CYSTIC FIBROSIS AS A GENETICALLY COMPLEX DISEASE

CYSTISCHE FIBROSE: EINE GENETISCH KOMPLEXE ERKRANKUNG

Vom Fachbereich Chemie
der Universität Hannover

zur Erlangung des Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

von

Dipl.-Biochem. Frauke Mekus
geboren am 12.03.1970 in Essen

Referent:	Prof. Dr. Dr. Burkhard Tümmler
Korreferenten:	Prof. Dr. Galla
	Prof. Dr. Niemann
	Dr. Jan Bijman
	Prof. Dr. Jean-Jaques Cassiman
Tag der Promotion:	02.02.2000

Für meine Eltern.

The shape of DNA, it is popularly said [on Discworld], owes its discovery to the chance sight of a spiral staircase when the scientist's mind was just at the right receptive temperature. Had he used the lift, the whole science of genetics might have been a good deal different.*

* *Although, possibly, quicker. And only licensed to carry fourteen people.*

Terry Pratchett "Sourcery. A novel of Discworld."
published by Penguin Books Ltd, Harmondsworth, Middlesex, England (1989)

DANKSAGUNG

Diese Arbeit wurde vom November 1994 bis Juli 1998 in der klinischen Forschergruppe "Molekulare Pathologie der Mukoviszidose" unter der Anleitung von Prof. Burkhard Tümmler durchgeführt.

Mein besonderer Dank gilt Herrn Prof. Tümmler für die Überlassung des vielschichtigen Themas der Dissertation und die hervorragende wissenschaftliche Betreuung der Arbeit.

Für die Übernahme der Korreferate bedanke ich mich bei Herrn Prof. Niemann (MHH, Hannover), Herrn Prof. Cassiman (Katholieke Universiteit, Leuven) und Herrn Dr. Bijman (Erasmus Universität, Rotterdam).

Herrn Prof. Maaß und Herrn Prof. von der Hardt danke ich für ihr stetiges Interesse am Werdegang der Arbeit.

Herzlich möchte ich mich bei den Kolleginnen und Kollegen aus der "forschenden Reisegruppe" bedanken, mit denen ich in Hannover, Rotterdam, Innsbruck, Verona und London an der Erhebung der Primärdaten zum CF Basisdefekt und an der Sammlung von DNA Blutproben beteiligt war, ohne die diese Arbeit nicht möglich gewesen wäre. Im einzelnen danke ich

Frau Ulrike Laabs (MHH, Hannover) für die Isolierung ungezählter DNA Proben und die kompetente Verwaltung von Patientendaten

Frau Inez Bronsveld (Erasmus Universität, Rotterdam) für die kollegiale Zusammenarbeit und die hilfreiche Diskussion der NPD und ICM Daten

Herrn Dr. Henk Veeze (Sophia Children's Hospital, Rotterdam) für die Einführung in die ICM-Meßmethode sowie die Interpretation der Primärdaten

Herrn Dr. Manfred Ballman (MHH, Hannover) für Beistand bei Kurzausflügen auf das klinisch-medizinische Parkett

und

Herrn Dr. Jan Bijman (Erasmus Universität, Rotterdam) für die kompetente wissenschaftliche Beratung bei allen elektro-physiologischen Fragestellungen.

Ein besonderer Dank gilt jenen, die den Aufenthalt an den CF Zentren in Innbruck, Verona und London möglich machten: Herrn Dr. Ellemunter (Innsbruck), Herrn Prof Mastella, Herrn Dr. Pradal und Herrn Dr. Castellani (Verona) sowie Herrn Dr. Thomas (London). Ferner danke ich allen CF Kliniken des "European CF Twin and Sibling Study Consortiums" für das zur Verfügung stellen klinischer Daten.

Herrn Christian Becker danke ich für die Isolierung von DNA-Proben während der Anfangsphase der Studie und die tatkräftige Hilfe bei der Datenverwaltung.

Bei Herrn Dr. Joachim Greipel bedanke ich mich für die engagierte Unterstützung bei der computergestützten Auswertung der ICM-Daten und viele damit verbundenen Diskussionen.

Herrn Hildebrandt danke ich für die Berechnung der FEV1%pred Werte.

Frau Jean Zeysig möchte ich für das Korrekturlesen danken.

Bei Frau Dr. Annette Demmer bedanke ich mich für Hilfe bei großen und kleinen Problemen rund um's Heißlabor.

Allen nicht erwähnten Mitarbeitern der "Klinische Forschergruppe Molekulare Pathologie der Mukoviszidose" und der Abteilung "Paediatrische Pneumologie" danke ich für das angenehme Arbeitsklima.

Allen Freunden und besonders meinen Eltern danke ich für die Unterstützung und die Geduld bei zeitweiliger Abwesenheit der Tochter, der Lebensgefährtin oder der vierten Frau zur Doppelkopfrunde.

Den Patienten und ihren Familienangehörigen danke ich für ihre Mitwirkung an den Forschungsvorhaben. Ihnen gelten meine besten Wünsche.

ABSTRACT

Cystic fibrosis (CF) is the most frequent disease inherited in an autosomal recessive fashion within the Caucasian population. Though CF is considered to be a monogenic disorder transmitted by lesions in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, the disease severity is characterised by a high degree of variability. This thesis reports on two studies addressing the relative impact of the *CFTR* mutation genotype and other inherited factors on the course of CF disease.

Patients participating in the study "*rare genotypes and atypical CF*" were recruited from the Hannover clinic. A total of 19 patients were investigated by intestinal current measurements (ICM), nasal potential difference measurement (NPD) and a pilocarpine iontophoresis sweat test to assess the basic defect, being defined as an impaired CFTR transmitted Cl^- conductance. Mutation analysis of the *CFTR* gene was done by single strand conformation polymorphism analysis, direct sequencing and Southern blotting. The two deletions CFTRdelEx2 and CFTRdelEx2,3 were identified by analysis of genomic DNA. Different residual Cl^- conductances, distinguishable by their sensitivity to DIDS and cAMP, were seen by ICM in patients participating in the study "*rare genotypes and atypical CF*" but pancreatic sufficiency was found to be associated with a residual current indicative of CFTR mediated Cl^- conductance. The diagnosis of CF was confirmed by investigation of the basic defect for three patients presenting CF symptoms but lacking evidence for two *CFTR* disease causing lesions as judged by extensive mutation screening and, in two of these cases, sustained by the family history or genetic analysis within the family.

In the "*European CF twin and sibling study*", disease severity and intrapair discordance for more than 300 patient pairs was quantified by combining nutritional (weight for height %, wfh%) and pulmonary (FEV1%pred percentiles, FEVPerc) status. A significantly higher concordance of monozygous twins in wfh% and FEVPerc compared to dizygous patient pairs indicated the impact of non-*CFTR* genetics on the CF disease phenotype. 38 patient pairs, selected due to an extreme phenotype while being homozygous for the most frequent CF causing allele ΔF508, were analysed at candidate gene loci. The typing of polymorphic markers employing multiwell PCR reactions and continuous direct blotting electrophoresis was optimised. Significant association of an allele at the marker locus D12S889 with a milder clinical phenotype was detected. Hence, SCNN1A, encoding for the α-subunit of the amiloride-sensitive sodium channel (ENaC) and located near D12S889, might modulate CF disease. The analysis of markers within the *CFTR* linkage group pointed to the existence of hitchhiking genes which modulate the course of CF disease and thus might be responsible for a heterozygote advantage among CF carriers. A significant influence on wfh% was detected at MetH and at D7S495 due to an association of alleles at these loci with a milder phenotype. Candidate genes for the modulation of the nutritional status located within the investigated region are the receptor tyrosine kinase MET, the paternally imprinted gene PEG1 and the human homologue of the murine obesity gene LEP. An association with the phenotype "discordance" in FEVPerc, indicative of a modulation mediated by a gene encoded in trans, was seen at J3.11. This effect was attributed to a trans encoded factor acting on a responsive element linked to J3.11, thereby regulating *CFTR* expression. At the loci D7S514 and D7S495, an association of alleles with the phenotype "discordant" was seen in wfh%. As serum levels of leptin, encoded for by the gene LEP located near D7S514 and D7S495, were shown to be in linkage to D2S1788, this marker was analysed among CF twins and sibs. Sharing of alleles at D2S1788 was significantly associated with higher concordance in wfh%. Summarising these findings gained by genetic analysis of ΔF508 homozygous CF twins and siblings, evidence for a role of the ENaC and hitchhiking genes near *CFTR* in modulating CF disease severity has been provided.

KEYWORDS: *Genotype-Phenotype Relation*
Association Study
Modulating Genes

ZUSAMMENFASSUNG

Cystische Fibrose (CF) wird durch Mutationen im "Cystic Fibrosis Transmembrane Conductance Regulator" (*CFTR*) Gen ausgelöst und ist die häufigste autosomal rezessiv vererbte Erkrankung in der kaukasischen Bevölkerung. CF wird als monogene Erkrankung angesehen, obwohl der Schweregrad der CF durch eine hohe Variabilität gekennzeichnet ist. Diese Dissertation beschreibt Ergebnisse zweier Studien, deren Ziel es ist, den relativen Einfluß des *CFTR* Mutationsgenotypes neben weiteren vererbten Faktoren auf den Verlauf der Erkrankung zu erfassen.

19 Patienten der CF Klinik in Hannover wurden im Rahmen der Studie "*Seltene Genotypen und atypische CF*" mithilfe der Messung der intestinalen Leitfähigkeit (ICM), der Nasalpotentialdifferenzmessung (NPD) und dem Pilocarpin-Iontophorese Schweißtest hinsichtlich ihres Basisdefektes untersucht, der sich als Störung der *CFTR* vermittelten Chloridleitfähigkeit manifestiert. Mutationsanalyse des *CFTR* Gens wurde mithilfe der Analyse von Einzelstrangkonformationspolymorphismen, Direktsequenzierung und Southern-Blot Verfahren durchgeführt. Die Deletionen CFTRDelEx2 und CFTRDelEx2,3 wurden auf genomischer Ebene nachgewiesen. Verschiedene Restleitfähigkeiten, die aufgrund ihrer Sensitivität gegenüber DIDS und cAMP unterschieden werden konnten, wurden mit ICM im Rahmen der Studie "*Seltene Genotypen und atypische CF*" beobachtet. Pankreassuffizienz war dabei mit CFTR vermittelter Restleitfähigkeit assoziiert. Durch die Analyse des Basisdefektes wurde die Diagnose der CF bei drei Patienten bestätigt, für die eine Anwesenheit von krankheitsauslösenden Läsionen auf beiden *CFTR* Allelen nach ausführlicher Mutationsanalyse unwahrscheinlich schien. Für zwei dieser Patienten wurde der genetische Befund durch klinische oder genetische Analyse der Familie unterstützt.

Im Rahmen der "*Europäischen CF Zwillings-und Geschwisterstudie*" wurde für mehr als 300 Patientenpaare der Schweregrad der Erkankung und die Intrapaardiskordanz quantitativ durch einen aus dem Ernährungsstatus (wfh%) und dem pulmonalen Status (FEVPerc) zusammengesetzten Parameter beschrieben. Monozygote Zwillinge waren signifikant konkordanter in wfh% und FEVPerc als dizygote Patientenpaare, so daß der Einfluß des genetischen Hintergrundes auf den Schweregrad der Erkrankung nachgewiesen werden konnte. Aufgrund eines extremen klinischen Phänotyps wurden 38 Patientenpaare, die für das häufigste CF verursachende Allel ΔF508 homozygot waren, zur Analyse von Kandidatengenorten ausgewählt. Die Typisierung polymorpher Marker wurde mithilfe von Multiwell-PCR Reaktionen und kontinuierlicher "Direct Blotting" Elektrophorese optimiert. Ein Allel des Marker Locus D12S889 war mit einem milden klinischen Verlauf signifikant assoziiert. SCNN1A, das Gen welches für die α–Untereinheit des Amilorid-sensitiven Natriumkanals codiert, liegt in der Nähe von D12S889 und könnte daher als Modulator der Erkrankung CF eine Rolle spielen. Eine Analyse von Markern innerhalb der *CFTR* Kopplungsgruppe ergab Hinweise auf die Existenz benachbarter Gene, die den Verlauf der CF beeinflussen und daher auch für den Heterozygotenvorteil von CF Genträgern verantwortlich sein können. Eine Assoziation von Allelen der Loci MetH und D7S495 mit einem milden Krankheitsbild konnte für wfh% beschrieben werden. Kandidatengene dieser Region mit einer möglichen Funktion bei der Modulation des Ernährungsstatus sind MET, PEG1 und LEP. Mit dem Phänotyp "diskordant" konnte eine Assoziation eines J3.11 Allels in FEVPerc gezeigt werden. Dieser Effekt stellt einen Hinweis auf den Einfluß eines in trans kodierten Gens dar und wurde auf einen regulatorischen Faktor zurückgeführt, der durch ein an J3.11 gekoppeltes Element die Expression des *CFTR* Gens beeinflußt. Eine Assoziation von Allelen der Loci D7S514 und D7S495 mit dem Phänotyp "diskordant" wurde in wfh% beobachtet. Da der Serumspiegel von Leptin – einem durch das zwischen D7S514 und D7S495 liegende Gen LEP kodierten Peptid – mit dem Marker D2S1788 in genetischer Kopplung steht, wurde D2S1788 bei den ausgewählten Patientenpaaren typisiert. Ein signifikanter Zusammenhang zwischen der Zahl gemeinsamer D2S1788 Allele und der Konkordanz in wfh% wurde beobachtet. Zusammenfassend läßt sich aus den Ergebnissen der Typisierung von ΔF508 homozygoten Zwillingen und Geschwistern ein modulierender Einfluß des ENaC sowie mehrerer Gene aus der *CFTR* Kopplungsgruppe auf den Schweregrad der Erkrankung ableiten.

SCHLAGWÖRTER: *Genotyp-Phänotyp-Beziehungen*
Assoziationsstudie
Modulierende Gene

TABLE OF CONTENTS (OVERVIEW)

TABLE OF CONTENTS (DETAILED)

1 INTRODUCTION

Cystic fibrosis (CF) is the most common autosomal recessive disease within the Caucasian population, exhibiting an incidence of 1 in 2500 births (Welsh *et al.* 1995, Davies *et al.* 1996). The symptoms of the disorder are caused by an impaired function of exocrine glands in many organs, predominantly within the gastrointestinal and respiratory tracts. Major manifestations of the disease are a consequence of chronic obstruction: bronchitis and bronchiectasis (obstruction of the airways) combined with chronic infection, pancreatic insufficiency (obstruction of the pancreatic duct), meconium ileus (obstruction of the intestine), biliary cirrhosis (obstruction of the biliary duct) and reduced fertility or infertility in both genders (obstruction within the reproductive organs) (Welsh *et al.* 1995, Davies *et al.* 1996).

The diversity of symptoms associated with CF and the variety of disease conditions which exhibit comparable features provide a challenge to diagnosis, especially when mild symptoms are presented by the patient. Since the 1950s, an elevated concentration of Cl^- in the sweat is used to discriminate CF from other disease conditions (Di Sant'Agnese *et al.* 1953, Gibson and Cooke 1958). The measurement of the ionic composition of sweat gland secretion is directly linked to the basic defect in cystic fibrosis, which was recognised as impaired Cl^- permeability in sweat glands (Schulz 1969). Other diagnostic methods relying on the basic defect have been developed and employed clinically for respiratory tissue (potential difference measurements on the epithelium of the lower nasal turbinate, NPD; Knowles *et al.* 1981) and intestinal tissue (short circuit current measurement of rectal biopsies, ICM; Veeze *et al.* 1991).

1.1 Terminology of human genetics: an application to cystic fibrosis

The field of human genetics examines the relationship between inherited factors (comprising the *genotype*) and the characters or traits of an individual (defining the *phenotype*; Johannsen 1909). In the case of cystic fibrosis, the gene which causes the disease when both chromosomal copies are impaired has been identified in 1989 and termed the Cystic Fibrosis Transmembrane conductance Regulator (*CFTR)* gene (Rommens *et al.* 1989).

1.1.1 Chromosomes and genes

The human genome is organised on pairs of homologous chromosomes, each of them harbouring numerous genes. In the *Guidelines for Human Gene Nomenclature* (1997), a gene is defined as "a DNA segment that contributes to phenotype/function". The localisation of a given gene can be expressed in cytogenetic terms which are based on a band pattern generated by staining of the chromosomal material that is specific for each of the different chromosomes. For instance, the *CFTR* gene is localised at 7q31, an abbreviation for the gene's position at band 31 on chromosome 7 with q referring to the longer of both chromosomal arms that are separated by the chromosome's centromer.

1.1.2 The unidirectional flow of genetic information

The carrier of genetic information is desoxyribunucleic acid (DNA), a macromolecule consisting of a linear sequence of nucleotides. The information content of the genome is encoded by the sequence of the four different nucleotides A, G, C and T (an abbreviation of the monomer's bases adenine, guanine, cytosine and thymidine) of the DNA molecule. Specific base pairing of A to T and G to C is not only a key element for the maintenance of genetic information through replication of the DNA molecule, but also for the expression of the genetic information in the form of a linear sequence of amino acids through transcription and translation.

Transcription generates mRNA (messenger ribonucleic acid), formed by aligning complementary bases to the DNA strand, thereby ensuring that the nucleotide sequence between DNA and mRNA contains the same information. During translation of the nucleotide sequence to an amino acid sequence, the successive bases of the mRNA are interpreted according to the rules of the genetic code: a sequence of three contiguous bases, the codon, is translated to one out of a set of 20 amino acids. Each of the amino acids is linked to a specific tRNA (transfer ribonucleic acid) molecule and specific base pairing between the three nucleotides of the mRNA codon and three nucleotides of the tRNA anticodon again ensures that the gene's information is preserved in the gene-product.

The *CFTR* gene sequence (Riordan *et al.* 1989) encompasses 240 000 bases (240 kb) containing 27 sections of coding sequences in sizes ranging from 38 to 724 bases (so-called *exons*) separated by much larger stretches of non-coding sequences (so-called *introns*). Both, exons and introns, are transcribed into primary mRNA. However, while the nucleotide sequence of the *CFTR* exons determines the amino acid sequence of the CFTR protein, introns are not translated but contain signals that are vital to the expression of the *CFTR* gene, such as the splice site consensus sequences at the exon/intron boundaries. In the splicing process, introns are excised from the primary mRNA and exons are joined to form mRNA that consists of an uninterrupted coding sequence. Translation of the transcribed exon sequences results in the chain of 1480 amino acids of the CFTR protein.

1.1.3 Nomenclature of DNA sequence alterations

The sequence of the four different nucleotides A, G, C and T of the DNA molecule defines the genetic information. Differences between two DNA sequences – so-called *loci* (Morgan *et al.* 1915) – can be described as either a deletion, an insertion or an exchange of nucleotides. A change of the DNA gives rise to alternative sequences at a locus which are called *alleles* (Bateson & Sounders 1902, Johannsen 1909). Rules for naming different alleles were published by a Nomenclature Working Group in *Human Mutation* 11: 1-3 (1998).

By comparing the *CFTR* gene sequences of CF patients and non-CF controls, numerous DNA sequence variations have been described (available at: http://www.genet.sickkids.on.ca/cftr/). They were classified as *neutral polymorphisms* if no influence on the CFTR protein was expected as a consequence of the DNA sequence alteration or as *mutations* whenever an impairment of CFTR function seemed likely.

Even though the terms "sequence alteration", "sequence variation" and "mutation" are all valid to designate a change in the nucleic acid sequence, most human geneticists aim at a distinction between sequence alterations that cause a change in phenotype – i.e. qualify as a cystic fibrosis causing lesion of the *CFTR* gene – and sequence alterations which do not influence the phenotype. Most frequently, this is based on comparisons of the amino acid sequences that will result from the two alleles. In other cases, the decision of whether a sequence alteration is called a neutral polymorphism instead of a mutation is based on the observation that both alleles occur on CF-chromosomes as well as on non-CF chromosomes. Analogously, sequence alterations have been classified as mutations if they were observed on a CF-chromosome, but not on non-CF chromosomes.

1.1.3.1 Sequence alterations of the *CFTR* gene

The DNA sequence alterations are named according to their position with respect to the *CFTR* coding sequence: nucleotide residues from the *CFTR* coding sequence and amino acids from the CFTR protein are enumerated consecutively. A change of a nucleotide of the coding sequence is then described as, for instance, 2694 T/G (Zielenski *et al.* 1991) to denote that the residue 2694 of the *CFTR* coding sequence was found to be T or G. Since the genetic code is redundant, not all nucleotide exchanges lead to a change of the amino acid sequence of the CFTR protein. In the case of the *neutral polymorphism* 2694 T/G, the resulting CFTR protein is not altered by the exchange: both resulting codons – ACT of the allele 2694T as well as ACG of the allele 2694G – encode for the same amino acid threonine (T) at position 854 within the amino acid chain of the CFTR protein.

Other nucleotide alterations within the CFTR coding sequence result in an amino acid exchange, such as the mutation E92K (Nunes *et al.* 1993) which refers to an exchange at position 92 of the amino acid sequence of glutamic acid (E) to lysine (K). E92K is the consequence of the nucleotide exchange from G to A at position 406 of the DNA coding sequence. Hence, the codon 92 is changed from GAA, translated to glutamic acid from allele 406G, to AAA translated to lysine from allele 406A. All sequence alterations that lead to an amino acid exchange are termed *missense mutations*. If a nucleotide exchange leads to an exchange of an amino acid to a termination signal for the translation machinery (so-called stop codon), the mutation is called a *nonsense mutation.* An example is the *CFTR* mutation G542X (Kerem *et al.* 1990) where a nucleotide exchange leads to a stop codon (X) instead of the amino acid glycine (G).

The deletion of nucleotides from the *CFTR* gene sequence are described as frameshift mutations or *out-of-frame deletions* when the reading frame of the *CFTR* coding sequence is disrupted due to a deletion of one or two nucleotides from a codon due to a loss of $3n + 1$ or $3n + 2$ ($n \in \mathbb{N}$) nucleotides, respectively. The most frequent CFTR mutation in central European populations is an *in-frame-deletion* of three nucleotides resulting in a CFTR protein that lacks the residue phenylalanine (F) at position 508 within the amino acid sequence and is abbreviated ΔF508.

Unless the full length genomic sequence is known, the nucleotides of the introns are not enumerated . Consequently, nucleotide exchanges within the non-coding sequences are described in relation to the nearest nucleotide of the coding sequence, e.g. the nucleotide exchange 1898+3A⇨G (Ferrari *et al.* p.c.) which implies that the A following as a third nucleotide in wild type *CFTR* on the nucleotide 1898 – the last residue of exon 12 of the *CFTR* gene – was found to be exchanged to a G in the patient's *CFTR* gene. Nucleotide exchanges near the exon-intron junctions of the *CFTR* gene may affect the splicing process by destroying the splice site consensus sequence and thus are called *splicing mutations*. The first two nucleotides and the last two nucleotides of each intron are highly conserved and nucleotide exchanges at these positions within the *CFTR* gene are considered to be a disease-causing lesion. However, even at a position +5 the splicing mechanism can be affected, as has been shown for the mutation 2789+5G⇨A (Highsmith *et al.* 1997). Nucleotide exchanges at less conserved positions than +1, +2 or –1, – 2 are generally assumed to be associated with the generation of wild-type mRNA and alternatively spliced transcript. In contrast to the alteration of a present splice site as in the case of 2789+5G⇨A, the splicing mutation 3849+10kbC⇨T generates a novel splice site consensus sequence in intron 19 of the CFTR gene. Processing of the altered transcript was shown to yield a truncated CFTR protein as the newly inserted exon bears a stop codon (Highsmith *et al.* 1994).

1.1.3.2 Dimorphic and polymorphic markers

For practical reasons, loci at which nucleotide sequence alterations occur that can easily be monitored by standard procedures are called a "marker". Two types of markers are discriminated by the number of alleles observed: single nucleotide polymorphisms (SNP) refer to an exchange of single nucleotides. SNP loci are characterised by the occurrence of two alleles within the population and thus are termed *dimorphic markers*. Insertions and deletions of more than one nucleotide are observed at loci where short DNA sequence motifs are repeated and consequently, the number of repeats varies between individuals. Most of these short tandem repeat polymorphisms – so-called microsatellites – are characterised by the occurrence of more than two alleles within the population and thus are termed *polymorphic markers*.

