JY, Veissiere D, Mergey M, Capeau J, Soubrane O, Housset C: Expression of delta F508 cystic fibrosis transmembrane conductance regulator protein and related chloride transport properties in the gallbladder epithelium from cystic fibrosis patients. *Hepatology*. 1999 Jun;29(6):1624-34

37. Kunzelmann K, Milenkovic VM, Spitzner M, Soria RB, Schreiber R: Calcium-dependent chloride conductance in epithelia: is there a contribution by Bestrophin? *Pflugers Arch.* 2007 Mar 15; [im Druck]
38. Suzuki M: The *Drosophila* tweety family: molecular candidates for large-conductance Ca^{2+} -activated Cl^- channels. *Exp Physiol.* 2006 Jan;91(1):141-7.
39. Barro Soria R, Spitzner M, Schreiber R, Kunzelmann K: Bestrophin 1 enables Ca^{2+} activated Cl^- conductance in epithelia. *J Biol Chem.* 2006 Sep 26; [im Druck]
40. Cunningham SA, Awayda MS, Bubien JK, Ismailov II, Arrate MP, Berdiev BK, Benos DJ, Fuller CM: Cloning of an epithelial chloride channel from bovine trachea. *J Biol Chem.* 1995 Dec 29;270(52):31016-26.
41. Fuller CM, Ismailov II, Keeton DA, Benos DJ: Phosphorylation and activation of a bovine tracheal anion channel by Ca^{2+} /calmodulin-dependent protein kinase II. *J Biol Chem.* 1994 Oct 28;269(43):26642-50.
42. Elble RC, Widom J, Gruber AD, Abdel-Ghany M, Levine R, Goodwin A, Cheng HC, Pauli BU: Cloning and characterization of lung-endothelial cell adhesion molecule-1 suggest it is an endothelial chloride channel. *J Biol Chem.* 1997 Oct 31;272(44):27853-61.
43. Zhu D, Cheng CF, Pauli BU: Blocking of lung endothelial cell adhesion molecule-1 (Lu-ECAM-1) inhibits murine melanoma lung metastasis. *J Clin Invest.* 1992 Jun;89(6):1718-24.
44. Pauli BU, Abdel-Ghany M, Cheng HC, Gruber AD, Archibald HA, Elble RC: Molecular characteristics and functional diversity of CLCA family members. *Clin Exp Pharmacol Physiol.* 2000 Nov;27(11):901-5.
45. Gruber AD, Elble RC, Ji HL, Schreur KD, Fuller CM, Pauli BU: Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca^{2+} -activated Cl^- channel proteins. *Genomics.* 1998 Dec 1;54(2):200-14.
46. Gruber AD, Schreur KD, Ji HL, Fuller CM, Pauli BU: Molecular cloning and transmembrane structure of hCLCA2 from human lung, trachea, and mammary gland. *Am J Physiol.* 1999 Jun;276(6 Pt 1):C1261-70.
47. Gruber AD, Pauli BU: Molecular cloning and biochemical characterization of a truncated, secreted member of the human family of Ca^{2+} -activated Cl^- channels. *Biochim Biophys Acta.* 1999 Mar 19;1444(3):418-23.
48. Agnel M, Vermat T, Culouscou JM: Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea. *FEBS Lett.* 1999 Jul 23;455(3):295-301.
49. Pawlowski K, Lepisto M, Meinander N, Sivars U, Varga M, Wieslander E: Novel conserved hydrolase domain in the CLCA family of alleged calcium-activated chloride channels. *Proteins.* 2006 May 15;63(3):424-39.