Most commonly used names of genetic markers contain little information with respect to the nucleotide exchange involved. For instance, the intron polymorphism 4006-200A/G within the *CFTR* intron 20 is known among researchers within the CF field as TUB20. Other dimorphic markers frequently used in CF research are MetH, XV-2c, KM.19 and J3.11 which were named arbitrarily by

the researchers and therefore, these abbreviations are non-conclusive for most other scientists. Restriction enzymes that cleave DNA specifically near short sequence motifs such as TCGA (recognised by *Taq*I), CTGCAG (recognised by *Pst*I) or CCGG (recognised by *Msp*I or *Hpa*II) are employed in the detection of these SNPs as the nucleotide exchange of the SNP alters a recognition sequence for a restriction enzyme. The alleles are differentiated by the observations "sequence is not cleaved" (defined as allele 1) or "sequence is cleaved" (defined as allele 2). SNPs that can be detected with restriction enzymes are termed "restriction fragment length polymorphisms" (RFLPs). The sequence that differs between the alleles 1 and 2 at a dimorphic marker locus can be inferred if the restriction enzyme used for detection of the dimorphic marker is named, e.g. MetH/*Msp*I, XV-2c/*Taq*I, KM.19/*Pst*I and J3.11/*Hpa*II.

The genome database, a reference for information on the human genome accessible at http://www.gdb.org, applies an unequivocal nomenclature for DNA loci: the letter D abbreviates the source (DNA), followed by a number to denote the chromosomal assignment and a symbol to indicate the complexity of the DNA segment, whereby S denotes a single copy segment. Finally, a sequential number is assigned "to give uniqueness to the above characters" (*Guidelines for Human Gene Nomenclature* 1997). For example, three repeat polymorphisms investigated within this thesis are D7S495, D7S514 and D7S525. While the localisation of these markers on chromosome 7 can be inferred based on their name in the D-nomenclature system, neither their subchromosomal localisation nor the nucleotide sequence alterations that confer polymorphic character to these sequences can be perceived. These three repeat polymorphisms are located near the *CFTR* gene in the order 7cen – D7S525 – *CFTR* – D7S514 – D7S495 – 7qtel, whereby D7S525, D7S514 and D7S495 have been assigned to the cytogenetic bands 7q22-7q31, 7q31-7q32 and 7q32-7q33, respectively (Bouffard *et al.* 1997).

1.1.4 Genetic linkage

While alterations of the DNA sequence that are classified as a mutation infer a change of the gene product's functionality, neutral polymorphisms do not influence a gene product by definition. Nevertheless, the study of polymorphic sequences has several applications in human genetics as it provides information on the genomic region defined by the location of the polymorphism on the physical map.

Both sets of homologous chromosomes are passed from cell to cell at each mitosis (division of somatic cells). A reduction to one set of homologous chromosomes takes place during the meiosis which forms the gametes. Prior to the separation of two chromosomes of a homologous pair during meiosis, sections of chromosomes are exchanged between the two homologous chromosomes in a process called "crossing over" (Morgan & Cattrell 1912) which leads to *recombination* (Bridges & Morgan 1923). Due to the recombination events, the genome of gametes that is passed to the next generation contains combinations of chromosomal sections which do not exist in the parents' somatic cells.

The analysis of recombination events within families facilitates the construction of a *genetic map* (Haldane 1919). Most polymorphisms were identified without knowledge about their chromosomal, much less subchromosomal localisation. Recombination events between two loci were monitored by analysing the inheritance of alleles of genetic markers at these two loci within a family. If the alleles at both loci were not transmitted in parallel to the next generation, recombination has occurred. Generally, recombination is expected to occur more frequently between loci that are located in a larger physical distance. Thus, the observation of recombination events between two loci in many families can be taken as an indication of a larger physical distance between them.

The term "*genetic linkage*" refers to loci which are not separated by a recombination event. Alleles at two loci that are inherited in parallel are called "*associated*". If two alleles are associated more frequently than expected due to the allele frequencies, they define a chromosomal region that is in "*linkage disequilibrium*" (Kimura 1956). The combination of alleles at two or more linked loci is called "*haplotype*" (Cepellini *et al.* 1967).

1.1.4.1 Mapping of the *CFTR* gene: linkage and association

Cystic fibrosis is often regarded as a model example among human geneticists of what is called "reverse genetics". The *CFTR* gene, causing CF when present in a defective state on both chromosomal copies, was identified by studying the transmission of chromosomal regions within large family pedigrees wherein one or more of the members were affected by the disease. Monitoring of the inheritance of chromosomal segments within these families was carried out by the analysis of dimorphic markers.

Due to the fact that alleles at certain dimorphic loci such as MetH and J3.11 were transmitted parallel to the disease in these families (Beaudet *et al.* 1986), researchers were able to conclude that the disease-causing gene was localised near these markers: the inheritance of alleles at a marker locus in parallel with the disease implies that no recombination has occurred between marker locus and disease locus. Consequently, marker loci MetH and J3.11 and the CF disease locus were assumed to be located near each other on the physical map, as the probability of observing a recombination event between two loci increases with the physical distance between them. The identified chromosomal region was further analysed by physical cloning and examination of the DNA sequences for candidate genes (Rommens *et al.* 1989). The identification of the *CFTR* gene among the candidates was due to its predicted biochemical properties based on the deduced amino acid sequence (Riordan *et al* 1989). Analysis of the *CFTR* coding sequence in CF families provided another argument for the role of the *CFTR* gene in CF disease: the three base pair deletion ΔF508 was found to be associated exclusively with CF disease among CF families (Kerem *et al.* 1989a).

1.1.4.2 An extragenic haplotype common to most CF chromosomes

In contrast to the situation 15 to 20 years ago, the physical map of the *CFTR* gene region is now known in detail. At four loci near *CFTR*, the dimorphic markers MetH, XV-2c, KM.19 and J3.11 have been studied within many populations (Estivill *et al.* 1987, Schmidtke *et al.* 1987). At these three markers, located on the genetic map in the sequence 7cen–MetH–XV-2c–KM.19–*CFTR*–J3.11, an association with CF of one out of the two possible alleles 1 or 2 has been observed. For instance, at MetH 54% of normal chromosomes, but 70% of CF chromosomes, carried the allele 1. At J3.11, 43% of normal chromosomes but 60% of CF chromosomes carried the allele 1. The allelic association is even more pronounced at the loci XV-2c and KM.19 which are closer to *CFTR* on the physical map: 89% of CF chromosomes, but only 45% of normal chromosomes, carried the allele 1 at XV-2c and 88% of CF chromosomes, but only 27% of normal chromosomes, carried the allele 1 at KM.19. The resulting MetH – XV-2c – KM.19 – J3.11 haplotype 1-1-2-1 is frequent on CF-chromosomes but rather rare on non-CF chromosomes. The allele sequence 1-1-2-1 at these loci is known as the "CF haplotype". Thus, CF chromosomes bearing the same haplotype can be considered to share a common chromosomal background.

1.2 The CFTR protein and the basic defect in CF

The CFTR gene codes for a protein of 1480 amino acids (Riordan *et al.* 1989) which was, due to sequence similarities and structure predictions, recognised as a member of the ATP binding cassette (ABC) transporter protein family. Like other ABC transporters, CFTR is composed of two cytosolic nucleotide binding domains and two membrane spanning domains containing 6 transmembrane helices. Unique to CFTR is the regulatory domain (R-domain). In addition to being able to transport a substrate across a membrane, ABC transporters are known to impose regulation on other membrane channels (reviewed by Higgins 1995). With respect to the modulation of the activity of other ion channels by CFTR, articles on interaction with the outwardly rectifying Cl^- channel ORCC (Egan *et al.* 1992, Gabriel *et al.* 1993), the amiloride-sensitive epithelial Na^+ channel ENaC (Ismailov *et al.* 1996, Stutts *et al.* 1997) and the K^+ channel have been published (Loussouarn *et al.* 1996, McNicholas *et al.* 1997).

The CFTR protein was identified as a cAMP regulated Cl^- channel expressed in the apical membrane of epithelial cells (Denning *et al.* 1992a). The secretory properties transmitted by CFTR are regulated by the β_2 adrenergic pathway in exocrine tissues. The molecular pathology, connecting the basic defect to the symptoms of the disease, is described comprehensively for the sweat gland (Quinton 1990): the impermeability of the apical membrane of sweat duct cells results in a decreased reabsorption of Cl^- and, parallel to this, a decreased reabsorption of the counterion Na^+. Reduction of NaCl absorption thus leads to elevated NaCl concentrations in sampled sweat.

The most frequent disease-causing lesion within the *CFTR* gene is the in-frame deletion ΔF508 (European Working Group on Cystic Fibrosis Genetics 1990). The molecular phenotype of ΔF508 CFTR protein was assessed in heterologous expression systems (Cheng *et al.* 1990) and in ΔF508 transgenic mice (French *et al.* 1996) as a maturation defect. The absence of apical CFTR was also demonstrated for patient tissue in sweat glands (Kartner *et al.* 1992), but in nasal polyps no difference between ΔF508 homozygous patients and non-CF controls was observed (Dupuit *et al.* 1995). Furthermore, missing or mislocalised CFTR was detected in non-CF lung biopsies (Brezillon *et al.* 1997) and remodelled non-CF nasal epithelium (Brezillon *et al.* 1995), depending on the dedifferentiation status of the tissue. The channel properties of ΔF508 CFTR were distinguishable from wild type channels by a decreased open probability (Haws *et al.* 1996).

For organs other than the sweat gland, the link between the basic defect and the disease symptoms is still not fully resolved. Basically, epithelial cells expressing defective CFTR have been characterised by failure to respond to stimulation of secretion. The interdependence of ion and water transport is thought to result in an altered composition of the fluid secreted by the epithelia of the gastrointestinal, respiratory and reproductive tracts, which in turn leads to obstruction of the excretory ducts of these organs. However, the vulnerability of CF lungs to opportunistic pathogens, determining the life span of CF patients, cannot be explained solely on the basis of altered fluid viscosity. Hypotheses aiming at an explanation of the impaired protection of CF airways were focussed on the unspecific host defence: it has been suggested that a higher NaCl concentration in the airway fluid, a condition presumed to exist in CF, impairs the antibacterial effect mediated by epithelial beta-defensin-1 (Bensch *et al.* 1995, Goldmann *et al.* 1997).

Compared to the efforts spent on analysing the Cl^- transport properties of the protein, few research groups focussed on functions of CFTR other than involvement in ion transport. CFTR was reported to play a role in regulation of plasma membrane recycling (Bradbury *et al.* 1992), water transport (Hasagawa *et al.* 1992), efflux of neutral amino acids (Rotoli *et al.* 1994), membrane endocytosis (Spilmont *et al.* 1996), transport of glutathione adducts of pro-inflammatory agents (Lallemand *et al.* 1997), HCO_3^- conductance (Illek *et al.* 1997), apoptosis (Maiuri *et al.* 1997) and glutathione transport (Lindsdell *et al.* 1998).

In summary, the extent to which the disease symptoms can be derived from the basic defect reflects the complexity of the tissue and the complexity of Cl^- conductances displayed by the individual cell: for the absorptive duct of the sweat gland, expressing exclusively the CFTR transmitted Cl^- conductance, the observed deviations between CF and non-CF individuals can be explained solely on the basis of defective CFTR mediated Cl^- conductance. For organs like the lung, displaying several cell types and expressing Cl^- conducting properties other than those mediated by CFTR, the link between the pathophysiology of CF disease and the basic defect is not fully understood.

1.3 The role of the *CFTR* gene in disease

While the mode of inheritance for CF was in 1910 first assumed to be autosomal recessive based on the analysis of a consanguineous family (Garrod *et al.* 1912), this mode of inheritance became a generally accepted fact in the 1950s (Lowe 1949, Carter 1952). The disease is considered to be monogenic due to a lack of evidence for the involvement of multiple loci (Scambler *et al.* 1985, Tsui *et al.* 1985, Lander and Botstein 1986).

While no evidence for genetic heterogeneity was established, the impact of the allelic heterogeneity at the *CFTR* locus has been shown to contribute to the phenotypic variation as demonstrated by an association of *CFTR* spanning haplotypes with the pancreatic status (Kerem *et al.* 1989b) and sweat sodium values (Witt *et al.* 1991). Even before detailed analysis of the *CFTR* gene allowed the identification of more than 700 sequence variants (The Cystic Fibrosis Genetic Analysis Consortium 1998), the variability of the CF phenotype was attributed to allelic heterogeneity based on haplotype analysis at the *CFTR* locus (Kerem *et al.* 1989a, Gasparini *et al.* 1990). Based on the patient's pancreatic status, mutations are termed "mild" (e.g. A455E, R347P) or "severe" (e.g. ΔF508) with pancreatic sufficiency resulting as a consequence of at least one "mild" *CFTR* mutation. However, even for a group of patients homozygous for the most common CF mutation ΔF508, phenotypic diversity has been noticed (Johansen *et al.* 1991, Lester *et al.* 1994).

The in-frame deletion ΔF508 accounts for 60-80% of CF chromosomes within European populations (European Working Group on Cystic Fibrosis Genetics 1990). The comparatively high frequency of the ΔF508 allele has been attributed to genetic drift or to an advantage of the heterozygous state. A complex effect on the family size of *CFTR* mutation carriers was shown (deVries *et al.* 1997) and a reproductive advantage of ΔF508 heterozygotes was demonstrated (Dahl *et al.* 1998a). The benefit of the carrier state was ascribed to a longevity advantage due to resistance to viral and bacterial pathogens (Stuart and Burton 1974, Shier 1979, Meindl 1987, Hollander 1982, Cassano 1985, Rodman and Zamundio 1991, Gabriel *et al.* 1994, Pier *et al.* 1998).

Apart from disease-causing lesions, sequence variants are known within the *CFTR* gene. These polymorphisms are termed neutral due to their occurrence in the non-CF population, but it has been shown that polymorphic variants such as IVS8-6(T) and the missense variant M470V alter the splicing of CFTR mRNA, the processing of CFTR protein and the electrophysiological properties of the CFTR channel (Strong *et al.* 1993, Cuppens *et al.* 1998).

The role of CFTR in diseases other than CF is evident for congenital bilateral absence of the vas deferens (CBAVD) as demonstrated by an increased frequency of *CFTR* mutant alleles among CBAVD patients (Dumur *et al.* 1990b, Chillon *et al.* 1995). Similarly, it has also been suggested that CFTR may play a role in allergy (Warner 1976), chronic bronchial hypersecretion (Dumur *et al.* 1990a), allergic bronchiopulmonary aspergillosis (Miller *et al.* 1996), disseminated bronchiectasis (Pignatti *et al.* 1996), sinopulmonary disease of ill-defined etiology (Friedman *et al.* 1997), asthma (Dahl *et al.* 1998b) and chronic pancreatitis (Sharer *et al.* 1998).

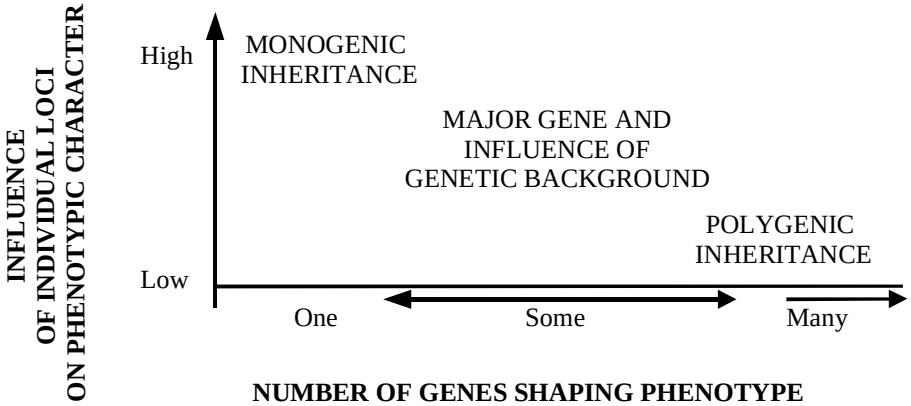
1.4 Analysis of complex inherited traits

Since the work of Mendel in 1865, reporting that characters are transmitted from generation to generation due to independently assorted entities called genes, the genetic basis of phenotypic diversity has been studied. It was realised that groups of genes are ordered linearly along chromosomes and recombination events were taken as a measurement of the distance between genes (Sturtevant 1913). The discovery of the molecular basis of inheritance subsequently enabled the direct investigation of chromosomal sequences, and the analysis of the inheritance of DNA polymorphisms facilitated the construction of genetic maps (Haldane 1919). Distance units on genetic maps are based on observed recombination frequencies, assuming that the recombination frequency between two loci will increase with distance. However, as recombination breakpoints are not distributed equally throughout the genome, the relationship is not linear (e.g. accounted for by Kosambi 1944) and consequently, genetic distances and physical distances between two loci are different.

The genetic diversity is complex on the level of the isolated gene as well as within the 'gene network' (reviewed by Tanksley 1993, Romeo and McKusick 1994, Frankel and Schork 1996). Phenotypic variation based on isolated gene sequences occurs due to sequence variants (allelic heterogeneity) which may express variability in their impact on the phenotype (penetrance) or mode of transmission (dominant or recessive). Considering several gene loci, they can either interact synergistically (epistasis) or act through independent pathways (genetic heterogeneity). Whether interaction between genes will be recognised when studying a phenotypic trait depends on the number of loci involved and on the impact of each gene (figure 1): the inheritance is perceived as monogenic when one locus accounts for nearly all variations of the phenotypic trait studied. If the trait is modified by few additional loci, the role of the extragenetic background will be noticed in phenotypic variation. When many loci have a similar impact on the phenotype, the mode of inheritance is recognised as polygenic.

FIGURE 1: MONOGENIC AND POLYGENIC INHERITANCE

The number of loci involved and the impact of each locus on the phenotype determines whether a trait is perceived as monogenic, influenced by genetic background or polygenic.



Strategies to determine the genetic basis of a trait have been applied successfully to map more than 500 genes to their chromosomal regions (Lander and Kruglyak 1995). Methods involved pursue two different concepts, namely linkage and association (reviewed by Lander and Schork 1994). Linkage studies rely on the observation of the inheritance of alleles and the transmission of a phenotypic character within a pedigree. If an allele is transmitted in concordance with the phenotype, linkage between the locus studied and the gene involved in determining the character is assumed. Association studies compare allele frequencies between a group expressing a trait and a control group. Allelic imbalance, e.g. an increased frequency for an allele in one group, is considered as an indication of association between the locus studied and the gene responsible for shaping the trait.

Linkage and association methods detect the genetic distance between a marker and the gene transmitting the phenotype under investigation. However, when either the influence of the gene on the character studied is too small or the marker locus investigated is too far away from the gene of interest, neither method will provide evidence for the gene's location. Both elements can be taken into account by the study design: when analysing a quantitative trait, the influence of minor genes is expected to be more pronounced at the extremes of the distribution (reviewed by Tanksley 1993) and, consequently, selection of extreme phenotypes can increase the number of individuals who either possess or lack the gene of interest (Risch and Zhang 1995). The selection of the marker locus also varies: genome-wide screens, employing markers typically spaced in 20 cM intervals, cover the entire genome but might miss a gene of interest due to the space between the loci tested. Alternatively, markers can be placed with higher density in gene-rich regions of the genome (Antonorakis 1994, Inglehearn 1997), thereby decreasing the distance between the marker locus studied and the gene of interest when it is located within gene-rich regions. Finally, candidate genes, selected due to a hypothesis on their involvement in shaping the phenotype, can be tested directly when a sequence variant of sufficient informativeness is either intragenic or close to the candidate gene.

1.5 Study designs and objectives

Within this thesis, results from two research trials investigating genotype-phenotype relations in CF are presented: the study *"Rare genotypes and atypical CF"* analyses patients with uncommon *CFTR* alleles while the *"European CF Twin and Sibling study"*, carried out under the auspices of the European Working Group on Cystic Fibrosis, is focussed on the most common *CFTR* allele ΔF508. Both trials were designed as collaborative studies between the CF centers of Rotterdam, The Netherlands and Hannover, Germany. Recruitment of patients for the *European CF Twin and Sibling study* involved CF clinics in Austria, Belgium, France, Germany, Italy, the United Kingdom, Spain, Sweden, Switzerland and The Netherlands. Investigation of patient pairs was conducted in Hannover, Germany; Innsbruck, Austria; London, United Kingdom; Rotterdam, The Netherlands and Verona, Italy.

For both studies, the investigation of the clinical phenotype – as evaluated in accordance with the guidelines of the BIOMED genotype-phenotype workshop in 1994 by pulmonary function, characterisation of nutritional status by weight and height, serum analysis and identification of bacterial pathogens in sputum cultures – was combined with an investigation of the basic defect in the sweat gland and in the major affected organs, that is the respiratory and intestinal tracts. Respiratory tissue was analysed by nasal potential difference measurement (NPD) and rectal tissue was analysed by intestinal current measurements (ICM). Both ICM and NPD are sensitive to the basic defect in CF by assessing the chloride conductance defect transmitted by non-functional CFTR protein. Expertise with these electrophysiological methods was provided by the CF research group in Rotterdam, as was the equipment to measure NPD and ICM.

1.5.1 Rare genotypes and atypical CF: genotype-phenotype correlation on a case-to-case basis

The term "atypical cystic fibrosis" refers to an uncommon manifestation of the disease CF. This can appear as a general mild course of the disease or as a tissue specific expression – as described for CBAVD – where helpful diagnostic methods such as sweat test and mutation analysis do not provide clear evidence for CF. A patient, presenting mild disease symptoms and a close-to-normal sweat test, might not be considered as a CF patient if the analysis of the *CFTR* gene does not sustain the diagnosis. On the other hand, if the same patient is shown to carry two rare *CFTR* alleles with unknown impact on the CFTR protein, he might be considered to be a CF patient even if the chain of evidence between the molecular lesion within the *CFTR* gene and the disease symptoms presented is far from being conclusive. Thus, a characterisation of the phenotype inflicted by uncommon *CFTR* mutations and a characterisation of the *CFTR* gene in patients with atypical CF was attempted by the study *"Rare genotypes and atypical CF"* with the aim of correlating *CFTR* genotype and CF phenotype on a case-to-case basis.

CFTR mutation genotypes have been categorised in accordance with the anticipated effect of the molecular lesion (reviewed by Zielenski and Tsui 1995): no CFTR protein is expected for most nonsense mutations, partly impaired regulation or function has been suggested for some missense mutations and reduced amounts of wild type CFTR protein can result from some splice site mutations. Within the study *"Rare genotypes and atypical CF"*, patients carrying uncommon *CFTR* alleles that cover the broad spectrum with respect to the probability of residual CFTR activity were investigated. A total of 19 patients were examined: 7 patients were homozygous for a non-ΔF508 mutation, 7 patients carried at least one rare *CFTR* allele and five patients were suspected of bearing at least one functional *CFTR* allele in spite of CF symptoms. Three of the patients analysed bear sequence variations which have not previously been found within the *CFTR* gene: two novel deletions are presented in chapter 3.1.1.1 and a novel intronic sequence variant is described in chapter 3.1.1.2.

A widely accepted approach to correlating genotype and phenotype in cystic fibrosis is based on explaining the latter only on the basis of the former, i.e. on deriving the clinical outcome only from the impaired chloride conductance of CFTR expressing epithelia and on deriving a defect in ion conductance solely on the basis of the underlying *CFTR* genotype. This concept can only be successfully applied to explain a patient's clinical state when the *CFTR* genotype is assumed to be predictive for the basic defect in all affected tissues and, likewise, when the impaired chloride conductance is assumed to be predictive for the clinical outcome. Two chapters of the thesis address this issue: chapter 3.1.3.1 examines whether the basic defect in three tissues – namely the sweat gland, the nasal tissue and the rectal tissue – correlates due to the CFTR defect transmitted to all tissues or whether other, tissue specific ion channels diminish this correlation. Chapter 3.1.3.2 inspects whether the results from the basic defect measurements can be explained entirely on the basis of the *CFTR* genotype and whether the disease severity can be based entirely on the observed – if any – residual chloride conductance in investigated tissues.