50. Whittaker CA, Hynes RO: Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol Biol Cell*. 2002 Oct;13(10):3369-87.
51. Abdel-Ghany M, Cheng HC, Elble RC, Pauli BU: The breast cancer beta 4 integrin and endothelial human CLCA2 mediate lung metastasis. *J Biol Chem*. 2001 Jul 6;276(27):25438-46.
52. Abdel-Ghany M, Cheng HC, Elble RC, Pauli BU: Focal adhesion kinase activated by beta(4) integrin ligation to mCLCA1 mediates early metastatic growth. *J Biol Chem*. 2002 Sep 13;277(37):34391-400.
53. Abdel-Ghany M, Cheng HC, Elble RC, Lin H, DiBiasio J, Pauli BU: The interacting binding domains of the beta(4) integrin and calcium-activated chloride channels (CLCAs) in metastasis. *J Biol Chem*. 2003 Dec 5;278(49):49406-16.
54. Elble RC, Pauli BU: Tumor suppression by a proapoptotic calcium-activated chloride channel in mammary epithelium. *J Biol Chem*. 2001 Nov 2;276(44):40510-7.
55. Nakanishi A, Morita S, Iwashita H, Sagiya Y, Ashida Y, Shirafuji H, Fujisawa Y, Nishimura O, Fujino M: Role of gob-5 in mucus overproduction and airway hyperresponsiveness in asthma. *Proc Natl Acad Sci U S A*. 2001 Apr 24;98(9):5175-80.
56. Hoshino M, Morita S, Iwashita H, Sagiya Y, Nagi T, Nakanishi A, Ashida Y, Nishimura O, Fujisawa Y, Fujino M: Increased expression of the human Ca²⁺-activated Cl⁻ channel 1 (CaCC1) gene in the asthmatic airway. *Am J Respir Crit Care Med*. 2002 Apr 15;165(8):1132-6.
57. Schlotterer C: The evolution of molecular markers--just a matter of fashion? *Nat Rev Genet*. 2004 Jan;5(1):63-9.
58. Ellegren H: Microsatellites: simple sequences with complex evolution. *Nat Rev Genet*. 2004 Jun;5(6):435-45.
59. Hudson RR, Kaplan NL: Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*. 1985 Sep;111(1):147-64.
60. Wang N, Akey JM, Zhang K, Chakraborty R, Jin L: Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am J Hum Genet*. 2002 Nov;71(5):1227-34.
61. Assoziationsstudien und das Kopplungsungleichgewicht. In: Strachan T, Read AP (Hrsg.): *Molekulare Humangenetik*. 3. Auflage. Spektrum, Heidelberg, München 2005: 517-524
62. Ritzka M, Stanke F, Jansen S, Gruber AD, Pusch L, Woelfl S, Veeze HJ, Halley DJ, Tummler B: The CLCA gene locus as a modulator of the gastrointestinal basic defect in cystic fibrosis. *Hum Genet*. 2004 Nov;115(6):483-91.

63. Bridges RJ, Worrell RT, Frizzell RA, Benos DJ: Stilbene disulfonate blockade of colonic secretory Cl⁻ channels in planar lipid bilayers. *Am J Physiol.* 1989 Apr;256(4 Pt 1):C902-12.
64. Schwiebert EM, Flotte T, Cutting GR, Guggino WB: Both CFTR and outwardly rectifying chloride channels contribute to cAMP-stimulated whole cell chloride currents. *Am J Physiol.* 1994 May;266(5 Pt 1):C1464-77.
65. Anderson MP, Sheppard DN, Berger HA, Welsh MJ: Chloride channels in the apical membrane of normal and cystic fibrosis airway and intestinal epithelia. *Am J Physiol.* 1992 Jul;263(1 Pt 1):L1-14.
66. Botstein D, White RL, Skolnick M, Davis RW: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet.* 1980 May;32(3):314-31.
67. Kitts A, Sherry S: The Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation. In: *The NCBI Handbook*. Bethesda (MD): National Center for Biotechnology Information; 2006
68. Vincze T, Posfai J, Roberts RJ: NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res.* 2003 Jul 1;31(13):3688-91; <http://tools.neb.com/NEBcutter2/index.php>
69. Gross-Bellard M, Oudet P, Chambon P: Isolation of high-molecular-weight DNA from mammalian cells. *Eur J Biochem.* 1973 Jul 2;36(1):32-8.
70. Konzentrationsbestimmung von Nucleinsäuren. In: Lottspeich F, Zorbas H. (Hrsg.): *Bioanalytik*. 1. Auflage. Springer, Heidelberg, Berlin 1998: 574-5.
71. Erlich HA, Arnheim N: Genetic analysis using the polymerase chain reaction. *Annu Rev Genet.* 1992;26:479-506.
72. Steve Rozen and Helen J. Skaletsky: Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (Hrsg.): *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ 2000: 365-386; <http://frodo.wi.mit.edu/>
73. Chester N, Marshak DR: Dimethyl sulfoxide-mediated primer T_m reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Anal Biochem.* 1993 Mar;209(2):284-90.
74. Bookstein R, Lai CC, To H, Lee WH: PCR-based detection of a polymorphic BamHI site in intron 1 of the human retinoblastoma (RB) gene. *Nucleic Acids Res.* 1990 Mar 25;18(6):1666.
75. Veeze HJ, Halley DJ, Bijman J, de Jongste JC, de Jonge HR, Sinaasappel M: Determinants of mild clinical symptoms in cystic fibrosis patients. Residual chloride secretion measured in rectal biopsies in relation to the genotype. *J Clin Invest.* 1994 Feb;93(2):461-6.