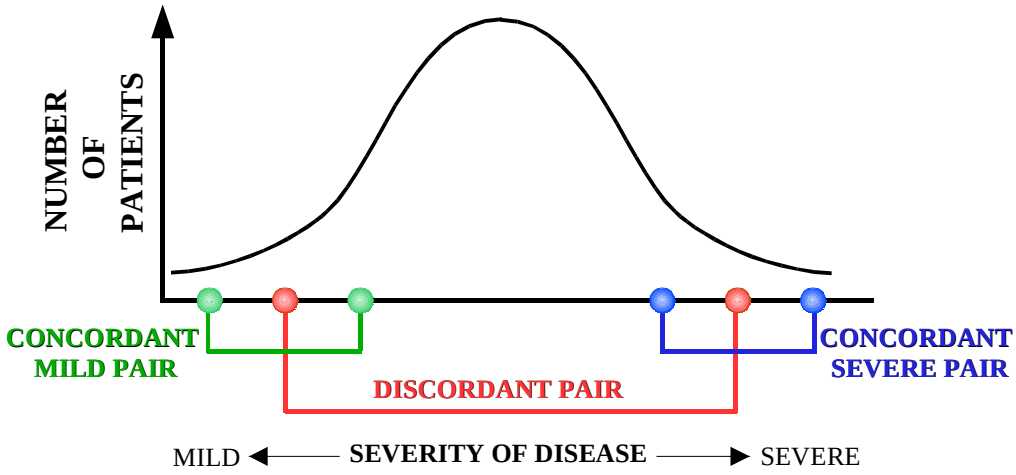
1.5.2 European CF twin and sibling study: genotype-phenotype correlation through patient pair cohorts

The variability of CF disease severity can be attributed to the *CFTR* genotype, epigenetic factors and disease modulating genes. The *"European Cystic Fibrosis Twin and Sibling Study"* pursues a classical approach to address the impact of each factor by investigating affected patient pairs. The influence of epigenetic factors can be estimated by analysing monozygous twin pairs. Being genetically identical, they can only vary in phenotype based on non-inherited factors. Dizygous CF patient pairs harbour the same *CFTR* mutation, share mostly the same environment but only part of their genetic background. Thus, when comparing monozygous twins and dizygous pairs with respect to their intrapair variation – i.e. the difference between two siblings or twins of a pair in a phenotypic character – the impact of the differences in the genetic background can be assessed. Furthermore, in central European populations, approximately 70% of CF alleles bear the same *CFTR* mutation, ΔF508 (European Working Group on Cystic Fibrosis Genetics 1990). Consequently, 50% of all patient pairs are homozygous for the same disease-causing lesion, a fact that facilitates analysis of the disease severity in a group with a homogeneous mutation genotype in the major disease-causing gene. CF is the only disease where a sufficient number of patient pairs that fulfil these criteria can be recruited, due to the prevalence of one mutation genotype in a so-called monogenic disease that follows an autosomal recessive trait.

For identifying a gene involved in shaping a complex trait, the extreme phenotypes — presenting either an extremely mild or an extremely severe clinical picture — are considered to be most informative. Three groups of pairs can be defined based on the phenotype of the individual (see figure 2): discordant pairs, composed of two sibs with highly different phenotypes (DIS), concordant pairs, consisting of two sibs with mild phenotypes (CON+), and concordant pairs, comprised of two

sibs with severe phenotypes (CON-). When selecting CF twin and sibling pairs exhibiting extreme phenotypes, the clinical phenotype, the basic defect — being accessible by nasal potential difference and intestinal current measurement — and selected candidate gene loci composing part of the genetic background were investigated. This thesis reports on the following results from the study: the selection of CF twin and sibling pairs for analysis of selected candidate genes is described in chapter 2.2.1. Conclusions drawn from the clinical epidemiological data are shown in chapter 3.2.1. Results of the ICM experiments are discussed in chapter 3.2.2. In chapter 3.2.3, the modulation of CF disease severity by the genetic background is examined.

FIGURE 2: CONCORDANT AND DISCORDANT PATIENT PAIRS



The pathophysiology of CF provides numerous opportunities to make an "intelligent guess" about the involvement of a gene product in the modulation of this disease. However, to test the hypotheses with a genetic approach, the localisation of the gene must be known in order to select DNA markers for an association study: in order to detect the influence of a functional polymorphism within a candidate gene on the phenotype by analysis of a closely linked neutral polymorphism, marker locus and candidate gene locus must be in linkage disequilibrium.

Within this thesis, the amiloride sensitive, epithelial sodium channel ENaC was analysed for its role as a candidate gene. This sodium channel is a likely candidate for a modulation of CF disease severity due to two observations: on the molecular level, an interaction of CFTR and the ENaC has been demonstrated. Furthermore, a disease called pseudohypoaldosteronism type I (PHA) is caused by mutations within the ENaC coding sequences. Some PHA patients share characteristic symptoms with CF, e.g. an increased NaCl concentration in sweat. Results from the association study on polymorphic markers near the ENaC coding sequences are reported in chapter 3.2.3.4.

Besides candidate genes located anywhere in the genome, sequences near the *CFTR* gene might play a role in the modulation of CF disease severity due to their physical location which results in a tight linkage between the CF disease causing gene and flanking sequences, the so-called hitchhiking genes. Hitchhiking genes have been the subject of discussions as to whether they might confer an advantage to the chromosomal background shared by many mutant *CFTR* alleles, as indicated by the common CF haplotype. 3.2.3.5 shows the results of the association study with markers localised near *CFTR*. As will be illustrated in detail within chapter 3.2.3.6, the results obtained from markers within the *CFTR* linkage group have pointed to a modulation mediated by the proopiomelanocorticoin gene POMC that is most likely involved in the regulation of the expression of one *CFTR* hitchhiking gene.

An association study primarily detects an allelic imbalance, i.e. in the case of the "*European CF twin and sibling study*" an overrepresentation of one allele at given marker locus among pairs ranked DIS, CON+ or CON-. The modulation of the disease phenotype that can be inferred from an allelic imbalance is directly accessible through clinical parameters that describe the disease severity. Parameters that describe the pulmonary and the nutritional status were available for the ΔF508 homozygous patients typed. In chapter 3.2.3.7, the attempt to dissect whether a candidate gene region influences the CF nutritional or the pulmonary state – or both – was made.

Finally, chapter 3.2.3.8 introduces hypotheses that might explain the observed results on the genomic level. While the hypothesis is straightforward for the interpretation of the results gained on markers near the ENaC coding sequences because a direct hypothesis facilitated the choice of the ENaC as a candidate gene, more than one gene within the *CFTR* gene region serves as a likely candidate for the modulation of the CF disease phenotype within a complex gene network.

1.5.3 Aim of the thesis

Cystic fibrosis is known as the most severe inherited disease that follows an autosomal recessive trait. The course of the disease varies between "mild" and "severely affected". However, CF is a lethal disease and the disease symptoms as well as the considerable therapeutic efforts definitely reduce the patient's quality of life. Several implications for CF therapy arise from the fact that CF is considered as a monogenic disease: most basic research of the last decade was focussed on the *CFTR* transmitted defect, e.g. by somatic gene therapy (reviewed by Boucher 1999) or through pharmacological approaches attempting to correct specific mRNA or protein defects of CFTR mutants (reviewed by Zeitlin 1999). While substantiated by *in vitro* data and aiming at a permanent cure for CF, both approaches have run into technical complications that might well jeopardise their practicability *in vivo*. Besides CFTR directed interventions, therapeutic measures concentrate on specific CF symptoms, e.g. antibiotic treatment of patients infected by *Pseudomonas aeruginosa* or supplementation of pancreatic enzymes for pancreatic insufficient patients. While these therapeutic regimes have considerably lengthened the patient's life span, the clinical outcome is characterised by an interindividual variability of which the basis is poorly understood. Thus, it is of interest to the CF community to identify the basis of disease variability, thereby enabling the identification of factors

which (a) might be more amenable to therapeutic intervention or (b) have to be taken into account when applying current therapeutic measures. However, a *clinically significant* impact on the CF disease phenotype has to be verified. This thesis has the intention of evaluating whether besides the *CFTR* gene the genetic background influences the severity of CF disease in a substantial manner.
The thesis reports on results from two clinical studies which have been introduced in detail within the last chapters, viz the study "*Rare genotypes and atypical cystic fibrosis*" and the "*European cystic fibrosis twin and sibling study*". While the former is focussed on a case-to-case-approach, the latter is based on the study of patient pair cohorts. However, both studies aim at a correlation of the CF disease phenotype to the underlying *CFTR* genotype, the extragenetic background and epigenetic factors — three elements which can be considered to account for the variability of CF disease severity that characterises this multi-organ disorder. The data obtained within each study are discussed separately in chapter 3 of this thesis. The final section of this thesis, chapter 4, entitled "Synopsis", is given over to a discussion that links the results from both studies. In an attempt to summarise and draw conclusions from the two studies of this thesis as well as material published by many authors in the field of CF research, four questions are addressed in chapters 4.1, 4.2, 4.3 and 4.4 concerning the relative impact of the relative impact of the *CFTR* genotype, the genetic background and epigenetic factors on the course of CF disease.

CFTR genotype- the basis of CF disease phenotype?

For a disease that is considered to be monogenic, a straightforward genotype-phenotype correlation is expected. If only one locus – the *CFTR* gene – plays a dominant role for the disorder CF, it should be possible to predict the course of the disease on the basis of the disease-causing lesion. As CF is characterised by allelic heterogeneity, it has been speculated that most of the variation observed among CF patients results from compound heterozygous genotypes. Moreover, the basic defect of impaired chloride conductance in CFTR expressing epithelia was often considered as the sole basis of phenotypic variation since the impairment of *CFTR* function can be estimated from the underlying disease causing lesion. The investigation of CF patients who are homozygous for rare *CFTR* genotypes presents the opportunity to address the issue of whether different *CFTR* mutation genotypes predispose to a presentation of specific trademarks in CF diagnosis and clinical prognosis – i.e. the sweat chloride concentration, the pancreatic status or the severity of pulmonary disease. This approach is targeted by the study "*Rare genotypes and atypical cystic fibrosis*" through an analysis of the clinical picture and the basic defect in CF of non- ΔF508 homozygous patients.

CF disease variability- the basis of non-CFTR genetics?

If the variation of CF disease severity cannot be explained on the basis of the *CFTR* mutation genotype, the remaining elements which can be held responsible are epigenetic factors and the genetic background. Evidence for the relative impact of inherited vs non-inherited elements can only be gained by comparing the intrapair variation of CF disease severity in monozygous twins to

dizygous patient pairs: as monozygous twins are genetically identical, all variation observed when comparing two twins of a pair has to result from epigenetic factors. The "*European CF twin and sibling study*" is at present the only clinical study where information on the nutritional and pulmonary status of a sufficient number of patient pairs has been collected to pursue this issue.

Advantage of CF carriers- none, CFTR transmitted or hitchhiking genes?

The frequency of CF causing alleles in the general population has often been the subject of debate as to whether a selective advantage is associated with the carrier state. Alternatively, a random fixation by genetic drift – a mechanism which does not rely on a hypothetical benefit of mutant *CFTR* alleles – has been taken into account. If the maintenance of the allele frequency of *CFTR* mutants is based on a selective advantage, two alternative interpretations are possible: it has been speculated that mutant CFTR protein itself is advantageous by providing protection against viral and bacterial infections. In contrast, genes at loci near *CFTR* which are transmitted in parallel to the mutant *CFTR* allele – the hitchhiking genes – were considered to transmit a selective advantage. The latter is sustained by the observation of a chromosomal background common to many frequent *CFTR* mutations that is indicated by the CF haplotype. If a selective advantage is conveyed by sequences adjacent to *CFTR*, their effect should be seen not only in carriers, but also in the affected patients who harbour two CF alleles. In order to test the hypothesis as to whether *CFTR* flanking sequences influence the CF disease phenotype and thus may be able to convey a heterozygote advantage, the genomic region near *CFTR* was investigated in an association study on ΔF508 homozygous patient pairs recruited for the "*European CF twin and sibling study*". As ΔF508 homozygous patient pairs have to be considered as identical with respect to the *CFTR* gene itself, any association of alleles at flanking loci with the disease phenotype has to be caused by hitchhiking genes.

CF- monogenic disease or complex trait?

The state of a monogenic disorder, by definition caused by one gene locus, is opposed by two conditions, i.e. genetic heterogeneity and polygenic inheritance. While genetic heterogeneity infers that the phenotype can be caused independently by two loci, polygenic inheritance indicates that many different loci are simultaneously involved in shaping the trait. Neither of the studies this thesis reports on, examines CF as a strictly monogenic disease. Within the study "*Rare genotypes and atypical cystic fibrosis*" causes for CF other than two defective *CFTR* alleles are targeted by investigating patients displaying CF symptoms while lacking evidence for – or even providing evidence against – the presence of two disease-causing lesions within the CFTR gene. Whether CF behaves as a monogenic disease or has to be considered as substantially influenced by the genetic background is challenged by clinical data from the "*European CF twin and sibling study*".

2 METHODS

Unless described differently, all chemicals were purchased as p.a. substances from one of the following companies: Boehringer (Mannheim), Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg), Sigma (Deisenhofen). Composition of common reagents and buffers supplied by the manufacturer of enzymes are listed in chapter 8.3.

2.1 Analysis of genotype

Most methods used within this thesis to analyse a genomic sequence rely on polymerase chain reaction (PCR) methods, either to screen for and to identify a sequence variation of the *CFTR* gene (chapter 2.1.2.1) or to identify alleles at a marker locus for multiple samples (chapter 2.1.2.2). A deletion involving exon 2 and exon 3 of the *CFTR* gene was verified by Southern blotting (chapter 2.1.3).

2.1.1 DNA isolation

2.1.1.1 Preparation of high molecular weight DNA

High molecular weight DNA was isolated as described by Gross-Bellard *et al.* (1973). 5-10 ml K-EDTA blood, either fresh or thawed on ice after storage at –20°C, were incubated with 40 ml lysis buffer (50 mM Tris-HCl, 109.5 g/l saccharose, 1% w/v Triton X-100) for 30' on ice. Nuclei of lymphocytes were separated from cell fragments by centrifugation (15', 600 × g). Upon incomplete lysis of erythrocytes, detectable by the dark red colour of the pellet, the lysis procedure was repeated. The pellet was treated with 5 ml of 0.25 mg/ml proteinase K, 0.5% w/v SDS in STE for 8 to 14 h at 56°C in a shaking water bath. Digested proteins and high molecular weight DNA were separated by phenol/chloroform extraction: 3 ml chloroform/isoamylalcohol (29:1) and 3 ml phenol were added to the digest. After cautious mixing for 15' and centrifugation (10', 600 × g) the organic phase collected at the bottom of the tube was carefully removed. This extraction step was repeated once, followed by an extraction with 6 ml chloroform to remove phenol remaining in the aqueous phase. DNA was precipitated on ice by adding 1/10 of the volume of the aqueous phase of 3 M sodium carbonate (pH 5.5) and 30-40 ml of 70% ethanol precooled to -20°C. High molecular weight DNA was transferred to an Eppendorf cup containing 1 ml 70% ethanol and centrifuged (3', 12000 × g). The supernatant was discarded and the pellet washed twice with 70% ethanol. After dissolving DNA in 300 µl TE (4°C, 1 week), a 1:25 dilution of the stock solution was prepared and the DNA concentration determined by OD at 260 nm.

2.1.1.2 Chelex 100

A fast method yielding enough DNA to perform 20 to 40 PCR reactions was employed using Chelex 100 (Walsh *et al.* 1991), a cation exchanger binding divalent metal ions which are known to cause DNA degradation at higher temperatures in aqueous solution. 3-10 µl of K-EDTA blood were shaken at room temperature for 30 min. After centrifugation at 13000 × g the supernatant was discarded and the pellet dissolved in 200 µl of a 5% w/v Chelex 100 suspension. The mixture was incubated at 56°C for 30', vortexed for 10', incubated for 8' in a boiling water bath, vortexed for 10' and again centrifuged at 13000 × g. 5-10 µl of the supernatant were used as a template for PCR reactions.

2.1.2 PCR methods

The polymerase chain reaction is used for amplification of DNA sequences. The specificity of the reaction is ascertained by the precise binding of the primer molecules to their target sites adjacent to the sequence of interest. All factors that determine the association and dissociation of complementary DNA strands – base composition, salt concentration, temperature – have an influence on the accuracy of the repetitive annealing procedure and thus on the formation of PCR product.

The base composition of the primer molecules is largely dependent on the target sequence site. However, the 3' end of a primer, being elongated during PCR reactions if bound to a template, is preferably composed of a sequence unique to the chosen site and rare throughout the genome studied. Griffais *et al* (1991) have provided an algorithm to assign primer locations on given DNA sequences which takes the frequency of octamers throughout the genome into account, favouring rare octamers on the primer's 3' end. Primer binding can be facilitated by providing counterions for the negatively charged DNA backbone and thereby reducing the repulsion of phosphate groups. The salt concentration within PCR reactions is modified by $MgCl_2$ which is also a cofactor for the Taq polymerase. Besides the buffer composition, the stringency of PCR reactions is controlled most rigidly by the temperature conditions governing the annealing step of primer to template. This includes the annealing temperature itself as well as the temperature gradient applied when cooling from the denaturation step to the annealing temperature.

Apart from the thermodynamic conditions, the probability of a primer binding to its target site as opposed to binding at any other site with sufficient sequence similarity will depend on the ratio of unbound primer to DNA sequences available. This is influenced by the amount of DNA used as a template and changes rapidly during PCR amplification because products that contain the primer sequences are formed.

The standard protocol using 400 ng of DNA and 25 pmol of each primer in a 50µl reaction mix containing 2 mM $MgCl_2$ and 200 µM of each dATP, dGTP, dCTP and dTTP was varied for different applications. Less $MgCl_2$ (1-1.5 mM) and less dNTP's (50-100 µM) lower the product yield but hinder the formation of unspecific primed by-products, which is used for amplifying SSCP-products (chapter 2.1.2.1.1) and for PCR amplification of probes for Southern hybridisation (chapter 2.1.3).

The commercially available Taq polymerases differ in fidelity; consequently not all Taq polymerases are suitable for all applications. To reduce the experimental costs, a polymerase without proof reading activity was used when the size of the PCR product was analysed, such as RFLP and microsatellite analysis (chapter 2.1.2.2). All methods relying on amplifying the exact sequence, such as sequencing reactions (chapter 2.1.2.1.2), were performed with a Taq polymerase possessing an 5' to 3' polymerisation dependent exonuclease activity (Dap Goldstar, Eurogentec). Details of the reaction conditions used for amplification and primer sequences can be found in chapter .

2.1.2.1 Identification of sequence alterations

Sequence variants within a gene can be identified by direct sequencing of the complete coding sequence. However, depending on the equipment available, it may be more cost-effective and less time-consuming to screen the complete sequence for alterations and subsequently perform sequencing only for those fragments which have provided evidence for a sequence variant.

The screening technique used within this thesis is the single strand conformation polymorphism (SSCP) analysis (chapter 2.1.2.1.1). The fragment of interest, encompassing the respective exon and its flanking intron sequences, was PCR amplified and the DNA strands separated by heat denaturation. Rapid cooling to –20°C allowed each single strand to form secondary structures which depend on the nucleotide sequence. When the sequence is changed, e.g. by a point mutation, the resulting secondary structure will differ from the wild type molecule when the change has occurred at a position responsible for the formation of a secondary structure. If both single strand conformations — assumed by the wild type strand and the corresponding mutant strand — are sufficiently different from each other to be separated on a non-denaturing polyacrylamide gel, the resulting band shift will signal the presence of a mutant allele in a sample, thus leading to sequencing of the fragment (chapter 2.1.2.1.2).

The size of the fragment studied and the gel conditions have been shown to influence the resolution of the SSCP. Fragments of 100-150 bp have been reported to show the best results when looking for band shifts in SSCP gels; low temperature (4°C to 10°C) and glycerol in the polyacrylamide matrix stabilise single strand conformations. The sensitivity can be increased by using two different restriction digests per analysed PCR fragment to increase the probability that a nucleotide change is positioned within a sequence responsible for the formation of secondary structures. An overall detection rate of 100% for 134 known *CFTR* mutations, screened under optimised SSCP conditions, has been shown (Ravnak-Glacik *et al.* 1994).

2.1.2.1.1 Analysis of single strand conformation polymorphisms

400 ng of DNA, 25 pmol of each primer and 1 U of Invitaq polymerase (Invitek) were placed in a total reaction volume of 50 µl reaction buffer (Invitek) containing 0.08 mM dATP, dGTP, dCTP, dTTP and 5-7 mM $MgCl_2$. The mixture was subjected to 20 cycles of PCR amplification beneath a paraffin oil overlay of 30 µl. While incubating the reaction mixture at the annealing temperature, 20 µl were transferred on 1 µCi α^{33}PdATP (Amersham) and 0.5 µl of 2 mM dGTP, dCTP, dTTP and amplified for additional 15 cycles beneath a paraffin oil overlay of 20 µl. The remaining 30 µl of the unlabelled reaction mixture were also amplified for another 15 cycles and the products were separated on a 2% agarose gel and stained with ethidium bromide to access the quality. The labelled products were digested with restriction enzymes giving fragments of 150-200 bp. To obtain single strands, the restriction digest was mixed 1:1 with loading buffer (0.2% w/v of xylenecyanol and bromphenolblue in formamid), heated to 95°C for 5 min and directly cooled in an aluminium block precooled to –20°C. Separation of single strand conformations was carried out on a 40 cm gel casted with 0.2-0.4 mm wedge spacers, the matrix containing 6% acrylamide / bisacrylamide 29:1 and 10% glycerol in TBE. One glass plate was precoated with silane A-174 (BDH Chemicals) while the other was precoated with repel silane (dimethyldichlorsilane, Merck).

Electrophoresis was performed at 10°C with 1900 V for 7-9 h. After electrophoresis, the glass plates were separated and the gel, adherent to the silane A-174 coated plate, fixed by incubation in 10% acetic acid for 1 h and subsequent drying for 2 h at 80°C. X-Omat AR (Kodak) films were exposed to the surface of the dried gel from between 12 h to 3 days.

2.1.2.1.2 Sequencing of PCR products

The fragment of interest was amplified in a 100 µl PCR reaction with 50 pmol of upstream and 50 pmol of downstream primer employing a polymerase with high fidelity (DapGoldstar polymerase, Eurogentec). The quality of the PCR product was evaluated and the concentration of the product was estimated by loading 10µl of the reaction mixture on a 2% agarose gel, then checking for by-products and assessing the signal intensity after ethidium bromide staining. Unincorporated primers were removed from the reaction mixture by ultrafiltration (Ultrafree-MC filter unit, Millipore) according to the manufacturer's instructions. The filtrate was adjusted to 10-20 µl, depending on the intensity of the ethidium bromide stained band of the unpurified PCR product. 5-10 µl of the purified PCR product were amplified in an 100 µl reaction with 50 pmol of one primer only, resulting in an excess of linearly amplified single stranded PCR product of one template strand. 10 µl of the asymmetric product were checked for background smear or by-products by agarose gel electrophoresis and the remaining 90 µl were purified with Ultrafree-MC filter units and the volume of the filtrate was adjusted to 10-20 µl. Sequencing was performed manually with the Sequenase 2.0

kit (United States Biochemical). 5-10 µl of purified asymmetrical product and 5 pmol of the primer that was not used for the asymmetric amplification were mixed in a total volume of 20 µl reaction buffer. Annealing was done in a PCR thermocycler programmed to 2 min 80°C and cooling to 40°C within 30 min. Labelling with α^{33}PdATP (Amersham) and termination were carried out according to the manufacturer's instructions.

The products were separated on a 4%-6% acrylamide/bisacrylamide 29:1 gel with 50g urea (Gibco BRL, electrophoresis grade) freshly dissolved to make a total volume of 25 ml of acrylamide solution in TBE. Electrophoresis was performed at 55°C with 1900V for 3-6h. The radioactive labelled products were visualised as described in chapter 2.1.2.1.1, page 33.

2.1.2.2 Genotyping of polymorphisms in PCR multiwell plates

96 well PCR plates (multiwell plates) make it possible to process multiple DNA samples in parallel. However, compared to the conventional set-up in single tubes with 50 µl reactions beneath an oil overlay, several modifications became necessary as the standard protocol for several PCR reactions failed to give rise to a product (figure 3a): conditions that have been shown to amplify a single band within 50 µl tube reactions produced multiple by-products that are characteristic for a PCR with an unbalanced primer to template relationship and high $MgCl_2$ concentrations when used on multiwell plates.

Evaporation during the reaction time is prevented in single tube reactions even without paraffin oil overlay when used in a thermocycler with a heated lid. For multiwell plates, the following approach was used (figure 3b): the outer wells were filled with 50µl of distilled water. After the remaining inner 60 wells were set up for the PCR reaction, the outer rim of the multiwell plate was coated with 1000 µl of molten 1% agarose by pipetting a continuous trail surrounding all wells and pressing the lid on the still liquid agarose.

The stringency of the annealing process, in single tube reactions mostly regulated by the annealing temperature, was adjusted by modification of the $MgCl_2$ concentration (figure 3c). For a reaction mixture containing 50 ng of DNA, 15 pmol of upstream and 15 pmol of downstream primer in a reaction buffer containing 20 µM dNTP's, products below 500 bp could be amplified with 0.8 mM $MgCl_2$ at annealing temperatures between 50°C and 60°C. The same reactions failed to give a product at higher $MgCl_2$ concentrations. Products of 1000 bp or larger were amplified with 2.5 mM $MgCl_2$ at annealing temperatures between 55°C and 65°C.