76. Der Vergleich zweier unabhängiger Stichproben: U-Test nach Wilcoxon, Mann und Whitney. In: Sachs L. (Hrsg.): Angewandte Statistik. 11. Auflage. Springer, Berlin, Heidelberg, New York 2004: 381-389
77. Sham PC, Curtis D: Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet.* 1995 Jan;59(Pt 1):97-105.
78. Ewing B, Hillier L, Wendl MC, Green P: Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 1998 Mar;8(3):175-85.
79. Ewing B, Green P: Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 1998 Mar;8(3):186-94.
80. Hall, TA: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 1999;41:95-98.
81. Ornstein RL, Rein R, Breen DL, Macelroy RD: An optimized potential function for the calculation of nucleic acid interaction energies. I. Base Stacking. *Biopolymers* 1978 17: 2341-2360
82. Bolshoy A, McNamara P, Harrington RE, Trifonov EN: Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc Natl Acad Sci U S A.* 1991 Mar 15;88(6):2312-6.
83. Fye RM, Benham CJ: Exact method for numerically analyzing a model of local denaturation in superhelically stressed DNA. *Phys. Rev. E* 1999 59 (3): 3408 - 3426
84. Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T: MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics.* 2005 Jul 1;21(13):2933-42; <http://www.genomatix.de>
85. Pontius JU, Wagner L, Schuler GD: UniGene: a unified view of the transcriptome. In: *The NCBI Handbook.* Bethesda (MD): National Center for Biotechnology Information; 2003; <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=unigene>
86. Diehn M, Sherlock G, Binkley G, Jin H, Matese JC, Hernandez-Boussard T, Rees CA, Cherry JM, Botstein D, Brown PO, Alizadeh AA: SOURCE: a unified genomic resource of functional annotations, ontologies, and gene expression data. *Nucleic Acids Res.* 2003 Jan 1;31(1):219-23; <http://source.stanford.edu>; <http://smd.stanford.edu/help/SOURCE/normalization.html>
87. <http://www.ncbi.nlm.nih.gov/>
88. Pritchard JK, Rosenberg NA: Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet.* 1999 Jul;65(1):220-8.

89. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, Gabriel SB, Topol EJ, Smoller JW, Pato CN, Pato MT, Petryshen TL, Kolonel LN, Lander ES, Sklar P, Henderson B, Hirschhorn JN, Altshuler D: Assessing the impact of population stratification on genetic association studies. *Nat Genet.* 2004 Apr;36(4):388-93.
90. Connelly JJ, Wang T, Cox JE, Haynes C, Wang L, Shah SH, Crosslin DR, Hale AB, Nelson S, Crossman DC, Granger CB, Haines JL, Jones CJ, Vance JM, Goldschmidt-Clermont PJ, Kraus WE, Hauser ER, Gregory SG: GATA2 is associated with familial early-onset coronary artery disease. *PLoS Genet.* 2006 Aug 25;2(8):e139.
91. Campbell DB, Sutcliffe JS, Ebert PJ, Militerni R, Bravaccio C, Trillo S, Elia M, Schneider C, Melmed R, Sacco R, Persico AM, Levitt P: A genetic variant that disrupts MET transcription is associated with autism. *Proc Natl Acad Sci U S A.* 2006 Nov 7;103(45):16834-9.
92. Kim J, Klooster S, Shapiro DJ: Intrinsically bent DNA in a eukaryotic transcription factor recognition sequence potentiates transcription activation. *J Biol Chem.* 1995 Jan 20;270(3):1282-8.
93. Bash RC, Vargason JM, Cornejo S, Ho PS, Lohr D: Intrinsically bent DNA in the promoter regions of the yeast GAAL1-10 and GAL80 genes. *J Biol Chem.* 2001 Jan 12;276(2):861-6.
94. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD: Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* 2006 Jan 31;34(2):564-74.
95. Mielke C, Christensen MO, Westergaard O, Bode J, Benham CJ, Breindl M: Multiple collagen I gene regulatory elements have sites of stress-induced DNA duplex destabilization and nuclear scaffold/matrix association potential. *J Cell Biochem.* 2002;84(3):484-96.
96. Rowntree RK, Vassaux G, McDowell TL, Howe S, McGuigan A, Phylactides M, Huxley C, Harris A: An element in intron 1 of the CFTR gene augments intestinal expression in vivo. *Hum Mol Genet.* 2001 Jul 1;10(14):1455-64.
97. Pabst O, Zweigerdt R, Arnold HH: Targeted disruption of the homeobox transcription factor Nkx2-3 in mice results in postnatal lethality and abnormal development of small intestine and spleen. *Development.* 1999 May;126(10):2215-25.
98. Pabst O, Schneider A, Brand T, Arnold HH: The mouse Nkx2-3 homeodomain gene is expressed in gut mesenchyme during pre- and postnatal mouse development. *Dev Dyn.* 1997 May;209(1):29-35.
99. Buchberger A, Pabst O, Brand T, Seidl K, Arnold HH: Chick NKx-2.3 represents a novel family member of vertebrate homologues to the Drosophila homeobox gene tinman: differential expression of cNKx-2.3 and cNKx-2.5 during heart and gut development. *Mech Dev.* 1996 May;56(1-2):151-63.

100. Jacquemin P, Hwang JJ, Martial JA, Dolle P, Davidson I: A novel family of developmentally regulated mammalian transcription factors containing the TEA/ATTS DNA binding domain. *J Biol Chem.* 1996 Sep 6;271(36):21775-85.
101. Jacquemin P, Martial JA, Davidson I: Human TEF-5 is preferentially expressed in placenta and binds to multiple functional elements of the human chorionic somatomammotropin-B gene enhancer. *J Biol Chem.* 1997 May 16;272(20):12928-37.
102. Jiang SW, Wu K, Eberhardt NL: Human placental TEF-5 transactivates the human chorionic somatomammotropin gene enhancer. *Mol Endocrinol.* 1999 Jun;13(6):879-89.
103. Lowry JA, Atchley WR: Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain. *J Mol Evol.* 2000 Feb;50(2):103-15.
104. Minegishi N, Ohta J, Suwabe N, Nakauchi H, Ishihara H, Hayashi N, Yamamoto M: Alternative promoters regulate transcription of the mouse GATA-2 gene. *J Biol Chem.* 1998 Feb 6;273(6):3625-34.
105. Gao X, Sedgwick T, Shi YB, Evans T: Distinct functions are implicated for the GATA-4, -5, and -6 transcription factors in the regulation of intestine epithelial cell differentiation. *Mol Cell Biol.* 1998 May;18(5):2901-11.
106. Zhou H, Yoshioka T, Nathans J: Retina-derived POU-domain factor-1: a complex POU-domain gene implicated in the development of retinal ganglion and amacrine cells. *J Neurosci.* 1996 Apr 1;16(7):2261-74.
107. Di Renzo F, Doneda L, Menegola E, Sardella M, De Vecchi G, Collini P, Spreafico F, Fossati-Bellani F, Giavini E, Radice P, Perotti D: The murine Pou6f2 gene is temporally and spatially regulated during kidney embryogenesis and its human homolog is overexpressed in a subset of Wilms tumors. *J Pediatr Hematol Oncol.* 2006 Dec;28(12):791-7.
108. Perotti D, De Vecchi G, Testi MA, Lualdi E, Modena P, Mondini P, Ravagnani F, Collini P, Di Renzo F, Spreafico F, Terenziani M, Sozzi G, Fossati-Bellani F, Radice P: Germline mutations of the POU6F2 gene in Wilms tumors with loss of heterozygosity on chromosome 7p14. *Hum Mutat.* 2004 Nov;24(5):400-7.
109. Ron D, Habener JF: CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. *Genes Dev.* 1992 Mar;6(3):439-53.
110. Ubeda M, Wang XZ, Zinszner H, Wu I, Habener JF, Ron D: Stress-induced binding of the transcriptional factor CHOP to a novel DNA control element. *Mol Cell Biol.* 1996 Apr;16(4):1479-89.
111. Oyadomari S, Mori M: Roles of CHOP/GADD153 in endoplasmic reticulum stress. *Cell Death Differ.* 2004 Apr;11(4):381-9.