Multiwell plates were purchased from Greiner. Handling of solutions was done with an Eppendorf eight channel pipette. A Hybaid thermocycler with a heated lid was used for PCR reactions.

FIGURE 3: GENOTYPING OF POLYMORPHISMS IN MULTIWELL PLATES

(a) Amplification in 96 well plate without agarose seal *and outer wells filled with water as shown in (b). The PCR conditions used to generate these products were previously optimised for single tube reaction. As is evident from figure (a), these conditions were found to be unsatisfactory for multiwell PCR.*

(b) PCR set-up with agarose seal *and water within outer wells of PCR plate*

(c) Conditions optimised for multiwell reactions: RFLP markers within the CFTR linkage group

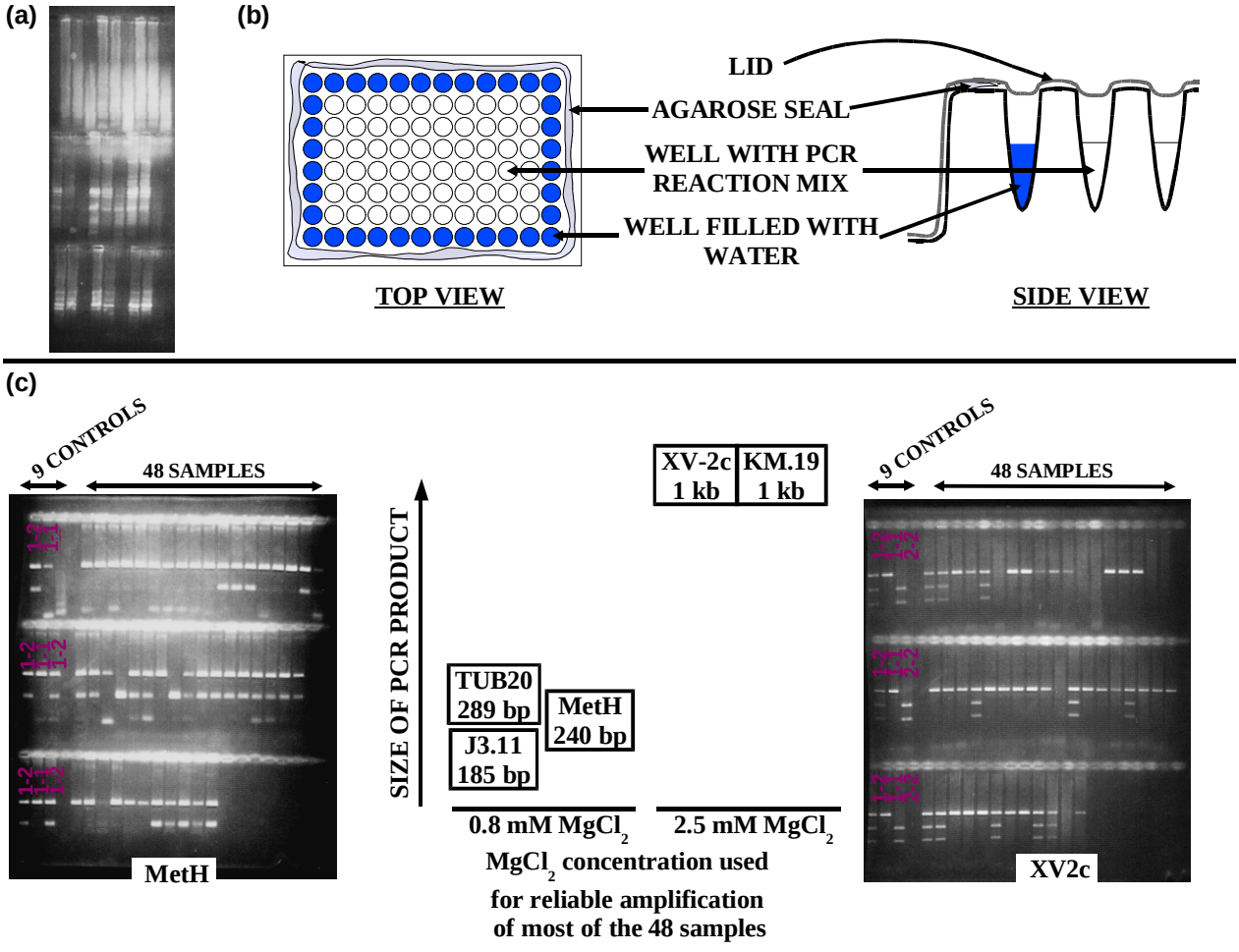
The analyses of dimorphic markers were carried out in the set-up described in (b) which allows PCR amplification of DNA in 60 wells per 96 well plate. To ensure that the patient's samples were PCR amplified and digested with the restriction enzyme appropriately, 12 wells on each 96 well plate were used to amplify four sets of three controls with known genotypes. The four sets of controls were used as follows: one set of three controls was tested prior to digestion with the restriction enzyme in order to ensure that the DNA was PCR amplified (not shown). The remaining three sets of controls were digested in parallel with the 48 samples of the CF patients and their parents and products were separated on the same gel in order to facilitate the assignment of marker genotypes to the samples based on the pattern seen for the controls.

Please note that most of the 48 samples give signals suitable for interpretation as can be estimated by comparison with the controls. A detail that explains the interpretation of the genotypes 1-1, 1-2 and 2-2 for dimorphic markers is shown in figure 5 on page 39.

LEFT: *48 samples and 9 controls analysed at* MetH. *Products after restriction digest separated on a 4% agarose gel. Allele 1: 240 bp, allele 2 (two fragments):100 bp*

MIDDLE*: optimised $MgCl_2$ concentration for PCR fragments of different size.* J3.11, MetH, TUB20, *and* KM.19 *were amplified with an annealing temperature of 55°C,* XV-2c *was amplified with an annealing temperature of 60°C*

RIGHT: *48 samples and 9 controls analysed at* XV-2c. *Products after restriction digest separated on a 2% agarose gel. Allele 1: 1000 bp, allele 2:600 bp and 400 bp*



2.1.2.2.1 Precoating of multiwell plates with DNA

Samples for typing were selected as described in chapter 2.2.1.4. The DNA was diluted in 96 well PCR plates to enable handling of the samples with a multichannel pipette. For each well, the stock solution of the sample was diluted by placing an aliquot of 1µg DNA at the bottom of the well. 100 µl of distilled water were added and the DNA dissolved and sheared with a multichannel pipette set to 80 µl. 5µl of the dilution were used as template for a reaction. A set of 15 plates was prepared and stored at 4°C.

2.1.2.2.2 Analysis of single nucleotide polymorphisms

The PCR was done with the set-up described in figure 3b. 30 µl of product were amplified using 0.3 U per well of Invitaq polymerase (Invitek) in a reaction mixture containing 50 ng of DNA, 15 pmol of upstream and 15 pmol of downstream primer in a reaction buffer provided by the manufacturer containing 20 µM dNTP's and 0.8 mM to 2.5 mM $MgCl_2$. The PCR products were digested on the multiwell plate used for PCR amplification by adding 10 µl of restriction mix containing 2-10 U of the restriction enzyme (NE Biolabs) and 4 µl of the supplied 10 × incubation buffer. The multiwell plates were incubated at the appropriate temperature after a fresh agarose seal, as described in chapter 2.1.2.2, was placed on the outer rim of the multiwell plate to avoid evaporation of samples. Products were separated on 2% or 4% agarose gels casted with wells spaced suitably for multichannel pipette loading of samples. For application into the wells, 10 µl of a 40% v/v glycerol buffer containing 0.05% w/v of bromphenolblue and xylenecyanol were added to the samples. 30 µl were loaded on the agarose gel, products separated by electrophoresis of 10 V/cm and stained with ethidium bromide.

2.1.2.2.3 Analysis of simple sequence repeat polymorphisms

Nucleotide repeat polymorphisms were amplified using an asymmetric primer ratio of 5 pmol biotin labelled primer and 25 pmol of unlabelled primer in an amplification volume of 15 µl containing 1 µM $MgCl_2$, 20 µM dNTP's and 0.25 U / well Invitaq polymerase (Invitek) in the reaction buffer supplied. PCR was carried out with an oil overlay of 10 µl in the set-up described in figure 3. 8 µl of PCR product were transferred to a new multiwell plate and dried by incubation for 6 h at 40°C without an agarose seal (see figure 3) in a thermocycler. Products were dissolved in 10 µl of formamid containing 0.2% w/v bromphenolblue and xylenecyanol. Prior to loading the products onto the gel, DNA fragments were denatured by heating to 95°C for five minutes and subsequently cooling the multiwell plates by placing each multiwell plate in a tight fitting form made of modelling clay precooled to –20°C and embedding the assembly in ice.

Amplified products were separated by direct blotting electrophoresis (Beck and Pohl 1984). The apparatus (GATC, Konstanz) consists of a vertical electrophoresis gel – anode in upper buffer chamber and cathode in lower buffer chamber – with a conveyor belt directly beneath the lower edge of the acrylamide gel (figure 4a). A membrane is mounted on the conveyor belt and transported beneath the lower edge of the acrylamide gel. Products that are separated on the gel exit the matrix at the lower end and are directly transferred onto the membrane. The resolution depends on the matrix used for separation but also on the velocity used for the transport of the membrane.

The direct blotter provides an opportunity for the processing of multiple samples in parallel: once a set of products is blotted on the membrane, the gel can be used for the next set of samples. The microsatellites described in this work had product lengths of 100-200 bp, providing a unique occasion for the consecutive blotting of 210 PCR products on a membrane of 50 cm in length within 5 h (figure 4b). One set of samples was loaded on the gel and separated by electrophoresis. When the bromphenolblue marker reached the lower edge of the gel after 50 min, the electrophoresis was stopped and a second set of samples was loaded. After restarting the electrophoresis, the products of the first set of samples were transferred to the membrane while the second set of samples was separated in the matrix. The blotting of the first set of samples was finished when the bromphenolblue marker of the second set of samples has reached the lower edge of the acrylamide gel. Then, electrophoresis and transport of the membrane was stopped and a third set of samples was loaded. This continuous direct blotting was possible for five to seven consecutive sets of samples with a length of 100-200 bp, equivalent to a time of five to seven hours electrophoresis.

The acrylamide gel was casted and mounted into the apparatus as described in the manufacturer's manual. Microsatellites were separated with 3.5% to 4% acrylamide gels. The velocity of the conveyor belt was set to 17 cm/h when products were transferred to the membrane. Membranes were dried and stored at 4°C until processing.

Biotinylated products were detected employing chemoluminescence detection. The membrane was incubated on a horizontal shaker for 5 min in 100 mM Tris pH 7.5 with 150 mM NaCl, twice for 30-60 min in 100 mM Tris pH 7.5 with 150 mM NaCl containing 1.5% w/v Blocking reagent (Boehringer) and for 1-4 h in 100 mM Tris pH 7.5 with 150 mM NaCl containing 1.5% Blocking reagent and 1 µl/10 ml anti-Biotin Fab fragments, alkaline phosphate conjugate (Boehringer). Detection was carried out after equilibrating the membrane in reaction buffer (100 mM Tris pH 9.5 with 100 mM NaCl and 50 mM $MgCl_2$) for 5 min. Next, the membrane was placed flat in a container and the surface covered with reaction buffer containing 10% v/v Sapphire II (Tropix) and 300 µ / 50 ml CDPstar (Tropix) for 5 min. The solution was retained and the membrane briefly rinsed in reaction buffer containing 1% v/v Sapphire II and 30 µl / 50 ml CDPstar. The membrane was wrapped in plastic foil and immediately exposed to X-omat AR films (Kodak) for 1 min to 60 min.

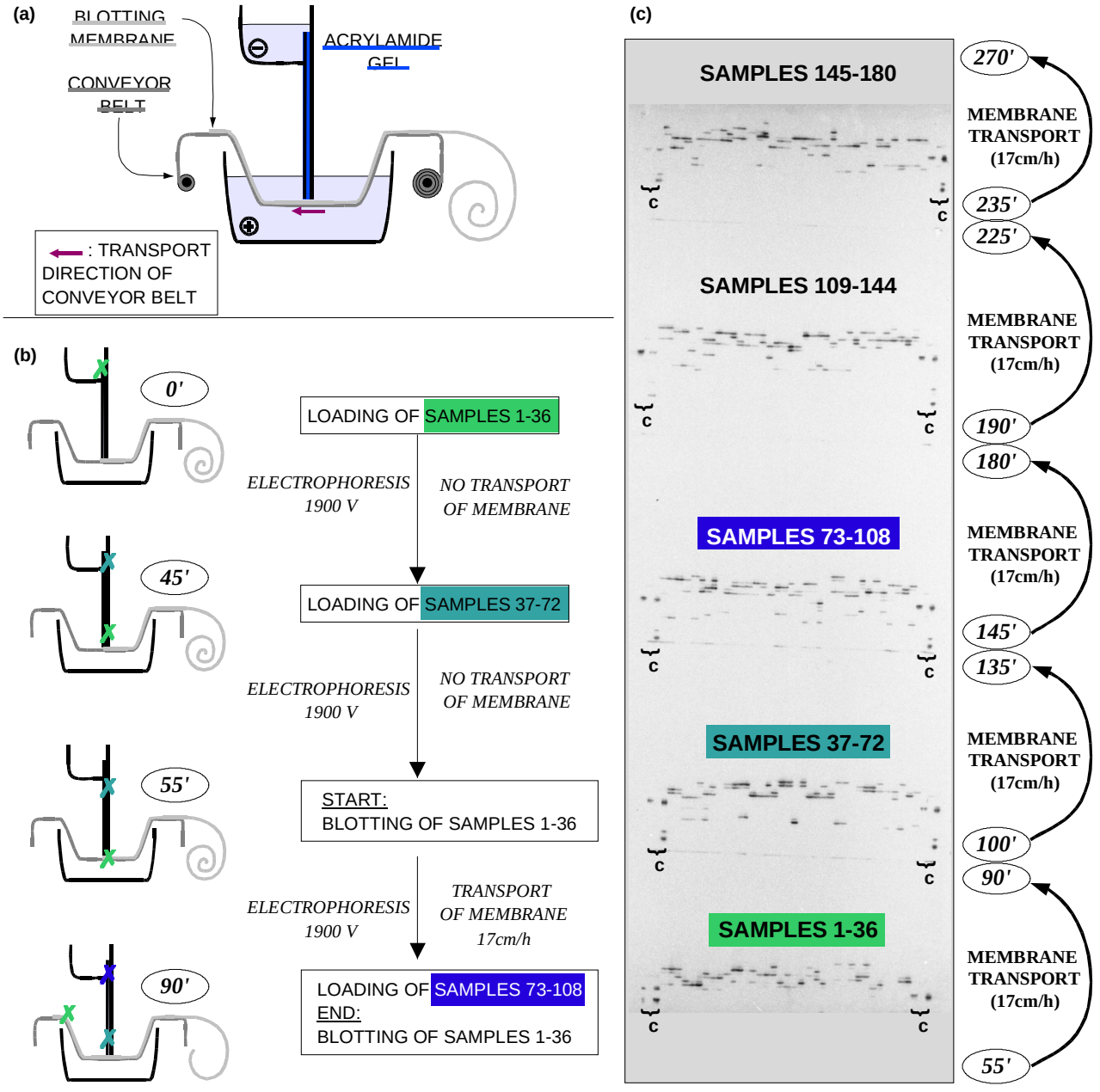
FIGURE 4: CONTINUOUS DIRECT BLOTTING ELECTROPHORESIS

(a) Direct blotting apparatus. The vertical electrophoresis gel is in contact with a conveyor belt at the lower edge of the acrylamide matrix. A membrane is mounted on the conveyor belt and transported beneath the lower edge of the acrylamide gel. Products separated on the gel exit the matrix and are directly transferred onto the membrane. The resolution depends on the matrix used for separation but also on the velocity used for the transport of the membrane.

(b) Transfer of three out of five sets of 36 samples to the membrane in a continuous direct blotting electrophoresis experiment.

(c) 180 samples of D2S1788, blotted with continuous DBE protocol. **c***: set of three control samples employed to calibrate the size of the alleles between the successive sets. The 180 samples represent the total number of CF twins and siblings and their parents analysed at polymorphic loci. A detail of samples 73-108 that demonstrates the assignment of alleles is shown in figure 8 on page 43.*

All time points given in this figure refer to minutes of electrophoresis time. The time needed for loading of the samples at ***45'****,* ***90'****,* ***135'****,* ***180'*** *and* ***235'*** *is not included in this time protocol.*



2.1.2.2.4 Marker analysis and interpretation

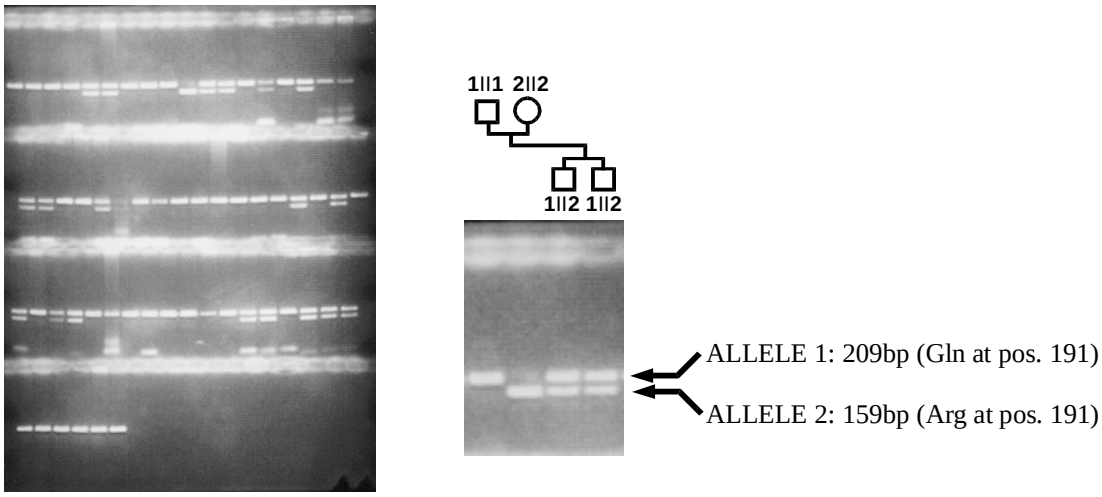
The PCR based study of genetic markers finds multiple implementations in human genetics: pedigree analysis aids in genetic counselling, it may be used for identification of parentage and allows the mapping of inherited traits. As all methods rely on the comparison of PCR generated signals between different individuals, the key to all applications of marker analysis is the unequivocal assignment of the allele's identity irrespective of the genomic source.

FIGURE 5: RFLP PATTERNS OF PON2

(a) PON2 sequence variant typed by DpnII digest of PCR product. The protocol by Adkins et al. (1993) was modified: a mismatch in the forward primer was generated to destroy a second DpnII site within the amplified sequence (see chapter 8.3) thus allowing easier discrimination of cleaved and uncleaved products on a 4% agarose gel.

LEFT: *54 samples separated after in-well digest as described in chapter 2.1.2.2.2*

RIGHT: *detail showing the three possible genotypes 1-1, 1-2 and 2-2 for this dimorphic marker*



Single nucleotide polymorphisms (SLPs), typed as restriction fragment length polymorphism (RFLP) when a site for an restriction endonuclease is either generated or destroyed upon nucleotide exchange at the SLP locus, are analysed by digestion of the PCR product and separation of the resulting fragments. The pattern can be interpreted unequivocally when the digested and the undigested product are separable under the gel conditions used and when the homozygous condition for the presence of a restriction site can be discriminated against the heterozygous state (figure 5).

Microsatellites normally present more than two alleles and, consequently, more than three genotypes can be observed. Furthermore, PCR amplification of repetitive sequences can give rise to by-products which differ by even numbers of repeat units in size, as do alleles at a polymorphic locus. The resulting pattern of alleles and by-products of both chromosomes studied can obscure the underlying genotype. Within this thesis, repeat polymorphism patterns were interpreted in two steps as described within the next two paragraphs: first, alleles and by-products were distinguished by analysing genotypes within pedigrees (figures 6 and 7). Next, the size of alleles was calibrated in arbitrary repeat units (ARU) by assigning a repeat unit of 10 to an allele and successively deducing the size of smaller or larger alleles throughout the gel (figure 8).

All figures shown are electronic images generated from X-ray films. However, the interpretation of band patterns was carried out solely on the basis of the autoradiograms and not of any photographic or electronic copy thereof. The electronic images presented within this chapter were adjusted for contrast and brightness to mimic the band pattern observed on the X-ray films after reproduction on a high resolution ink jet printer.

Interpretation of repeat polymorphism patterns (I): Identification of alleles

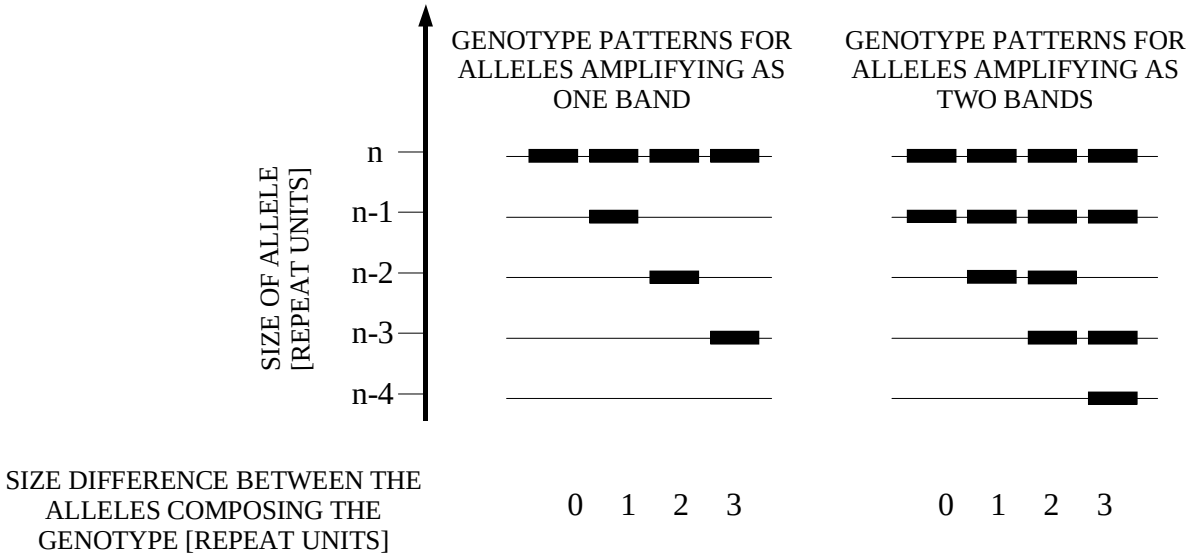
The pattern observed depends upon the products generated by PCR, which in turn will depend on the sequence of the repeat itself: repetitive sequences can be aligned asymmetrically by causing small segments of one DNA strand to form a loop. However, only structures with sufficient stability during the temperature conditions used for amplification that form a template-bound 3' end allowing elongation by the Taq polymerase will lead to PCR product formation (PCR artefacts). Additionally, the conditions used for separation of PCR products will determine the product pattern: sharp bands are produced by DNA fragments of equal size and identical conformation. Incomplete denaturation of products or reassembly of complementary sequences after denaturation, occurring especially when high concentrations of product are loaded on the gel, can form smeared bands due to inhomogeneous structural composition of the products (gel artefacts).

FIGURE 6: PREDICTED PATTERNS OF MICROSATELLITE GENOTYPES

Genotype patterns expected for repeats amplifying with one band per allele (left) and two bands per allele (right).

Examples for both conditions are shown in figure 7, page 42:

D7S495	*figure 7 (a), (b), (c)*	*PCR amplification yielded one band per allele*
D12S889	*figure 7 (d), (e), (f)*	*PCR amplification yielded two bands per allele*



Within this thesis, polymorphic markers were studied using direct blotting electrophoresis enabling simultaneous processing of 180 samples (see chapter 2.1.2.2.3 for details). This allowed the analysis of band patterns for several families and monozygous twin pairs, having been PCR amplified and separated under comparable conditions, thus being prone to similar PCR and gel artefacts.

First, the number of bands representing an allele was determined. It was assumed to be one (observed for D7S495, βENaCGT and D7S525, respectively) unless more than two bands were detected in nearly all lanes (observed for D12S889 and D7S514, respectively). Next, the resulting pattern was predicted for genotypes composed of two alleles of identical size, two alleles differing by one repeat unit in size, two alleles differing by three repeat units in size and genotypes with larger size differences between the alleles (figure 6). Comparison of predicted patterns to genotypes observed yielded rules used for the interpretation of all samples typed at that locus.

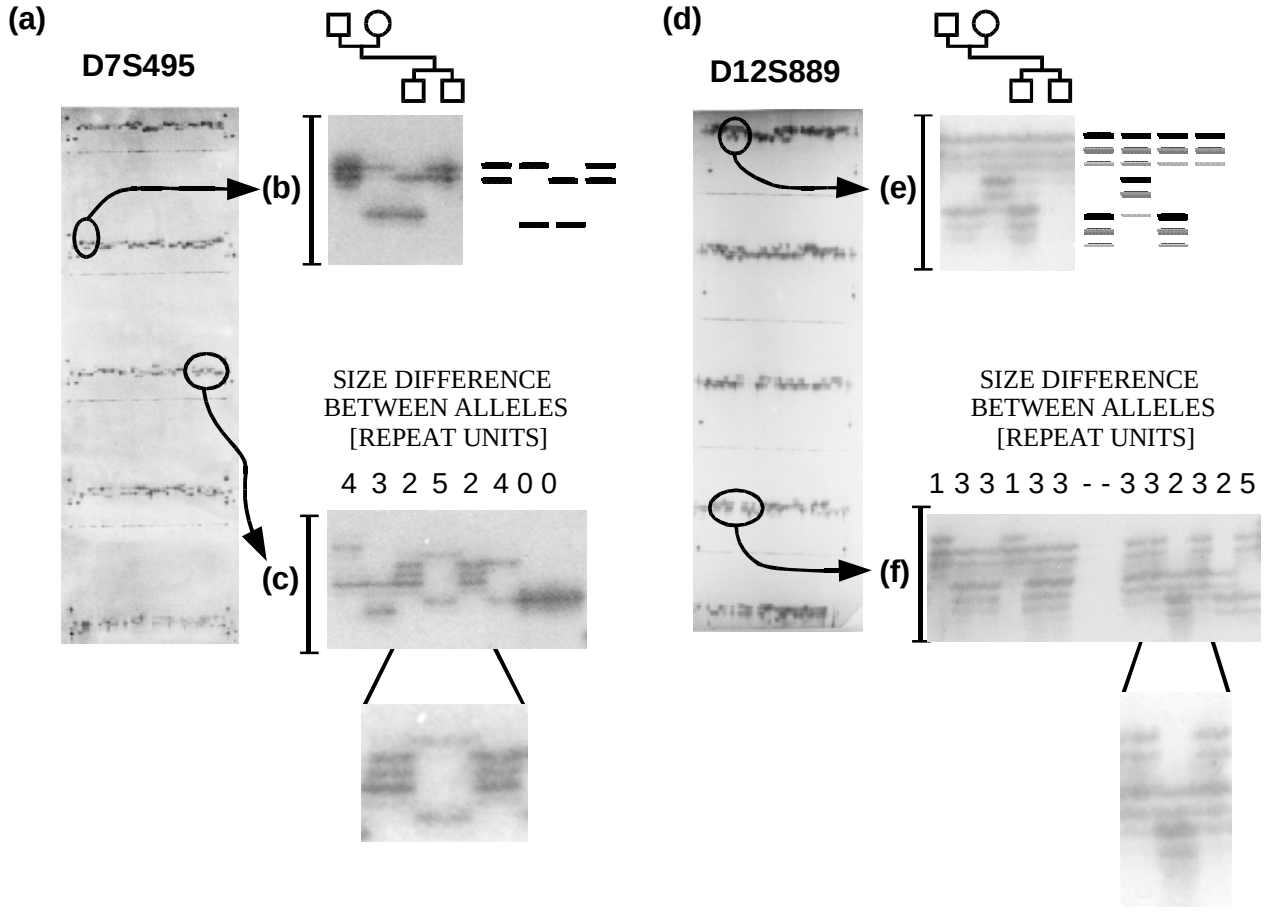
FIGURE 7: BAND PATTERN OF TWO DINUCLEOTIDE REPEATS

Band patterns for 180 samples in (a) (locus D7S495*) and (d) (locus* D12S889*) were generated by continuous direct blotting electrophoresis as described in chapter 2.1.2.2.3.*

D7S495 *(b) pedigree showing two bands per allele (c) band patterns for different genotypes*
D12S889 *(e) pedigree showing one band per allele (f) band patterns for different genotypes*

Figure (a)-(c): pattern as displayed in figure 6, page 41 for satellites amplifying with one band per allele
Figure (d)-(f): pattern as displayed in figure 6, page 41 for satellites amplifying with two bands per allele

See text for details.



Each microsatellite studied showed individual characteristics in terms of by-products and background smear. Two examples are shown in figure 7: the dinucleotide repeat D7S495 was predominantly amplified with one band per allele (shown by the pedigree in figure 7b) while the dinucleotide repeat D12S889 continuously showed two bands per allele (as shown by the pedigree in figure 7e). Consequently, as predicted by figure 6, the number of bands observed for genotypes composed of two identical alleles or alleles differing one, two or more repeat units in size was different for D7S495 and D12S889.

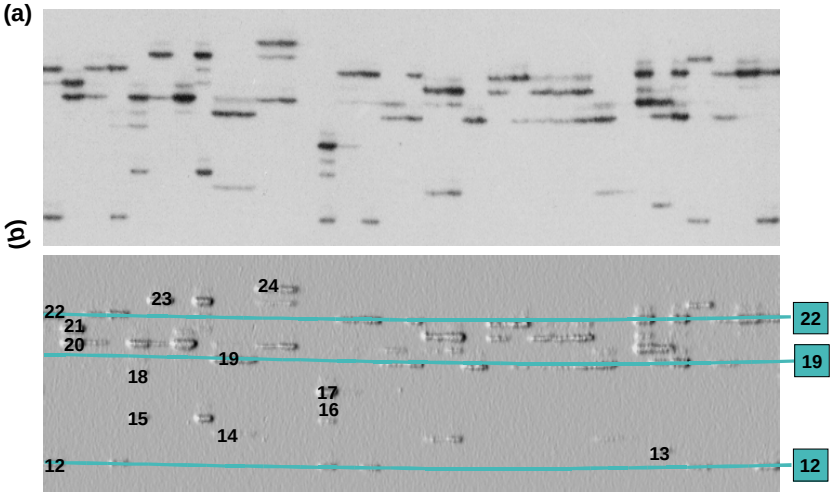
Genotypes with size differences between the alleles of 0 to 5 repeat units are displayed in figure 7c and figure 7f, for D7S495 and D12S889 respectively. The detail of figure 7f shows for D12S889 the four band pattern predicted for a genotype composed of two alleles with a difference in size of 2 repeat units. D7S495 however, showed an exception to the pattern predicted for a repeat where an allele is represented by one band only: whenever two alleles differing 2 repeat units in size were observed (detail of figure 7c), three bands were obtained instead of the expected two bands. For all D7S495 band triplets encountered within a pedigree, the pattern of inheritance was consistent with the hypothesis of the largest and the smallest band being the allele[1].

Interpretation of repeat polymorphism patterns (II): Assignment of arbitrary repeat units

Two alleles at a polymorphic locus are generally believed to be identical if they are of similar size. As the allelic diversity at a polymorphic locus is assumed to result from omission or insertion of one or more repeat units during DNA replication events, the size difference between any two alleles will be an even number of repeat units. Consequently, as long as all alleles that have to be compared in size are compared to the size of the same control allele(s), the natural dimension for sizing alleles at microsatellite loci is "number of repeat units". These can be arbitrarily calibrated by defining the size of a control allele for example as 10 arbitrary repeat units (ARU) (figure 8).

FIGURE 8: ASSIGNMENT OF ARBITRARY REPEAT UNITS (ARU)

(a) Separation of D2S1788 PCR samples in 35 lanes.
(b) Assignment of arbitrary repeat units 12-24 to bands of alleles and by-products of the samples displayed in (a). Desmiling lines are shown at 12 , 19 *and* 22 *arbitrary repeat units.*



1 A speculative explanation for the dependence of observable by-products on the size difference between the alleles typed is based on the observation that isolated D7S495 signals were accompanied by faint smears emerging upwards and downwards from the focussed band. If two alleles are separated by 2 repeat units in size, the "upward smear" of the lower band and the "downward smear" of the upper band would both be of equal size and might get focussed during electrophoresis.

2.1.3 Southern blotting

Southern blotting was performed to confirm a deletion of exon 2 and exon 3. *Eco*RI (NEBiolabs) and *Hind*III (NEBiolabs) were used to digest genomic DNA. In the reaction buffer supplied, 10 µg of DNA and 20 U of restriction enzyme were incubated overnight at 37°C. Restriction fragments were separated on a 1% agarose gel (2 V/cm, 4°C, 16-18 h) and stained with ethidium bromide to control for the loading of similar amounts of DNA in all lanes and the absence of non-specific degradation. The DNA was denatured by incubation in 1.5 M NaCl 0.5 M NaOH for 2 h. Transfer of DNA to a Hybond N+ membrane (Amersham) was carried out overnight with 0.4 M NaOH in a capillary blot device.

The exon 2 and exon 3 probes, encompassing the respective exons and its flanking intron sequences (Zielenski *et al.* 1991), were separately PCR amplified with a high fidelity polymerase (DapGoldstar Polymerase, Eurogentec) in two consecutive rounds: 50 ng of DNA were amplified with a stringent protocol (annealing at 57°C-65°C and $MgCl_2$ 1.5 mM or less) with 25 pmol of upstream and downstream primer (see appendix 8.3.1 on page 171) in a reaction volume of 50 µl. 10 µl of the product were checked by agarose gel electrophoresis for absence of by-products. For labelling with α^{32}PdATP, 5 µl of the PCR product were used as the template for the second amplification in 50 µl with 65°C annealing temperature and 1 mM $MgCl_2$. The probe was labelled in a PCR reaction: 10 pmol of each primer, 1 nmol dGTP, 1 nmol dCTP, 1 nmol dTTP, 25 nmol $MgCl_2$ and 2 µl 10 × reaction buffer provided by the supplier were mixed in a total volume of 8 µl and placed on ice. 6 µl α^{32}PdATP (5000 Ci/mmol, Hartmann Analytic) and 6 µl of PCR product were mixed in a separate tube. Both mixtures were combined, 2 U DapGoldstar Polymerase (Eurogentec) were added and the mixture was immediately placed in a 95°C preheated block of a PCR thermocycler. Labelling of the probe was carried out by PCR amplification with an annealing temperature of 65°C for 30 cycles. Unincorporated nucleotides were removed by ultrafiltration with a Sephadex G-50 (Pharmacia) column. The volume of the filtrate was adjusted to 1500 µl.

The hybridisation concentration of the probe was optimised prior to analysis of patient material on a set of control blots (figure 9). Next, the samples were processed using the optimised probe concentration. Ionic composition of the hybridisation and washing solutions were identical for control blots and sample blots: membranes were equilibrated at 67°C in a shaking water bath using 200 µl/cm^2 of 5 × SSC containing 1% Blocking reagent (Boehringer), 6% SDS and 100 µg/ml DNA from fish sperm (Boehringer). For hybridisation, aliquots from 50µl to 500µl of the purified probe were diluted in a final volume of 1 ml with TE. The probes were denatured by 5' incubation in a boiling water bath and cooled on ice. Each control blot was hybridised in a different hybridisation bag with 200 µl/cm^2 prehybridisation solution containing an aliquot of the probe. Hybridisation was accomplished overnight at 65°C in a shaking water bath. After hybridisation, the membranes were briefly rinsed in 2 × SSC at 65°C (1 ml/cm^2 membrane), washed 5 min in 2 × SSC at 65°C (2 ml/cm^2 membrane) and finally 5 min in 0.2 × SSC at 65°C. The wet membranes were sealed in plastic foil and exposed to X-Omat-AR films (Kodak) for 18 h up to 14 days.

FIGURE 9: SIGNAL TO NOISE RATIO FOR DIFFERENT PCR GENERATED PROBES

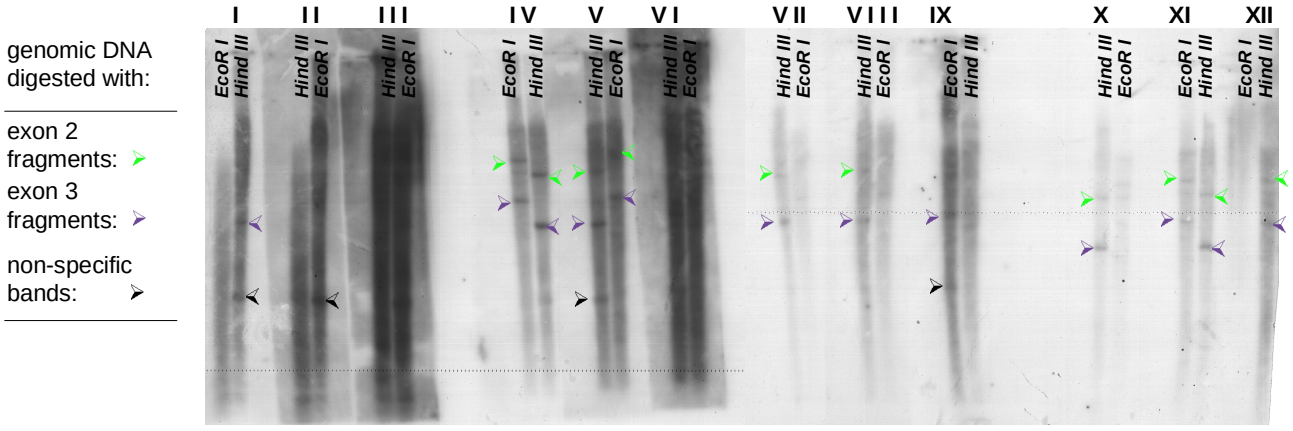
The signal to noise ration was depending on (a) the concentration of the template used for the PCR labelling reaction and on (b) the concentration of the probe used for the hybridisation. Both were optimised empirically on test blots prior to analysis of patient samples.

The probes and hybridisation conditions used for test blot IV shown below were used to perform the hybridisation shown in figure 16, page 62.

Four probes (A1, B1, A2, B2) were labelled by PCR according to the protocol described in the text. A1 and A2 were probes of exon 2, B1 and B2 were probes of exon 3. For A1 and B1, 5µl of a PCR reaction was used as a template for labelling. To generate probes A2 and B2, 5µl of a 1:10 dilution of the same PCR templates were used.

Probe	A1	A2	B1	B2
Template for labelling	PCR of exon 2	PCR of exon 2	PCR of exon 3	PCR of exon 3
Template PCR diluted 1: 10 prior to labelling?	no	yes	no	yes

The probes A1, A2, B1 and B2 were tested on a set of 12 control blots (I-XII). For this experiment, 8 different DNA control samples (named "control DNA 1" to "control DNA 8") were digested with EcoRI or HindIII. 12 hybridisation conditions, differing in (a) probe labelling conditions and (b) hybridisation concentration of the probe were tested on 12 blots. Probes were purified as described in the text and TE buffer was added to a final volume of 1500 µl. The experiments I-XII were hybridised in parallel and washed with the same protocol (see text for details).



Blot	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Volumes of labelled probes used for 10 ml hybridisation solution	50µl A1 50 µl B1 = (1)	100µl A1 100 µl B1 = (2)	500µl A1 500 µl B1 = (3)	50µl A2 50 µl B2 = (4)	100µl A2 100 µl B2 = (5)	500µl A2 500 µl B2 = (6)	1: 10 dilution of (1)	1: 10 dilution of (2)	1: 10 dilution of (3)	1: 10 dilution of (4)	1: 10 dilution of (5)	1: 10 dilution of (6)
Control DNA used for EcoRI and HindIII digest, respectively	Control DNA 1 and Control DNA 2	Control DNA 1 and Control DNA 2	Control DNA 3 and Control DNA 4	Control DNA 3 and Control DNA 4	Control DNA 5 and Control DNA 6	Control DNA 5 and Control DNA 6	Control DNA 7 and Control DNA 8	Control DNA 7 and Control DNA 8	Control DNA 7 and Control DNA 8	Control DNA 7 and Control DNA 8	Control DNA 7 and Control DNA 8	Control DNA 7 and Control DNA 8

➢,➢ : *Assignment of EcoRI and HindII fragments according to the physical map by Rommens et al. (1989). The size was estimated relative to the band pattern of a BstEII-Digest of lambda DNA (data not shown). Two bands are expected within each lane, one corresponding to exon 2 (➢) and one corresponding to exon 3 (➢).*

➢ : *non-specific band*

2.2 Analysis of phenotype

For both studies described in this thesis, a characterisation of the clinical phenotype according to the guidelines established at the Biomed genotype-phenotype workshop in 1994 (*Association Francaise de la Lutte contre la Mucoviscidose 1994: Genotype-Phenotype correlation in CF. AFLM, Paris. pp 4-6*) was attempted. This included assessment of pancreatic status by elastase from stool samples, measured in the Department of Gastroenterology, MHH according to standard laboratory procedures. To explore a bacterial colonisation of the patient's lung, sputum or throat swabs were obtained. Pathogens were analysed by the Department of Microbiology, MHH. Interventions and activities permitted only to medical professionals, such as the anamnesis and clinical investigation of the patients, were performed by the participating physicians of the study.

Parameters used to quantify the disease severity were derived from weight and height of the patient – used to characterise the nutritional status – and forced expiratory volume in 1 s – used to characterise the pulmonary status – by employing age corrected percentiles (chapter 2.2.1).

The basic defect, as defined by a defect in Cl^- conductance in CFTR expressing epithelia, was investigated by a sweat test (carried out according to the procedure used at the hospital where the research was performed), by nasal potential difference measurement (chapter 2.2.2.1) and by intestinal current measurement (chapter 2.2.2.2).

2.2.1 Evaluation of disease severity and selection of pairs with extreme phenotypes

For CF twin and sibling pairs, information on age, sex, weight, height, the lung function parameters forced expiratory volume in 1 s [FEV1] and vital capacity, pancreatic status and colonisation with *Pseudomonas aeruginosa* was required. Out of a total of 540 pairs recruited from more than 200 different European CF centers, complete information could be obtained for 324 pairs.

2.2.1.1 Selection of twin and sib pairs

The quantification of disease severity and overall clinical status was assessed using an unbiased mathematical approach. The aim of this procedure was to select pairs with an extreme phenotype. If this selection is not made by an objective approach but by manual evaluation, the pairs selected will express extreme phenotypes, but the selection will not be complete – simply because it is most likely that a pair displaying features similar to a manually selected pair will escape notice when reviewing clinical data of more than 300 patient couples. Consequently, evaluation done on manually selected pairs will result in describing merely a subgroup of all pairs with likewise features.

On the other hand, the computed approach demands the use of simplified data. As will be illustrated in chapter 2.2.1.2, the disease severity was quantified by equally weighing nutritional status and pulmonary status, achieved by a diagram with a parameter representing each status on one axis. The overall disease severity was then defined as distance from origin within that diagram, ignoring the information of direction that is contained within a vector. This was compensated for as

described in chapter 2.2.1.3. The result of the ranking procedure applied was the description of a pair's phenotype with five rank numbers, each taking into account disease severity and intrapair discordance. Selection of twin and sib pairs – for investigation of the basic defect and for candidate gene testing – was carried out based on these computed rank numbers and revealed pairs possessing non-average properties for their nutritional status, their lung function or the intrapair variation of one or both of these characters.

For evaluation of candidate genes, the selected pairs were categorised to define the most extreme ΔF508 homozygous pairs in the sequences with respect to the trait *"discordant"* or the combined traits *"concordant/mild disease"* and *"concordant/severe disease"*. Pairs ranked simultaneously in two conflicting categories were categorised separately as "*discordant and concordant mild disease*" (DC) and "*non-discordant*" (ND) as described in chapter 2.2.1.4.

2.2.1.2 Definition of age independent parameters used to describe disease severity

Patients and pairs, who were heterogeneous in both age and age difference, were compared using centiles. This allowed an age-independent description of disease severity and consequently enabled comparison of disease severity between patients of different ages, within as well as between pairs. The nutritional status was represented by weight expressed as a percentage of expected weight at the patient's height percentile (wfh%). wfh% was calculated on the basis of age and sex dependent centiles for weight and height published by Prader *et al.* (1989). Data for the complete patient cohort are shown in figure 10b. Lung function was evaluated from predicted values for FEV1 as described by Knudson *et al.* (1983) and converted to CF population centiles (FEVPerc). FEVPerc was computed from FEV1%pred on the data basis of the European CF registry report of 1995 which compiles lung function data from more than 20.000 CF patients. Figure 10a displays all data of the CF patients lung function parameter FEV1%pred whereby the measured FEV1 value is normalised for height, sex and age in a healthy non-CF population. Figure 10b shows the distribution of parameter FEVPerc that corrects for the CF specific decline in FEV1%pred with age. Regression formulas used for the calculation of wfh% and FEVPerc are listed in the appendix 8.3.3.

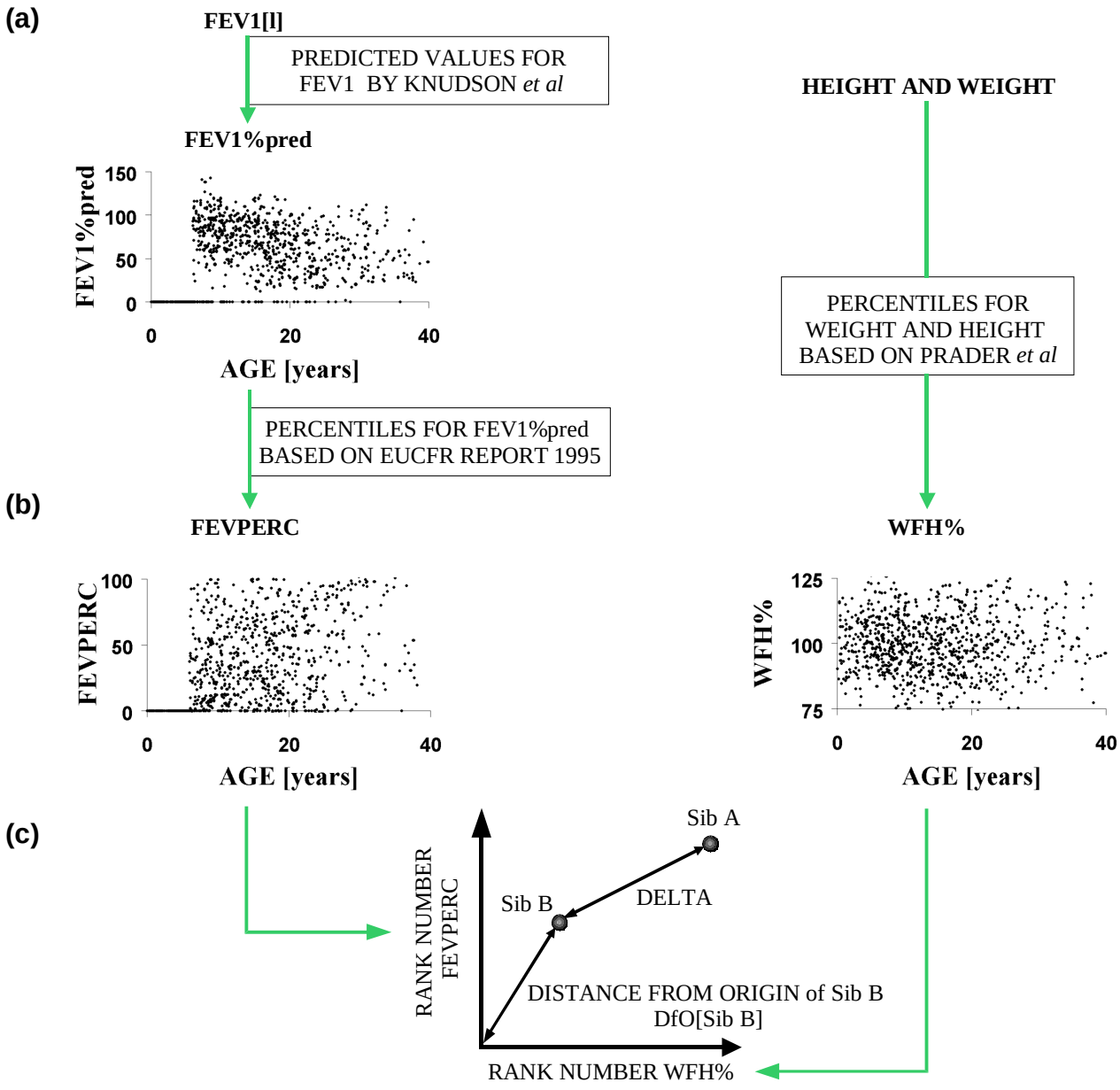
2.2.1.3 Ranking of pairs according to disease severity and discordance

The selection of twin and sib pairs with extreme phenotypes was based on the combination of nutritional and pulmonary status, represented by wfh% and FEVPerc, as defined in chapter 2.2.1.2. Disease severity was defined as "overall clinical shape" with equal weight on nutritional and pulmonary status. Nutritional status and pulmonary function were described non-parametrically by assigning ranking numbers for wfh% and FEVPerc within a group of 324 patient pairs. Both parameters were used to display a diagram with rank numbers of wfh% on the x-axis and rank numbers of FEVPerc on the y-axis (figure 10c). Each patient, represented as a point in that diagram, was then characterised in terms of disease severity by the distance from origin (DfO) in the plot. Low DfO corresponded to severe disease and high DfO corresponded to mild disease. Likewise, intrapair

distance was quantified by calculating the distance of the two data points representing two members of a pair within the same diagram. This parameter was named DELTA, with low values for DELTA corresponding to concordant pairs and high values of DELTA corresponding to discordant pairs (figure 10c).