112. Wang XZ, Lawson B, Brewer JW, Zinszner H, Sanjay A, Mi LJ, Boorstein R, Kreibich G, Hendershot LM, Ron D: Signals from the stressed endoplasmic reticulum induce C/EBP-homologous protein (CHOP/GADD153). *Mol Cell Biol.* 1996 Aug;16(8):4273-80.
113. Xu Y, Liu C, Clark JC, Whitsett JA: Functional genomic responses to cystic fibrosis transmembrane conductance regulator (CFTR) and CFTR(delta508) in the lung. *J Biol Chem.* 2006 Apr 21;281(16):11279-91.
114. Egan SE, Weinberg RA: The pathway to signal achievement. *Nature.* 1993 Oct 28;365(6449):781-3.
115. Khosravi-Far R, Der CJ: The Ras signal transduction pathway. *Cancer Metastasis Rev.* 1994 Mar;13(1):67-89.
116. Northwood IC, Gonzalez FA, Wartmann M, Raden DL, Davis RJ: Isolation and characterization of two growth factor-stimulated protein kinases that phosphorylate the epidermal growth factor receptor at threonine 669. *J Biol Chem.* 1991 Aug 15;266(23):15266-76.
117. Yan M, Dai T, Deak JC, Kyriakis JM, Zon LI, Woodgett JR, Templeton DJ: Activation of stress-activated protein kinase by MEKK1 phosphorylation of its activator SEK1. *Nature.* 1994 Dec 22-29;372(6508):798-800.
118. Barbacid M: ras genes. *Annu Rev Biochem.* 1987;56:779-827.
119. Malumbres M, Barbacid M: RAS oncogenes: the first 30 years. *Nat Rev Cancer.* 2003 Jun;3(6):459-65.
120. Giehl K: Oncogenic Ras in tumour progression and metastasis. *Biol Chem.* 2005 Mar;386(3):193-205.
121. Benito M, Porras A, Nebreda AR, Santos E: Differentiation of 3T3-L1 fibroblasts to adipocytes induced by transfection of ras oncogenes. *Science.* 1991 Aug 2;253(5019):565-8.
122. Nakagawa T, Mabry M, de Bustros A, Ihle JN, Nelkin BD, Baylin SB: Introduction of v-Ha-ras oncogene induces differentiation of cultured human medullary thyroid carcinoma cells. *Proc Natl Acad Sci U S A.* 1987 Aug;84(16):5923-7.
123. Noda M, Ko M, Ogura A, Liu DG, Amano T, Takano T, Ikawa Y: Sarcoma viruses carrying ras oncogenes induce differentiation-associated properties in a neuronal cell line. *Nature.* 1985 Nov 7-13;318(6041):73-5.
124. Seremetis S, Inghirami G, Ferrero D, Newcomb EW, Knowles DM, Dotto GP, Dalla-Favera R: Transformation and plasmacytoid differentiation of EBV-infected human B lymphoblasts by ras oncogenes. *Science.* 1989 Feb 3;243(4891):660-3.
125. Thiagalingam A, De Bustros A, Borges M, Jasti R, Compton D, Diamond L, Mabry M, Ball DW, Baylin SB, Nelkin BD: RREB-1, a novel zinc finger protein, is involved in the differentiation response to Ras in human medullary thyroid carcinomas. *Mol Cell Biol.* 1996 Oct;16(10):5335-45.