FIGURE 10: RANKING OF PATIENT PAIRS (I)
DEFINITION OF COMPOSITE PARAMETERS

The disease severity was quantified by equally weighing nutritional status and pulmonary status:
(a) raw data obtained from participating CF centers
(b) parameters with age independent centiles representing nutritional status and lung function
(c) definition of composite parameters DELTA (➔ intrapair discordance) and DfO (➔ disease severity)



Alternatively, the intrapair discordance was judged by the difference in DfO for both members of a pair. Two definitions for the intrapair discordance were employed because some pairs, being concordant and in good clinical status in their weight for height but discordant in their lung function, were scored as discordant when using DELTA. Discordance within a pair implies one sib being mildly diseased while the other is severely affected. In contrast, these pairs with high values for DELTA showed equally high DfO for sib A and sib B, demonstrating that both sibs are in good health. These pairs were scored as concordant when the difference between the distances from origin was chosen to define the intrapair discordance[2].

When plotting the parameter DfO on the y-axis against the parameter DELTA on the x-axis for all patients (figure 11a), pairs could be ranked according to their degree of discordance and their degree of disease severity by successively tracing pairs from the lower left edge (concordant severe pairs) or the upper left edge (concordant mild pairs) or the right rim (discordant pairs) of the diagram towards the middle of the plot where both, disease severity and degree of discordance, were average. However, this manual assignment was only reliable when ranking pairs according to their discordance. For the concordant – mild or severe – pairs, the emphasis can be put either on "disease severity" or on "discordance". Hence, both attributes of a pair were combined in a diagram with a parameter describing the discordance on the x-axis and a parameter describing disease severity on the y-axis. The rank of concordant pairs was derived from their distance from origin in that plot. With this definition, the impact of concordance and the disease severity were identical for the definition of the combined properties "concordant / mild disease" and "concordant / severe disease".

To sort all pairs according to the traits "discordant", "concordant / mild disease" and "concordant / severe disease", a total of five rank numbers were assigned to each pair (figure 11b-f). The rank defining the position of a pair according to its discordance (DISC) was the reverse rank number of the parameter DELTA, with DISC = 1 being assigned to the most discordant pair (figure 11b). The rank number defining the position of a pair according to its concordance and its disease mildness (CON+Δ) was calculated as rank number for the distance from origin in the plot of the rank number for DELTA on the x-axis and the inverse rank number for the sum of the DfO for both sibs on the y-axis. CON+Δ = 1 is referring to the most concordant and most mild pair (figure 11c). Likewise, the rank defining the position of a pair according to its concordance and its disease severity (CON-Δ) was calculated as rank number for the distance from origin in the plot of the rank number for DELTA on the x-axis and the rank number for the sum of the DfO on the y-axis. CON-Δ = 1 is referring to the most concordant and most severe pair (figure 11d). While CON+Δ and CON-Δ both employ the parameter DELTA for quantifying the discordance, in an alternative

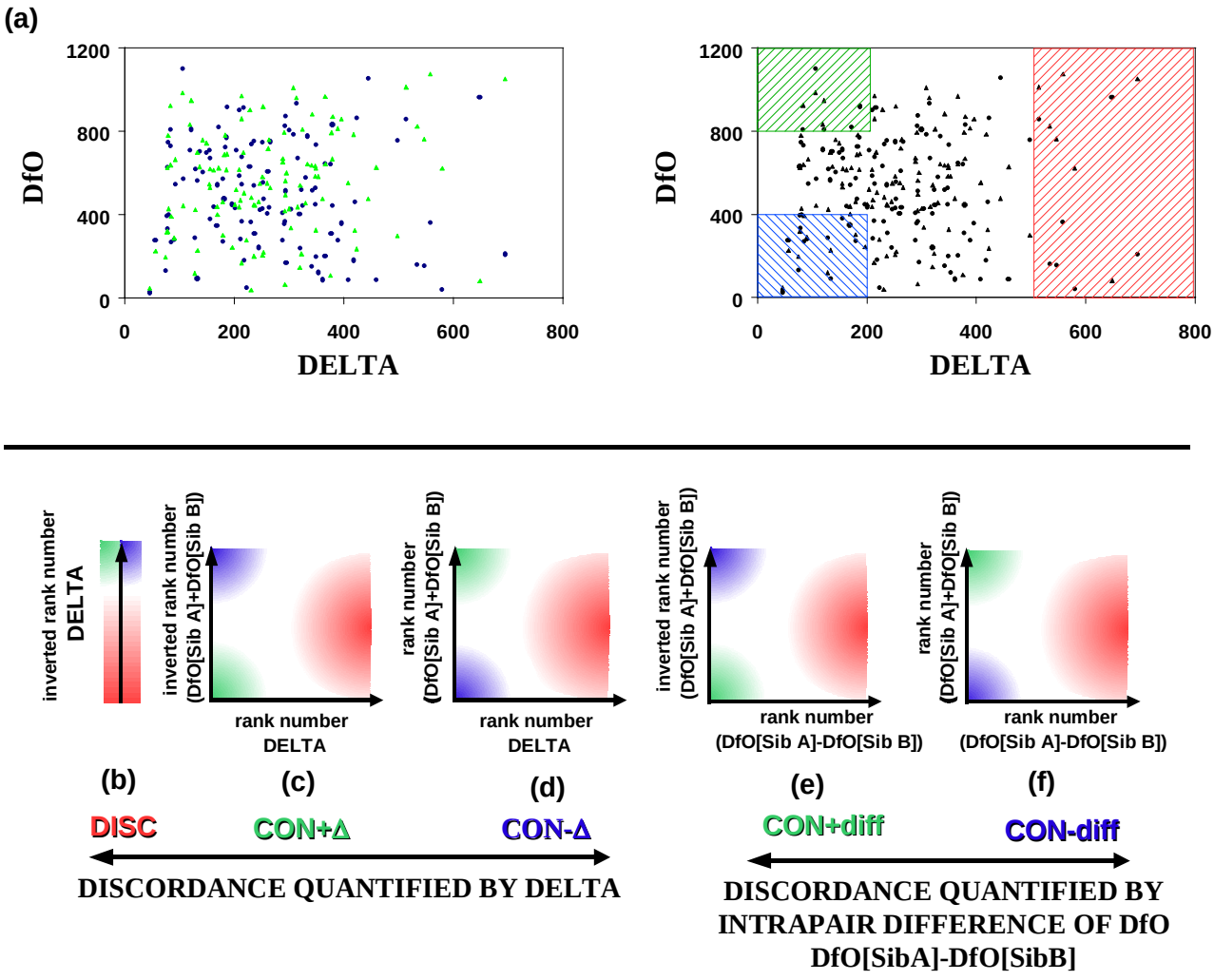
2 An example for such a pair is given by the following data set for an adult sib pair: sib A had a wfh of 120 %, corresponding to a rank of 578 for the nutritional status (maximal rank 648 within the evaluated 324 pair cohort). The lung function of Sib A revealed an FEV1%pred of 17% resulting in a rank number of 12 for the pulmonary status. The overall disease severity based on DfO was ranked as 502. Sib B showed a wfh of 120 %, ranking 582 in the nutritional status. The lung function of Sib B was 91%pred in FEV1, resulting in a rank of 526 for the pulmonary status. The overall disease severity evaluated via DfO was ranked 558.
This pair was categorised as "DC", as described in figure 12.

approach the difference in DfO for both members of a pair was used to define discordance in otherwise similar plots as described for CON+Δ and CON-Δ. Rank numbers derived from differences in DfO were called CON+diff and CON-diff with CON+diff = 1 corresponding to the most concordant and most mildly diseased pair (figure 11e) and CON-diff = 1 corresponding to the most concordant and most severely diseased pair (figure 11f).

FIGURE 11: RANKING OF PATIENT PAIRS (II)
DEFINITION OF RANK NUMBERS

(a) plot of disease severity vs discordance for all ΔF508 homozygous pairs
left: ● Sibling A of pair; ▲ Sibling B of pair
right: areas representing discordant pairs (■), concordant/mild disease pairs (■) concordant severe disease pairs (■)

(b)-(f): diagrams used to define rank numbers by ranking a pair's distance from origin in these plots
(b) discordance DISC *(discordance defined by DELTA)*
(c) concordance/mild disease CON+Δ *(discordance quantified by DELTA)*
(d) concordance/severe disease CON-Δ *(discordance quantified by DELTA)*
(e) concordance/mild disease CON+diff *(discordance quantified by difference between DfO[SibA]and DfO[SibB]*
(f) concordance/severe disease CON-diff *(discordance quantified by difference between DfO[SibA]and DfO[SibB]*



2.2.1.4 Categorising of pairs selected for candidate gene typing

Based on the ranking described in chapter 2.2.1.3, all ΔF508 homozygous pairs with rank numbers lower than 20 in any of the sequences CON+Δ, CON-Δ, CON+diff or CON-diff or lower than 25 in the sequence DISC were selected for PCR typing if DNA samples were available by December 1997. A total of 38 ΔF508 homozygous pairs could be selected. The rank numbers for the selected ΔF508 homozygous pairs are shown in table 1.

For ranking pairs, disease severity was defined by DfO with nutritional and pulmonary status balanced. As a combined parameter, DfO does not contain the information whether it was based on a more than average wfh%, a more than average FEVPerc, or both. Likewise, the parameter DELTA, defining the discordance of a pair, can be large even if both members of a patient pair have a equally high DfO as for pairs concordant in one parameter – e.g. wfh% – but highly discordant for the other parameter – e.g. FEVPerc. Furthermore, in defining rank numbers for concordant pairs, disease severity and discordance was given equal importance. Consequently, pairs being extremely concordant received comparably low rank numbers in CON+Δ and CON-Δ.

To assure that all pairs grouped within the same category are comparable with respect to both the intrapair difference of the wfh% and FEVPerc, the five rank numbers were combined to identify a category as displayed in figure 37:

- All pairs categorised as ***dis****cordant* pairs (**DIS**, 12 couples) had a low rank number for DISC, but not for CON+Δ, CON-Δ, CON+diff and CON-diff.
- 4 pairs selected based on the rank number for CON+diff also displayed a low rank number for DISC and were categorised as ***d****iscordant and* ***c****oncordant/mild disease* (**DC,** 4 couples).
- All pairs categorised as ***con****cordant mild* (**CON+**, 8 couples) had low rank numbers for CON+Δ and CON+diff but not for CON-Δ, CON-diff and DISC.
- All pairs categorised as ***con****cordant severe* (**CON-**, 8 couples) had low rank numbers for CON-Δ and CON-diff, but high rank numbers for CON+, CON+diff and DISC.
- 6 pairs had low rank numbers for CON+Δ, CON-Δ, CON+diff and CON-diff but a high rank number for DISC and were categorised as ***n****on-****d****iscordant* (**ND**, 6 couples)[3].

The similarity of intrapair difference for wfh% and FEVPerc within each of the resulting categories is demonstrated in figure 28. Diagrams displaying DELTA, DfO, wfh% and FEVPerc for all pairs within the categories DIS, CON+ and CON- are shown in figure 29 (Chapter 3.2.3.2, page 99 and page 100).

3 To discriminate pairs "CON+" and "CON-"from pairs "ND", the following borders were employed:

CON+ vs ND: {Rank CON+Δ — Rank CON-Δ} < 30 OR {Rank Con+diff — Rank Con-diff} < 30 ⇒ pair classified as "CON+"

CON- vs ND: {Rank CON+Δ — Rank CON-Δ} < 20 OR {Rank Con+diff — Rank Con-diff} < 20 ⇒ pair classified as "CON-"

whereby

{CON+Δ — Rank CON-Δ} was used if the pair was selected based on Rank CON+ Δ or Rank CON-Δ

and

{Rank Con+diff — Rank Con-diff} was used if the pair was selected based on rank CON +diff or CON-diff.

The borders employed to distinguish concordant mild or severe pairs from non-discordant pairs by the rank number differences, {Rank CON+Δ — Rank CON-Δ} and {Rank Con+diff — Rank Con-diff}, were set to 10, 20, 30 and 40. Other borders than those used in the definition above resulted either in omission of most pairs (e.g. border 40) or no pair (e.g. border 10).

TABLE 1: PAIRS SELECTED FOR GENETIC ANALYSIS (I)
RANK NUMBERS AND CATEGORIES

38 ΔF508 homozygous pairs were selected for genetic analysis (DNA available by December 1997) based on a rank number of 20 or lower (CON+Δ, CON+diff, CON-Δ, CON-diff) or 25 or lower (DISC). Rank numbers leading to the selection of a pair are printed in colour.

Categories DIS, CON+, CON-, DC and ND were defined as shown in figure 12.

For discriminating pairs "CON+" and "CON-" from pairs "ND" see text.

DISC	CON+ Δ	CON+ diff	CON- Δ	CON- diff

DISC	CON+ Δ	CON+ diff	CON- Δ	CON- diff

	LOW	HIGH	HIGH	HIGH	HIGH
DIS3	3	114	114	99	85
DIS5	5	110	106	103	100
DIS6	6	106	105	104	104
DIS8	8	102	88	105	87
DIS9	9	113	112	91	82
DIS14	14	112	113	79	75
DIS15	15	78	80	97	103
DIS16	16	85	91	96	97
DIS18	18	80	44	95	50
DIS19	19	81	83	92	94
DIS21	21	98	98	75	77
DIS24	24	103	109	62	67

	HIGH	LOW	LOW	HIGH	HIGH
CON+2	106	2	7	78	81
CON+5	99	5	8	65	68
CON+6	105	6	13	38	44
CON+7	78	7	24	82	93
CON+10	84	10	25	57	69
CON+15	74	15	33	77	96
CON+16	68	16	19	90	89
CON+19	79	19	17	51	62

	HIGH	HIGH	HIGH	LOW	LOW
CON-2	113	68	73	2	4
CON-5	103	65	66	5	2
CON-7	104	59	68	7	11
CON-8	97	79	89	8	15
CON-10	107	37	40	10	6
CON-12	86	50	53	12	10
CON-14	77	67	75	14	27
CON-25	76	45	35	25	8

	LOW	HIGH	LOW	HIGH	LOW
DC(1)	7	70	10	114	99
DC(2)	17	51	2	111	70
DC(3)	38	33	1	110	83
DC(4)	31	46	15	102	59

	HIGH	LOW	LOW	LOW	LOW
ND(1)	95	17	21	35	41
ND(2)	85	39	47	19	19
ND(3)	110	13	14	28	28
ND(4)	102	20	26	26	34
ND(5)	96	27	49	17	38
ND(6)	93	11	12	40	40

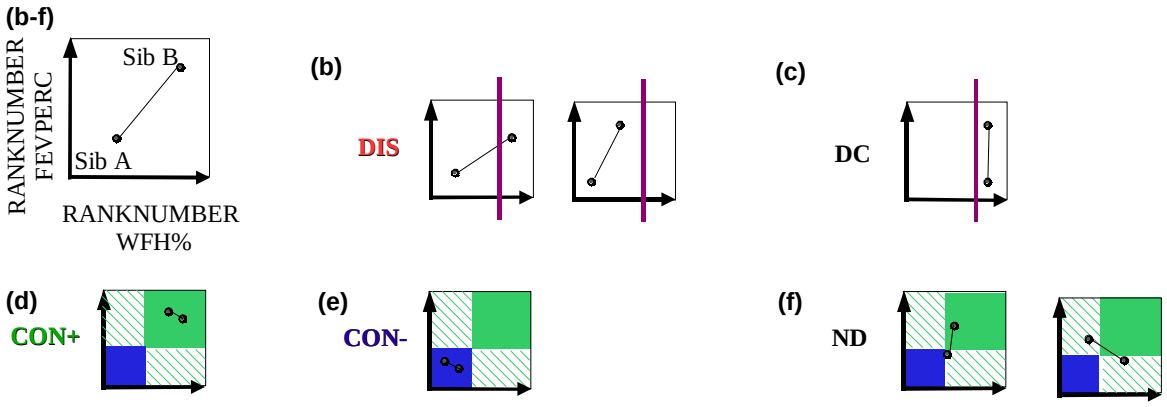
FIGURE 12: RANKING OF PATIENT PAIRS (III)
CATEGORISING OF SELECTED PAIRS

(a) rank number patterns defining the categories DIS (discordant), DC (discordant and concordant/mild disease), CON+ (concordant mild disease), CON- (concordant/severe disease), ND (non-discordant)

(b)-(f): examples of pairs within each category represented by plots of ranknumber of wfh% vs ranknumber of FEVPerc. This diagram was used to define the parameters DfO and DELTA representing the disease severity and the intrapair-discordance (figure 10c).

(a)

	RANK NUMBERS AS DEFINED IN FIGURE 11				
	DISC	**CON+Δ**	**CON+diff**	**CON-Δ**	**CON-diff**
DIS	LOW	HIGH	HIGH	HIGH	HIGH
DC	**LOW**	HIGH	**LOW**	HIGH	HIGH
CON+	HIGH	LOW	LOW	HIGH	HIGH
CON-	HIGH	HIGH	HIGH	LOW	LOW
ND	HIGH	**LOW**	**LOW**	**LOW**	**LOW**



2.2.2 Assessment of the basic defect

The impaired Cl^- permeability, regarded as the basic defect in CF, was assessed in three different tissues: in the sweat gland (by measuring Cl^- concentrations of the sweat in a pilocarpine iontophoresis sweat test according to Gibson and Cooke 1959), in the respiratory tract (by measuring the permeability of the epithelium of the lower nasal turbinate, chapter 2.2.2.1) and in the intestinal tract (by measuring the permeability of rectal suction biopsies, chapter 2.2.2.2). Gibson-Cooke-sweat test were performed by the technician of the pediatric clinic of the MHH according to approved guidelines (National committee for clinical laboratory standards [1994] Sweat testing: sample collection and quantitative analysis - approved guideline. Document C34-A. Wayne, Pennsylvania.). The expertise and the equipment to measure NPD and ICM was provided by the Rotterdam group. All NPD measurements were performed by I. Bronsveld, Rotterdam. The results discussed within this thesis refer to the study "*rare genotypes*". ICM measurements were done by H. Veeze, Rotterdam for all cases in the study "*rare genotypes*" and several cases of the "*European CF*

twin and sibling study". ICM results from both studies are presented and discussed within this thesis. Quantitative evaluation of ICM tracings was accomplished computer assisted, employing INTEG, programmed by J. Greipel, Hannover.

2.2.2.1 Nasal potential difference (NPD) measurement

Nasal potential difference measurements were performed with a protocol adapted from Knowles *et al.* (1981). The reference electrode was connected to a saline-filled needle which was placed submucosally in the forearm. The epithelium of the lower nasal turbinate was accessed by a catheter used to superfuse the nasal epithelium with different solutions. Both electrodes were connected to the voltage measuring device via a salt bridge and Ag/AgCl electrodes. The nasal epithelium was superfused subsequently with the solutions displayed in table 2.

TABLE 2: NPD PROTOCOL

(1) **Basal**	1.2×10^{-1} M NaCl 2.5×10^{-2} M NaGluconate 4×10^{-4} M NaH_2PO_4 2.4×10^{-3} M Na_2HPO_4	pH 7.4	basal value
(2) **Amiloride**	(1) with 10^{-4} M amiloride		blocking of Na^+ channels
(3) **Chloride free**	1.45×10^{-1} M NaGluconate 4×10^{-4} M NaH_2PO_4 2.4×10^{-3} M Na_2HPO_4 10^{-4} M amiloride	pH 7.4	block of anion permeability by gluconate
(4) **Isoproterenol**	10^{-4} M isoprenaline in (3)		stimulation of CFTR Cl^- channels

NPD is employed as a sensitive method for diagnosing CF. The condition CF, characterised by a reduced Cl^- permeability of the nasal epithelium, was identified by a high[4] basal value compared to non-CF controls (CF: PD = -39mV ± 14mV, non-CF: PD = -20mV ± 9mV). In borderline cases of basal PD values between -25 mV and –35 mV, the spontaneous Cl^- conductance assessed by superfusion of the nasal epithelium with chloride free solution after blocking of Na+ channels with amiloride was employed to discriminate between CF and non-CF condition: a large hyperpolarization ($\Delta PD \geq 16$ mV) was considered confirmatory for the non-CF status; CF patients exhibiting low basal values showed a smaller response ($\Delta PD \leq 7$ mV; Bronsveld *et al.* 1996). The protocol allows the assessment of residual chloride secretion in CF patients by measuring the responses to chloride-free solution and isoproterenol. Within chapter 3.1.2, the conditions "normal", "CF" and "CF-res" are distinguished accordingly.

4 A "high" basal refers to a negative potential difference with high absolute value by convention. -50mV is termed a high basal PD compared to –20 mV.

2.2.2.2 Intestinal current measurement (ICM)

Short circuit currents were measured in a micro Ussing chamber as described by Veeze *et al.* (1991). Two samples were analysed in parallel. Freshly obtained rectal biopsies were mounted within 5 minutes in the Ussing chamber. Tissue was incubated in Meyler solution at 37°C and gassed with carbogen (95% O_2, 5% CO_2). After stabilisation of the basal current, the protocol shown in table 3 was implemented.

TABLE 3: ICM PROTOCOL
Secretagogues (2)-(6) were added to the mucosal (M) or serosal (S) side.