126. Zhang L, Zhao J, Edenberg HJ: A human Raf-responsive zinc-finger protein that binds to divergent sequences. *Nucleic Acids Res.* 1999 Jul 15;27(14):2947-56.
127. Ray SK, Nishitani J, Petry MW, Fessing MY, Leiter AB: Novel transcriptional potentiation of BETA2/NeuroD on the secretin gene promoter by the DNA-binding protein Finb/RREB-1. *Mol Cell Biol.* 2003 Jan;23(1):259-71.
128. Lang T: Erkrankungen der Leber und der Gallenwege. In: Reinhard D, Götz M, Kraemer R, Schöni MH (Hrsg.): *Cystische Fibrose*. 1. Auflage. Springer, Berlin, Heidelberg 2001: 432 – 424
129. Evans SR, Thoreson WB, Beck CL: Molecular and functional analyses of two new calcium-activated chloride channel family members from mouse eye and intestine. *J Biol Chem.* 2004 Oct 1;279(40):41792-800.
130. Pauli BU, Abdel-Ghany M, Cheng HC, Gruber AD, Archibald HA, Elble RC: Molecular characteristics and functional diversity of CLCA family members. *Clin Exp Pharmacol Physiol.* 2000 Nov;27(11):901-5.
131. Ritzka M, Weinel C, Stanke F, Tummler, B: Sequence Comparison of the Whole Murine and Human CLCA Locus Reveals Conserved Synteny Between Both Species. *Genome Lett.* 2003;2: 149-154
132. Elble RC, Walia V, Cheng HC, Connon CJ, Mundhenk L, Gruber AD, Pauli BU: The putative chloride channel hCLCA2 has a single C-terminal transmembrane segment. *J Biol Chem.* 2006 Oct 6;281(40):29448-54. [im Druck].
133. Greenwood IA, Miller LJ, Ohya S, Horowitz B: The large conductance potassium channel beta-subunit can interact with and modulate the functional properties of a calcium-activated chloride channel, CLCA1. *J Biol Chem.* 2002 Jun 21;277(25):22119-22.
134. Knight JC: Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. *Clin Sci (Lond).* 2003 May;104(5):493-501.

Primersequenzen und Restriktionsendonukleasen zur Typisierung der SNPs im CLCA-Gen-Cluster

Marker	rs-Nummer	Primersequenz	Restriktionsenzym
1 ^{C2IN3}	rs3737672	TAGCAAGGCACATTGCTGAG TCATCGAACACACCCCAAC	Aci I CCgC
2 ^{C2IN11}	rs3765986	AGCATTTCAGAGCCTCGTCAT TGCCAAGAGGCAGGTCTACT	BamH I GGATCC
3 ^{C1IN1}	rs2791519	TCCAGACTGTTCTGCTTATTCA CATCACTTGCTAATGCTTTTGG	EcoRII CcWGG
4 ^{C1IN10}	rs2791498	CCACCAGTGTGAATGCTGAT TTCACCTCTGTGGACCTTCA	BsmA I GTCTcN
5 ^{C1EX15}	rs1882753	GGAACCTTATGGGGTGATGA CTTCTCAGAGTTGGCTTCC	BsrB I CcGCTC
6 ^{C1/4}	rs2791473	CCCACCAGTTCCAGATACAAAG AAGAGCACAAAGTGACAGACA	Taq I TcGA
7 ^{C1/4}	rs1321681	TGCTTCTGCTGCTGCTATTT CTTTTTC AATTCTTCTGAGCATA CAAA	Mwo I gC(N) ₇ GC
I ^{C4P}	Erstbeschreibung	GCCACATGGTGAAGAATGAA GCTTTTGGGTTTTTGGACAT	Ahd I GAc(N) ₅ GTC
II ^{C4P}	rs2791476	TCTATACAGCTGACAAAGGA CTTGCAAGAGTAGGGAGATA	EcoRV gATATC
III ^{C4P}	rs11807298	GTTGGTGGATCACGAGTTCA GTTGGTGGATCACGAGTTCA	Nsp I RCATGY
IV ^{C4P}	rs11807638	TGAAGAAGATACTTGACACACCG TTTCTGAAATTGGCTGCTG	TspR I NNcASTGNN
V ^{C4P}	Erstbeschreibung	AAAGGCATAAGAACGACACAA AAAGGCATAAGAACGACACAG AAAGGCATAAGAACGACACAG	allelspezifische PCR
8 ^{C4IN1}	rs1321680	AGAGGCTGCAAAGGATAGGG ACTCAAACATTTATCTTTCCATTG	Hpy188 III TCNNGA
9 ^{C4IN1}	rs974843	AGAGGATGGGTGACAAGTGG CAGGAGGAAGCATAGGGTGA	Mse I TtAA
10 ^{C4IN1}	rs972764	TCACCCCTCTTGTTGTTTCCT CCCAAACTTTAAAAATCCTTATGC	Dde I CTNAG
11 ^{C4IN1}	rs6684219	TTGCATGGTTTTTAGTGATGTT TCATCACCACATGGCACGTA	Rsa I gTAC
12 ^{C4IN6}	rs772602	AAATTGCATGGTGGCATAAA TCACATTGCACCCACTTGT	Xba I TCTaGA
13 ^{C4IN4}	rs1517899	GTCAGGGTACAGGGACAGGA GTGTCAGGCACTTTCAAGCA	BseR I GaGGAG(N) ₁₀