(1) **glucose**	10^{-2} M	M+S	10'-20'	stabilisation of basal current
(2) **amiloride**	10^{-4} M	M	5'	blocking of Na^+ channels
(3) **indomethacin**	10^{-5} M	M+S	5'	blocking of prostaglandin synthesis
(4) **carbachol**	10^{-4} M	S	5'-30'	raising intracellular Ca^{2+} concentration: † transient activation of conductance, e.g., through basolateral K^+-channels and apical Cl^- - channels
(5) **cAMP/forskolin**	10^{-3} M/10^{-5} M	M+S	5'-30'	raising intracellular cAMP
(6) **DIDS**	2×10^{-4} M	M	10'	e.g. blocking of Ca^{2+} activated Cl^- channels ‡
(7) **histamine**	5×10^{-4} M	S	5'-30'	raising intracellular Ca^{2+} concentration: † transient activation of conductance, e.g., through basolateral K^+-channels and apical Cl^- channels

Alternatively, DIDS was added prior to carbachol when two working biopsies were obtained as judged by the amiloride response.
‡ DIDS acts on multiple components of the ion transport system (Gögelein 1988)
† carbachol and histamine act through different, though both Ca^{2+} dependent, pathways

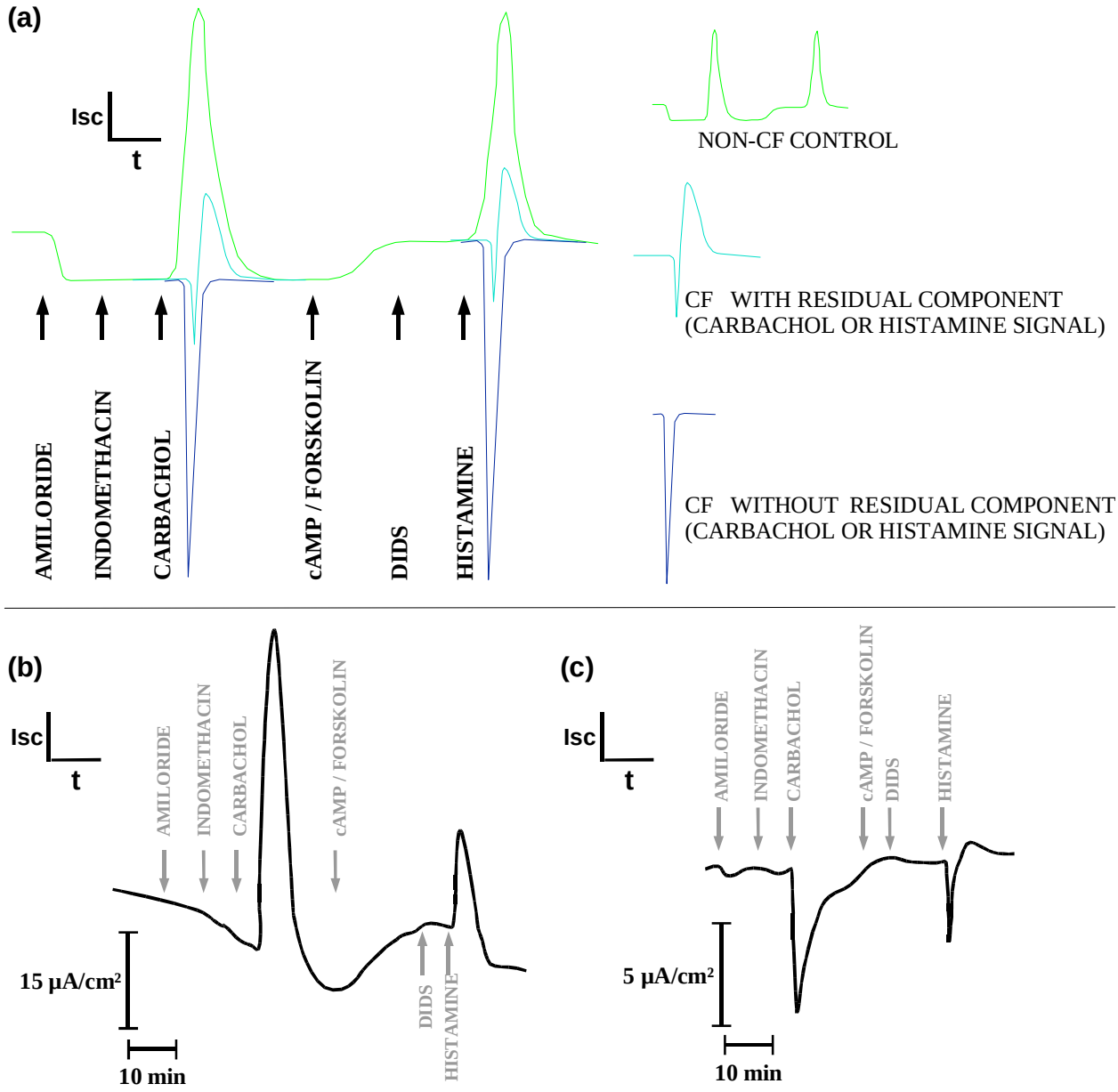
The employment of ICM as a diagnostic tool in CF was based on the response to carbachol (Veeze *et al.* 1991). In non-CF tissue, the CFTR mediated Cl^- secretory current provoked a transient increase in short circuit current while in most CF tissues a transient decrease in short circuit current was observed due to the efflux of K^+ ions through K^+ channels. In some CF patients a response in the non-CF Cl^- secretory direction was seen but the size of the responses differentiates between non-CF (mean peak response 26 µA/cm^2) and CF (mean peak response 3 µA/cm^2) conditions. Residual Cl^- secretion was also observed by CF patients exhibiting both the K^+ mediated reversed and the Cl^- mediated residual response.

The protocol permits a dissection of residual Cl^- secretion in proportions mediated by CFTR and mediated by alternative, Ca^{2+} activated Cl^- channels by the employment of DIDS which is known to inhibit the Ca^{2+} activated secretory currents (reviewed by Gögelein 1988). In order to assess the presence of alternative secretion, the residual component of responses obtained with carbachol (in the absence of DIDS) and histamine (after 10' incubation with DIDS) were compared. Finally, CFTR mediated chloride conductance was considered to be activated by addition of cAMP and forskolin.

Representative tracings for the CF and the non-CF condition are displayed in figure 13.

FIGURE 13: ICM RESULTS TYPICAL FOR NON-CF AND CF PATIENTS

(a) Idealised ICM results for non-CF controls and CF patients with protocol shown in table 3.
(b) Example for a non-CF ICM. The result was obtained on patient O (non-CF, described in chapter 3.1.2.5).
(c) Example for a CF ICM. The result was obtained on patient D (CF, described in chapter 3.1.2.1).
(b) and (c): scales for Isc and t are adjusted for better viewing.



Carbachol-evoked responses with substantial proportions of a residual Cl^- secretory component have been shown to correlate with a mild CF phenotype in a group of unrelated CF patients with several *CFTR* mutation genotypes (Veeze *et al.* 1994). Within this thesis, two approaches towards the evaluation of ICM results were pursued: in chapter 2.2.2.2.1 the responses obtained by carbachol-stimulation were analysed qualitatively based on the shape of the response in order to evaluate the results of the study "*rare genotypes and atypical cystic fibrosis*" (chapter 3.1.2). A quantification of

the residual component was attempted as described in chapter 2.2.2.2.2. The results gained on patients participating in the "*European CF twin and sibling study*" were evaluated quantitatively as shown in chapter 3.2.2.

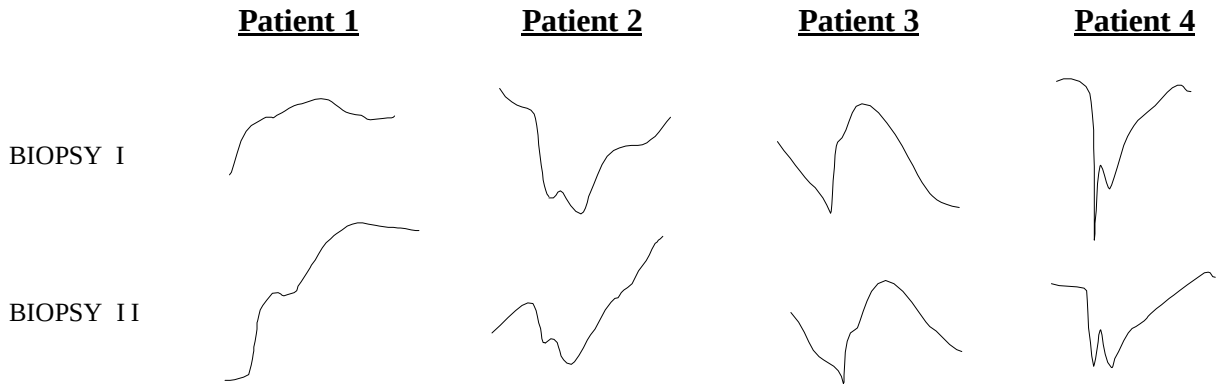
2.2.2.2.1 Shapes of carbachol-induced short circuit currents in rectal biopsies

Ideally, the baseline is supposed to be constant during a response. However, when studying rectal biopsies by ICM, most responses displayed either a change of the baseline level during the response (defined as offset) or a change in the baseline slope during the response (defined as drift). The nine shapes of carbachol- and histamine responses observed on analysing 66 tracings are displayed in table 4. A clear distinction between presence and absence of residual chloride secretion is feasible for five out of nine shape types observed (type I, II: without residual component; type VII, VIII, IX: with residual component). These unequivocally interpretable shapes were assigned to 22 out of 43 carbachol-induced peaks and for 26 out of 39 histamine-induced peaks. In some cases, results for two biopsies were obtained and different shape types were displayed by both biopsies. Especially type III, characterised by its broadness , was found in combination with other shapes displaying a residual component (type VII and VIII, respectively). Type III was observed only for carbachol- but not for histamine-induced responses.

Figure 14 displays some uncommon response shapes. Each of the four patterns was seen in only one patient, but was reproduced in both biopsies investigated.

FIGURE 14: UNUSUAL, BUT REPRODUCIBLE RESPONSE SHAPES IN ICM

Selected examples for unusually shaped ICM responses. Each of the four pairs of responses was obtained by measuring two biopsies from the same patients. All results were obtained by stimulation with carbachol.



2.2.2.2.2 Quantitative evaluation of ICM signals

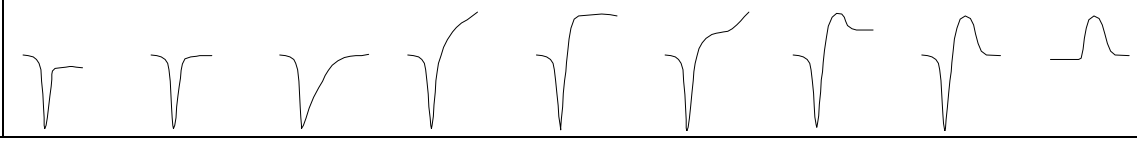
In order to evaluate all carbachol– and histamine-induced responses regardless of their shape, a computer-assisted approach was taken: for each peak, a linear baseline was assigned prior to the response and – independently of the former – following the response. The course of the baseline during the response was calculated as a spline function, defined by a third degree polynomial

equation starting with slope and baseline of the linear function defining the baseline prior to the response and ending with slope and baseline of the linear function defining the baseline after the response. After the spline function has been assigned as baseline to the response, height and area above and below the baseline were calculated and considered representative for the reversed and the residual proportions of the signal (data shown in figure 15, page 59).

TABLE 4: COMMON SHAPES OF CARBACHOL AND HISTAMINE INDUCED PEAKS

*Definition of shape types I to IX by offset and drift conditions. Offset: change of baseline level during response, either towards the reversed direction (*to rev*) or towards the residual direction (*to res*). Drift: change of baseline slope during response. Shape* III *is discriminated against shape* II *on the basis of the peak broadness.*

reversed	yes	yes	yes	yes	yes	yes	yes	yes	no
residual	no	no	unknown	unknown	unknown	unknown	yes	yes	yes
offset	to rev	no	no	no	to res	to res	to res	no	no
drift	no	no	no	yes	no	yes	no	no	no
SHAPE	I	II	III	IV	V	VI	VII	VIII	IX



CARBACHOL-INDUCED SIGNALS

Shapes of carbachol-induced signals observed in ICM tracings of 66 patients. In 7 cases, no conclusive signal was seen. 10 results were excluded from shape type analysis because of a high drift or too small response. Of the remaining 49 patients, 43 results were obtained from either one biopsy only or both measurements showed the same carbachol-induced shape. The number of observed shapes I-IX are displayed in the table below. In 6 cases, different shapes were observed in both biopsies. Of these 6 cases, three displayed a combination of shape III in one measurement and shape VIII in the other measurement. The following combinations of shapes were observed only once: shape III and shape VII, shape VI and shape VIII, shape V and shape VIII.

SHAPE	I	II	III	IV	V	VI	VII	VIII	IX
no. of tracings (n=43)	2	4	6	6	4	5	5	6	5

HISTAMINE-INDUCED SIGNALS

Shapes of histamine-induced signals observed in ICM tracings of 66 patients. In 19 cases, no conclusive signal was seen. 3 results were excluded from shape type analysis because of a high drift or too small response. Of the remaining 44 patients, 39 results were obtained from either one biopsy only or both measurements showed the same histamine-induced shape. The number of observed shapes I-IX are displayed in the table below. In 5 cases, different shapes were observed in both biopsies. Of these 5 cases, three displayed a combination of shape II in one measurement and shape VII in the other measurement. The following combinations of shapes were observed only once: shape V and shape VI, shape II and shape VI.

SHAPE	I	II	III	IV	V	VI	VII	VIII	IX
no of tracings (n=39)	1	3	—	2	7	3	9	9	4

FIGURE 15: RELATION OF PEAK AREA AND PEAK HEIGHT OF CARBACHOL –- AND HISTAMINE –- INDUCED SIGNALS IN ICM

All data displayed were obtained using baselines calculated by INTEG as described in chapter 2.2.2.2.2. For each response, four values were calculated: peak height below the baseline (parameter a), peak area below the baseline (parameter A), peak height above the baseline (parameter b) and peak area above the baseline (parameter B); see grey box for definition. Units are given as divisions on scale for short circuit current (Isc) and time (t).

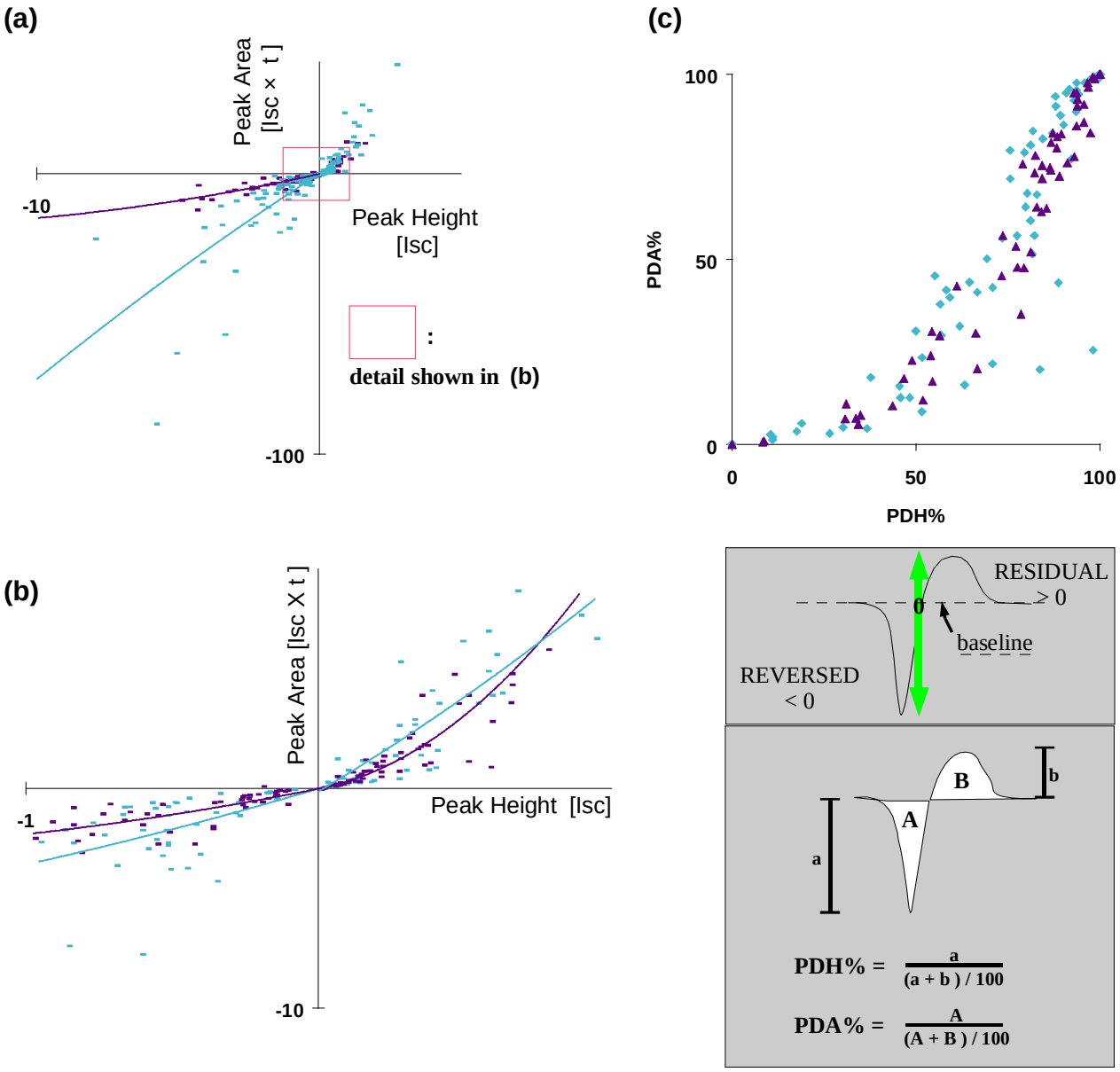
(a), (b) *Values below zero refer to the reversed proportion of a response (height* a, *area* A*) and values above zero refer to the residual proportion of a response (height* b, *area* B*).*

-: carbachol-induced responses *- : histamine-induced responses*

——— : average height to area ratio for carbachol-induced peaks (2 nd degree polynomial fit)

——— : average height to area ratio for histamine-induced peaks (2 nd degree polynomial fit)

(c) *Proportion of the reversed component of area (PDA%) and height (PDH%). Peaks displaying exclusively the reversed component have values of 100% PDH and 100% PDA.*



3 RESULTS AND DISCUSSION

3.1 The study "Rare genotypes and atypical Cystic Fibrosis"

The study *"Rare genotypes and atypical CF"* aims at correlating *CFTR* genotype and CF phenotype by investigating the effect of uncommon molecular lesions in the *CFTR* gene on the basic defect and the clinical phenotype. The patients selected for this study can be divided into three groups:

(a) *compound heterozygous patients carrying a rare CFTR allele*

nonsense mutations: G542X/Y1092X; R553X/L1059X

splice mutations: ΔF508/3272-26 A⇨G; ΔF508/3850-3 T⇨G; 4000-2 A⇨G/IVS8-6(5T)

complex allele: ΔF508-R553Q/R553X; ΔF508-V1212I/ΔF508

(b) *patients with homozygous CFTR genotypes other than ΔF508*

R553X, CFTRdelEx2[(a)]; CFTRdelEx2,3[(b)]; E92K; 1898+3 A⇨G; 3849+10kb C⇨T

(c) *patients with atypical CF suspected to carry at least one functional CFTR allele*

patient P, Q, R, S (table 5)

All genotypes but those indexed by footnote and described in chapter 3.1.1 were resolved previously. The states of CFTR mutation screening of the participants of unknown mutation genotype at the day of entry are listed in table 5.

TABLE 5: SSCP SCREENING CARRIED OUT UNTIL 1994 ON PATIENTS PARTICIPATING IN THE STUDY *"RARE GENOTYPES AND ATYPICAL CF"*

D, I, P, Q, R, S: patient identification *S: exon was sequenced after SSCP analysis*

1: SSCP analysis performed with one restriction digest; 2: SSCP analysis performed with two restriction digests

SSCP Screening of CFTR Exons

	1	2	3	4	5	6a	6b	7	8	9	10	11	12	13	14a	14b	15	16	17a	17b	18	19	20	21	22	23	24
D				1				1		S	S	S	2	1			2		1	1		S	1	S	1		
I				2	1			2	1	2	S	S	2	1	2				1	2		S	1	S	1		
P	2	1	S	2	1	2	2	2	1	2	S	2	2	2	2			2	1	2		1	1	1	1	2	
Q	2	S	1	2	1	2	2	2	1	1	S	2	1	1	2		2	2	1	1	1	1	2	2	2	1	
R	2	1	S	2	1	2	2	2	1	S	1	S	2	1	2		2	2	1	1	1	S	S	2	2	2	
S	2	S	S	2	1	1	S	1		S	S	S	2	1	2		2	2	2	2	1	S	S	S	1	1	

These patients suspected of carrying at least one functional *CFTR* allele were selected on the basis of the work of T. Dörk who has analysed more than 350 CF patients without being able to identify both disease-causing lesions in some cases. Within this thesis, SSCP screening of all *CFTR* exons and the promoter region was completed for these 6 cases and another 14 patients who were diagnosed as having CF but did not carry two mutations, as was known after mutation screening of most CFTR coding sequences and flanking intron sites until 1994.

(a) This mutation was identified in patient I (table 5).
(b) This mutation was identified in patient D (table 5).

3.1.1 Results of the genotype analysis

3.1.1.1 CFTRDelEx2 and CFTRDel Ex2,3

Among the 770 CF mutations submitted to the Cystic Fibrosis Genetic analysis consortium, (http://www.genet.sickkids.on.ca/cftr/), only seven large deletions have been reported. This is partly based on the method employed for mutation analysis: the common screening methods rely on SSCP or DGGE, which are both PCR based. Any deletion encompassing the analysed PCR fragment will not be detected by PCR unless presented in the homozygous state.

In two patients, both children of a consanguineous marriage, deletions of complete exons could be identified. The missing exons were first noticed during SSCP analysis where no product for exon 2 (patient I) or exon 2 and exon 3 (patient D) was amplified.

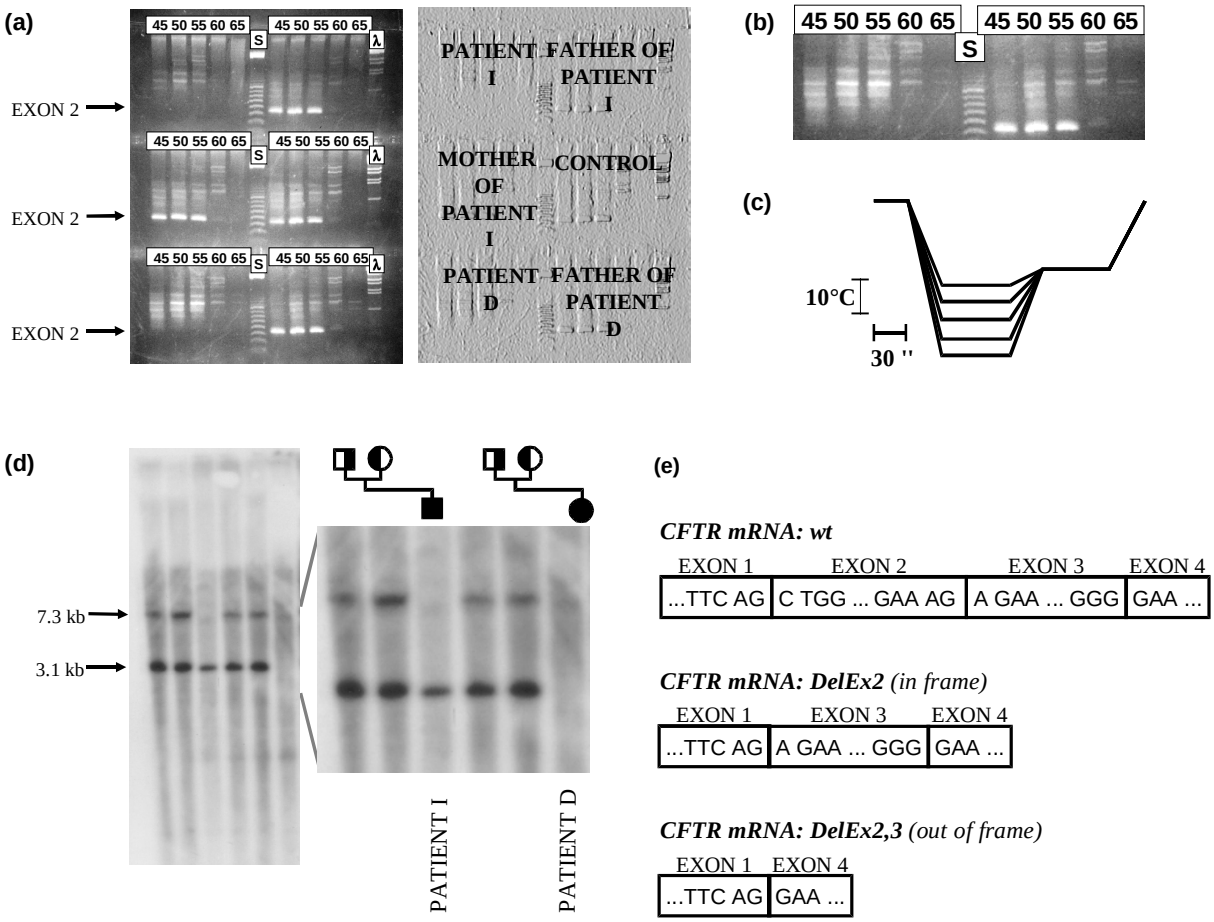
Amplification of exon 3 from patient D under optimised conditions did not show a product of the expected size, but showed several bands, the most prominent among them having a size of 450 bp, as judged on a 2% agarose gel. However, analysis of the excised band on a 12% polyacrylamide gel revealed a composition of more than one band. Direct sequencing of the 450 bp band excised from the agarose gel was carried out and the band pattern could be interpreted at several stretches of 10-20 bp length, none of them composed of exon 3 sequences (data not shown). Consequently, the PCR results were interpreted by assuming a deletion of exon 3 in patient D leading to a non-specific amplification of other genomic sequences unrelated to exon 3.

Amplification of exon 2 with optimised PCR conditions in patients I and D failed to amplify any product. The amplicon was further analysed by PCR employing protocols of increasing stringency by subsequently rising the annealing temperature (figure 16a-c). At 45°C, 50°C and 55°C, PCR product was amplified from DNA of a control, both parents of patient I and the father of patient D. The same samples failed to amplify exon 2 product at annealing temperatures of 60°C and 65°C. From the DNA of patients I and D, no exon 2 band could be amplified under any condition. However, an identical by-product pattern was generated when amplifying the DNA of either patient I or D at low annealing temperatures or the DNA of the control, of both parents of patient I or the father of patient D at higher annealing temperatures. The results were explained by assuming a deletion of exon 2 in patients I and D, resulting in binding of PCR primers to sites within sufficient homology to the target sequences to allow product formation. The same secondary binding sites are present in patients and controls, resulting in an identical band pattern when the target sequence is not available – either by deletion of the target sequence in the patient or by increasing the annealing temperature above optimum conditions in control samples.