Tabelle 15: Auflistung der Primersequenzen und der Restriktionsendonukleasen zur Typisierung der SNPs. Die zur Matrize nicht-komplementären Basen innerhalb der Primersequenzen sind unterstrichen. Alle Enzyme wurde von der Firma NEB bezogen, lediglich EcoRII war nur bei der Firma Fermentas verfügbar.

Sequenzierprimer

Bezeichnung	Sequenz
TM1	GGGCATCCAGATTGGTAAAAG
IN1F	AAGAAGCCTGCATAGCCAAA
IN2F	AACCCAAAAGCAAATGCAAC
IN3F	CTAAAAGTGGGCTGGGTGTG
IN4F	TGCTCATCAATCAACCAGTG
IN5F	TCCAACAGGAAAATGTGCAA
IN6F	CCAACATTTGAGCCAGGAAT
TM2	CCCGACACAATGCAAAGATA
IN1B	TGGCTGATTATGGGATGTTG
IN2B	TTTCTGAAATTGGCTGCTG
IN3B	CTGAACTCGTGATCCACCAA
IN4B	TCCCATTCGTGGGTTGT
IN5B	CCATTCGTCTATATGCCTGCT
IN6B	GAGCATGGGATGTGTTTTCA

Tabelle 16: Auflistung der Sequenzierprimer. Das PCR-Produkt für die Sequenzierung wurde mit dem Primerpaar TM1 und TM2 erstellt. Die internen Primer für die Erstellung der Reads in vorwärts-Richtung sind mit INxF und in rückwärts-Richtung mit INxB bezeichnet.

Lebenslauf

Name	Ernst Wolfgang Kolbe
Geburtsdatum	8. November 1978
Geburtsort	Göttingen
Staatsangehörigkeit	deutsch

Schulbildung

1986 - 1990	Grundschule Kolenfeld
1990 - 1992	Orientierungsstufe Wunstorf-Süd
1992 - Juni 1999	Hölty-Gymnasium Wunstorf, Abschluß: Abitur

Zivildienst

August 1999 - Juni 2000	Johanniter-Unfallhilfe e.V. Niedersachsen Mitte, Personenbeförderung
-------------------------	---

Studium

Oktober 2000 - November 2007	Humanmedizin an der Medizinischen Hochschule Hannover
September 2002	Ärztliche Vorprüfung
August 2003	Erster Abschnitt der Ärztlichen Prüfung
September 2005	Zweiter Abschnitt der Ärztlichen Prüfung
Oktober 2006 - September 2007	Praktische Jahr
November 2007	Dritter Abschnitt der Ärztlichen Prüfung

Promotion

Oktober 2005 - Juni 2006:

Durchführung der experimentellen Laborarbeiten zum Promotionsthema "Analyse der *CLCA*-Genfamilie als Modulator des Basisdefektes bei Patienten mit Mukoviszidose" unter der Betreuung von Prof. Dr. Dr. B. Tümmler in der Abteilung Pädiatrische Pneumologie und Neonatologie im Rahmen des StrucMed-Programmes der Hannover Biomedical Research School.

Hannover, den 03.04.2008

Ernst Kolbe

Erklärung nach § 2 Abs. 2 Nr. 5 und 6 PromO

Ich erkläre, daß ich die der Medizinischen Hochschule Hannover zur Promotion eingereichte Dissertation mit dem Titel

Analyse der *CLCA*-Genfamilie als Modulator des Basisdefektes bei Patienten mit Mukoviszidose

in der Abteilung Pädiatrische Pneumologie und Neonatologie des Zentrums für Kinder- und Jugendmedizin, Humangenetik und Dermatologie der Medizinischen Hochschule Hannover unter Betreuung von Prof. Dr. Dr. Burkhard Tümmler mit der Unterstützung durch Frau Dr. F. Stanke und der Kobetreuung von Prof. Dr. Hassan Y. Naim (TiHo-Hannover, Institut für Physiologische Chemie) ohne sonstige Hilfe durchgeführt und bei der Abfassung der Dissertation keine anderen als die dort aufgeführten Hilfsmittel benutzt habe.

Ich habe diese Dissertation bisher an keiner in- oder ausländischen Hochschule zur Promotion eingereicht. Weiterhin versichere ich, daß ich den beantragten Titel bisher noch nicht erworben habe.

Ergebnisse der Dissertation wurden in Posterform auf der 29th European CF Conference, Copenhagen, Denmark, 15-18 June 2006 und als Abstract veröffentlicht:

E. W. Kolbe, B. Siebert, U. Laabs, I. Leverkusöhne, A. v. Barneveld, M. Ballmann, S. Jansen, M. Ritzka, B. Tümmler and F. Stanke: Analysis of the *CLCA*-gene cluster as a modulator of cystic fibrosis. J Cyst Fibros. 2006 Jun;5 Suppl 1: S5

Hannover, den 11.06.2007

Ernst Kolbe

Danksagung

Herrn Prof. Dr. Dr. Burkhard Tümmler gilt mein besonderer Dank für die Überlassung des interessanten und vielschichtigen Themas und von ganzen Herzen für sein persönliches Engagement bei der Betreuung in allen Phasen dieser Arbeit. Seine gleich bleibend intensive Unterstützung und seine Ratschläge haben mir sehr geholfen und meine Arbeit stets vorangetrieben.

Frau Dr. Frauke Stanke möchte ich herzlich danken für die hervorragende wissenschaftliche Anleitung, die fachlich kompetente und fundierte Heranführung an die theoretischen Grundlagen der Genetik sowie für ihre immer währende Bereitschaft zu den vielen inspirierenden Diskussionen. Ihre fachkundige Weitsicht war für das Gelingen der Arbeit von entscheidender Bedeutung.

Danken möchte ich Frau Dr. Andrea van Barneveld für die fachlich kompetente Beantwortung meiner Fragen zur Antikörpersynthese und für die Bereitstellung der Transkriptomdaten.

Frau Silke Jansen gilt mein Dank für die Einführung in die experimentelle Arbeit und ihre hilfreiche Unterstützung im Laboralltag. Auch möchte ich Frau Dr. Ina Leverkoehne und Herrn Benny Siebert für die ersten Genotypisierungs-Daten meinen Dank aussprechen.

Herrn Prof. David W. Ussery (The Technical University of Denmark, Center for Biological Sequence Analysis, Department of Biotechnology, Lyngby, Denmark) danke ich für die Berechnungen der verschiedenen strukturellen Parameter der DNA.

Des Weiteren gilt mein Dank den Organisatoren des Programmes der Strukturierten Doktorandenausbildung (StrucMed) der Hannover Biomedical Research School, in welchem Rahmen diese Arbeit angefertigt wurde. Auch danke ich Herrn Prof. Dr. Hassan Y. Naim (TiHo-Hannover, Institut für Physiologische Chemie) für seine hilfreichen Ratschläge und Denkanstöße als Kobetreuer.

Überdies möchte ich mich bei allen Mitgliedern der Klinischen Forschergruppe ‚Molekulare Pathologie der Mukoviszidose‘ für die herzliche Aufnahme, die ausgezeichnete Atmosphäre und die ausgesprochene Hilfsbereitschaft bedanken.

Meinen Eltern, meinen beiden Schwestern Sabine und Magdalena und allen anderen, die mich während dieser Arbeit unterstützt haben, danke ich für ihre Toleranz und Nachsicht.