Both deletions were confirmed by Southern blotting as described in chapter 2.1.3 (figure 16d). Since the PCR probe used for hybridisation was generated with exon-flanking intron primers (Zielenski et al 1991), it is evident that complete exons are missing in the deletions and that the respective deletion breakpoints are located within the introns upstream and downstream of the oligonucleotide primer sequences. The sequence of the exon/intron boundaries from exon 1, 2, 3, and 4 as published by Zielenski *et al*. (1991) predicts that the deletion of exon 2 is an in frame deletion while the deletion of exon 2 and exon 3 will produce a frameshift (figure 16e).

FIGURE 16: ANALYSIS OF THE DELETIONS CFTRDelEx2 AND CFTRDelEx2,3

(a) Analysis of exon2 deletion with PCR, Assignment of exon 2 product was done by size comparison with a 100 bp ladder (lane S) and a lambda DNA BstEII digest (lane λ)
(b) Detail showing similarity of by-product patterns between (left)sample without exon 2 target sequence and (right) sample with exon 2 target sequence.
(c) Temperature profile of 35 extension cycles used to amplify products shown in figure (a).
(d) Detection of CFTRDelEx2 *(patient* I*) and* CFTRDelEx2,3 *(patient* D*) by Southern blotting as described in chapter 2.1.3. The PCR generated genomic probe encompasses the full exon and flanking intron sequences (Zielenski et al. 1991). Fragments were assigned according to the physical map of Rommens et al. (1989): the expected fragments are 7.3 kb (corresponding to exon2 — HindIII fragment) and 3.1 kb (corresponding to exon3 — HindIII fragment).*
(e) Predicted mRNA sequences for expression of CFTR-DelEx2 (in frame deletion) and CFTR-DelEx2,3 (out of frame deletion) based on the published exon/intron structures of the CFTR gene (Zielenski et al. 1991).



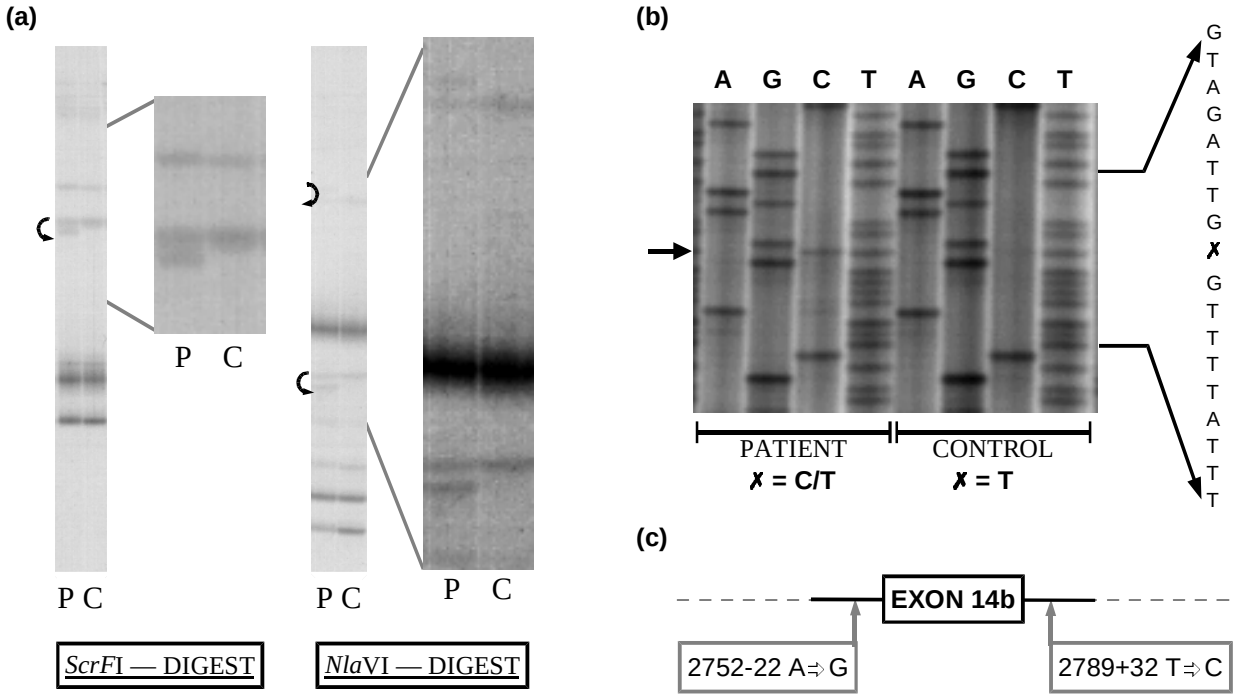
3.1.1.2 Polymorphisms in intron 14a and intron 14b

A novel intronic sequence variant was detected in intron 14b in patient P. The mutation was detected by SSCP in two different digest. Upon sequencing, two sequence alterations encompassing exon 14b were detected: 2752-22A⇨G in intron 14a[5] and 2789+32T⇨C in intron 14b (this work). Both nucleotide exchanges were located on the maternal chromosome. SSCP and sequencing analysis is shown in figure 17.

5 described by Marrigo et al. (personal communication)

FIGURE 17: EXON14b SSCP AND SEQUENCING

(a) SSCP analysis of exon14b and flanking intron sites.
The arrows denote band shifts observed comparing control (C*) and patient P (* P*) samples.*
(b) Sequencing of 2789+32T⇨C *and control sample. The arrow indicates the nucleotide exchange in the patient* (✗)
(c) Location of two sequence alterations of maternal chromosome in patient P.



3.1.2 Results of the phenotype analysis

The *CFTR* mutation genotypes have been categorised in five classes according to the anticipated effect of the molecular lesion (reviewed by Zielenski and Tsui 1995). Mutations which are expected to avert protein synthesis, such as nonsense and frameshift mutations, are combined in class I. The phenotype of four patients homozygous for class I mutations is described in chapter 3.1.2.1. Class II mutations are presumed to fail in producing functional CFTR based on a processing defect, with the most prominent class II mutation being ΔF508. Three patients carrying a second sequence variant on a ΔF508 chromosome were examined and the results are shown in chapter 3.1.2.2. Results on two patients homozygous for mutations which might give rise to proteins defective in processing (class II), regulation (class III) or to CFTR with an altered conductance (class IV) are reported in chapter 3.1.2.3. Class V mutations are defined as molecular lesions leading to missplicing of RNA implying reduced amounts of wild type CFTR. In chapter 3.1.2.4, data on patients carrying one or two splice mutations is reported. The last paragraph of this section deals with *CFTR* genotypes that fit into none of the above categories: for 5 patients participating in this study, only one or even no disease-causing lesion could be identified in the *CFTR* gene (chapter 3.1.2.5).

3.1.2.1 Nonsense mutation and out of frame deletion

The four adult patients carrying two nonsense mutations (A, B, C) or a frameshift mutation caused by a deletion (D, see figure 16 for details[6]) are all pancreatic insufficient and all showed highly elevated sweat Cl^- values. While their nutritional status, shown by wfh%, was average for all patients, the lung function parameters of patient A to D ranged from normal (A) to severe impairment (D). Patients B, C and D were colonised with *P. aeruginosa*. Patient A had *S. aureus*.

TABLE 6: NONSENSE MUTATION AND OUT OF FRAME DELETION
PI*: pancreatic insufficient*
ICM *intestinal current measurement Classification of ICM results as described in chapter 2.2.2.2.1, page 57:* **CF***: carbachol response displays only reversed component.* **n.c.**: *non-conclusive.*
NPD *nasal potential difference measurement Classification of NPD results as described in chapter2.2.2.1, page 54:* **CF**: *no residual secretion.* **CF-res**: *residual secretion.*

	AGE			*CFTR* GENOTYPE		WFH%	FEV1 %pred	FVC %pred	Sweat Cl^- [mval/l]	ICM	NPD
A	24 y	m	PI	G542X	Y1092X	94	111	98	115	n.c.	CF-res
B	23 y	f	PI	R553X	L1059X	116	68	87	106	CF	CF
C	25 y	f	PI	R553X	R553X	103	44	71	96	CF	CF-res
D	18 y	f	PI	CFTR DelEx2,3	CFTR DelEx2,3	92	46	64	103	CF	CF-res

References: Kerem 1990 *et al.* (G542X), Shoshani *et al.* p. c. (Y1092X), Bozon *et al.* p. c. (Y1092X), Cutting *et al.* 1990 (R553X), Dörk *et al.* 1994b (L1059X), this work (CFTR DelEx2,3)

The investigation of the basic defect in rectal tissue showed no evidence for residual secretion in any of the four patients: the responses of the rectal tissue to carbachol and histamine showed only the reversed response typical for CF. With NPD, residual secretion was detected in patients A and D upon superfusion of the nasal epithelium with chloride free solution in the presence of amiloride (ΔPD of 3 mV in patient A and 4 mV in patient D, respectively). In patients A and C, chloride secretion was stimulated by isoproterenol in chloride-free solution in the presence of amiloride (ΔPD of 2 mV in patient A and 6 mV in patient C, respectively). Patient B did not show any residual secretion in nasal tissue.

3.1.2.2 Complex alleles

Two adult siblings carrying the complex allele ΔF508-V1212I (E, F) and one patient bearing ΔF508-R553Q (G) have been investigated. On the day of clinical investigation, all patients had highly elevated sweat Cl^-, were pancreatic insufficient and had a normal wfh%. Patient G had sweat chloride values of 50-70 mval/l in childhood. The siblings E and F were discordant with respect to lung function parameters FEV1%pred and FVC%pred. Patients E, F and G showed no (E), moderate (F) or severe (G) impairment of pulmonary function. Siblings E and F were colonised with *P. aeruginosa*. The sib pair was discordant in pulmonary status and sweat Cl^- concentration.

6 Shown in chapter 3.1.1.1 on page 62

TABLE 7: COMPLEX ALLELES

Patients E and F are siblings.

PI*: pancreatic insufficient*

ICM *intestinal current measurement* *Classification of ICM results as described in chapter 2.2.2.2.1, page 57:*
CF-res*: carbachol response shows residual component*

NPD *nasal potential difference measurement* *Classification of NPD results as described in chapter2.2.2.1, page 54:*
CF: *no residual secretion.*

	AGE			*CFTR* GENOTYPE		WFH%	FEV1 %pred	FVC %pred	Sweat Cl^- [mval/l]	ICM	NPD
E	22 y	f	PI	ΔF508	ΔF508-V1212I	115	92	114	87	CF-res	CF
F	24 y	m	PI	ΔF508	ΔF508-V1212I	104	52	77	104	CF-res	CF
G	17 y	f	PI	R553X	ΔF508-R553Q	101	49	66	117	CF-res	CF

References: Rommens *et al.* 1989 (ΔF508), Macek Jr p.c. (ΔF508-V1212I), Cutting *et al.* 1990 (R553X), Dörk *et al.* 1991 (ΔF508-R553Q)

Patient E and F both showed residual secretion measured by ICM. The carbachol responses were composed of equal proportions of reversed and residual secretion in patient E. Patient F displayed only the residual component. In both siblings, no cAMP response was seen and the residual chloride conductance was inhibited by DIDS. In rectal tissue, patient G showed upon stimulation with carbachol and histamine a reversed response with a pronounced offset towards the residual direction indicating underlying residual secretion. No effect of cAMP or DIDS was observed. The investigation of nasal tissue by NPD did not give evidence for residual secretion in patient G. The ICM tracings of patients E, F and G are shown in figure 18, page 68.

3.1.2.3 Missense mutation and in frame deletion

Two patients homozygous for mutations that might give rise to non-wild-type CFTR protein were characterised. The patient carrying the missense mutation E92K (H) was pancreatic sufficient. Patient I carried the in-frame deletion of exon 2 (see figure 16 for details[7]) and was pancreatic insufficient. Both patients had highly elevated sweat Cl^-, a more than average wfh% and a normal lung function. On the day of clinical investigation, *P. aeruginosa* was detected for the first time in patient H. Patient I had *S. aureus*.

In patient H, residual secretion was detected in ICM and NPD. The rectal tissue showed residual secretion upon stimulation with carbachol and cAMP. The residual component was insensitive to DIDS. By NPD, a hyperpolarization of 11 mV was detected after superfusion with chloride-free solution in the presence of amiloride. In patient I, evidence for residual chloride secretion was seen by ICM, but not by NPD. In rectal tissue, a cAMP stimulated residual secretion was seen in the absence, but not in the presence, of DIDS. The ICM tracings of patients H and I are shown in figure 18, page 68.

7 Shown in chapter 3.1.1.1 on page 62

TABLE 8: MISSENSE MUTATION AND IN FRAME DELETION

PS*: pancreatic sufficient,* **PI***: pancreatic insufficient*

ICM *intestinal current measurement Classification of ICM results as described in chapter 2.2.2.2.1, page 57:*

CF-res*:carbachol response shows residual component* (H) *or cAMP response present* (I).

NPD *nasal potential difference measurement Classification of NPD results as described in chapter2.2.2.1, page 54:*

CF: *no residual secretion.* **CF-res**: *spontaneous residual secretion.*

	AGE			*CFTR* GENOTYPE		WFH%	FEV1 %pred	FVC %pred	Sweat Cl^- [mval/l]	ICM	NPD
H	12 y	m	PS	E92K	E92K	141	101	99	118	CF-res	CF-res
I	16 y	m	PI	CFTR DelEx2	CFTR DelEx2	121	133	130	102	CF-res	CF

References: Nunes *et al.* 1993 (E92K), this work (CFTR DelEx2)

3.1.2.4 Splice mutations

Four adults and an infant, being homozygous for a splice site sequence alteration (J, K, N) or heterozygous for splice mutation and ΔF508 (L, M) were investigated. Except for patient J, all had elevated sweat Cl^- values. All patients had a normal wfh%. Only patient M was pancreatic insufficient. All adult patients showed moderately (patient K) to severely impaired (patient L) pulmonary function. No lung function data was available for patient N. All adult patients J, K, L, M were *P. aeruginosa* colonised. Sweat test values for siblings J and K were discordant on the day of clinical investigation, but for patient K, lower values have been reported: sweat chloride values were non-pathological until the age of 21. When diagnosed as CF, patient K had a sweat test of 68 mval/l. By the age of 28 years, four sweat tests were within the borderline range (46, 51, 53 and 46 mval/l respectively. Patient K received a lung transplant.

TABLE 9: SPLICE MUTATIONS

Patients J and K are siblings.

PS*: pancreatic sufficient,* **PI***: pancreatic insufficient*

ICM *intestinal current measurement Classification of ICM results as described in chapter 2.2.2.2.1, page 57:*

CF-res*: carbachol response shows residual component* (J, K) *or cAMP-response* (L,M) **CF***: carbachol response displays reversed direction.* **n.c.***.: non-conclusive*

NPD *nasal potential difference measurement Classification of NPD results as described in chapter2.2.2.1, page 54:*

CF: *no residual secretion.* **CF-res**: *spontaneous residual secretion.*

	AGE			*CFTR* GENOTYPE		WFH%	FEV1 %pred	FVC %pred	Sweat Cl^- [mval/l]	ICM	NPD
J	20 y	m	PS	3849+10kb C⇨T	3849+10kb C⇨T	98	44	59	20	CF-res	CF
K	31 y	m	PS	3849+10kb C⇨T	3849+10kb C⇨T	97	82‡	68‡	93	CF-res	CF-res
L	58 y	f	PS	ΔF508	3272-26 A⇨G	99	29	55	81	CF-res	CF-res
M	36 y	m	PI	ΔF508	3850-3 T⇨G	103	32	39	116	CF-res	CF
N	3 y	f	PS	1898+3 A⇨G	1898+3 A⇨G	112	—	—	74	n.c.	CF-res

‡Patient K received a lung transplant.

References: Highsmith *et al.* 1994 (3849+10kbC→T), Rommens *et al.* 1989 (ΔF508), Fanen *et al.* 1992 (3272-26A→G), Dörk *et al.* 1993 (3850-3T→G), Ferrari *et al.* p.c. (1898+-3A→G)

In patients J, K and L, ICM measurement showed DIDS-insensitive residual secretion upon stimulation with carbachol and histamine. The residual secretion was stimulated by cAMP for patient L and M. The cAMP response of patient M was seen in the absence , but not in the presence, of DIDS. With NPD, spontaneous chloride secretion was detected in patients K, L and N upon superfusion of the nasal epithelium with chloride-free solution in the presence of amiloride (ΔPD of 12 mV in patient K, 2 mV in patient L and 3 mV in patient N, respectively).

3.1.2.5 Unresolved *CFTR* genotypes

Mutation analysis is carried out to confirm the suspected diagnosis of CF. For the recessive inherited disease, the identification of two mutant *CFTR* alleles is considered to be diagnostic. However, for most populations a detection rate of 100% is not achieved due to the allelic heterogeneity of CF disease-causing lesions. For the German population, extensive screening has enabled identification of 54 mutations on 94% of 700 CF chromosomes (Dörk *et al.* 1994b). In the Spanish CF population, 75 mutations were identified on 90% of the CF chromosomes from 640 families (Casals *et al.* 1997). 47 mutations were reported to account for 86% of CF chromosomes from 600 CF patients within the French non-celtic population (Chevalier-Porst *et al.* 1994).

The failure to identify disease-causing lesions in all *CFTR* alleles of CF patients may be due to technical reasons, as most screening protocols analyse the coding region and the flanking intron sites, but not the entire *CFTR* gene encompassing 240 kb. Hence, sequence variants at regulatory elements within the non-coding sequences might remain unnoticed. Further, deletions of entire exons cannot be detected by PCR based methods in the heterozygous state because the non-deleted copy on the other chromosome will always be amplified. Consequently, if a patient is diagnosed as having CF but at least one *CFTR* allele appears to be normal after thorough mutation screening, it cannot be distinguished whether a sequence variant within the *CFTR* gene has been missed due to technical reasons or whether the disease is caused by another entity.

Within this chapter, the clinical phenotype and the basic defect is described for four patients who do not carry two disease-causing *CFTR* mutations as judged by SSCP screening of all *CFTR* exons and flanking intron sites (patients P, Q, R, S; table 10). A case of atypical CF carrying a splice site alteration and the IVS8-6(5T) allele (patient O) is discussed in comparison to patient P. Patients O – S were pancreatic sufficient. The diagnosis of CF was unclear for patients O, P and Q.

TABLE 10: UNRESOLVED GENOTYPES

S*: Exon was sequenced after SSCP analysis*

1*: SSCP analysis performed with one restriction digest* 2*: SSCP analysis performed with two restriction digests*

	SSCP SREENING OF *CFTR* EXON																									GENOTYPE
	1	2	3	4	5	6a	6b	7	8	9	10	11	12	13	14a	14b	15	16	17a	17b	18	19	20	21	22	23	24	
P	2	1	S	2	1	2	2	2	1	2	S	2	2	2	2	2	2	2	1	2	2	1	1	1	1	2	2	IVS8-6(5T); [1]
Q	2	S	1	2	1	2	2	2	1	1	S	2	2	1	2	2	2	2	1	1	2	1	2	2	2	1	2	1898+3 A⇨ G
R	2	1	S	2	1	2	2	2	2	S	1	S	2	1	2	2	2	2	1	1	1	S	S	2	2	2	2	—
S	2	S	S	2	1	1	S	1	2	S	S	S	2	1	2	2	2	2	2	2	1	S	S	S	1	1	2	—

[1] *:* 2752-22A⇨ G – 2789+32A⇨ G

References: Chu *et al.* 1991 (IVS8-6(5T)), Ferrari *et al.* p.c. (1898+3A→G), Marigo *et al.* p.c. (2752-22A→G), this work (2789+32A→G)

FIGURE 18: ICM RESULTS FOR SELECTED PATIENTS WITH RARE GENOTYPES OR ATYPICAL CF

Examples for residual secretion in 6 CF patients. cAMP-sensitivity is assumed if a residual current is stimulated by addition of cAMP and forskolin. The DIDS-sensitivity is assessed by comparison of the carbachol- and the histamine-induced response when DIDS was added 10' prior to stimulation with histamine.

	TYPE OF RESIDUAL SECRETION			TYPE OF RESIDUAL SECRETION	
	cAMP-sensitive	DIDS-inhibitable		cAMP-sensitive	DIDS-inhibitable
PATIENT E	NO	YES	PATIENT F	NO	YES
PATIENT G	NO	NO	PATIENT H	YES	NO
PATIENT I	YES	NO	PATIENT R	NO	NO

Scales for Isc and t are identical for all tracings displayed in this figure.

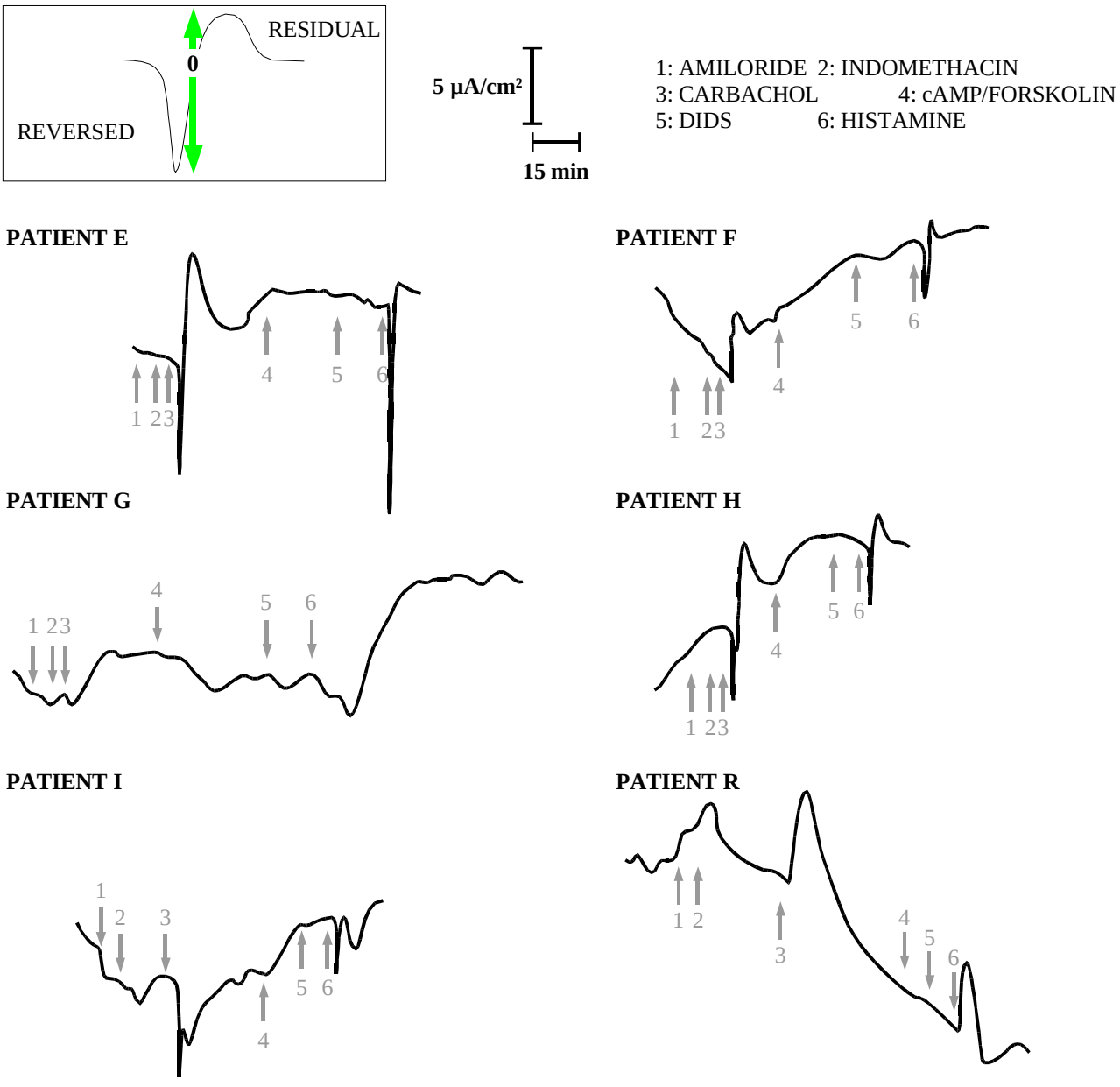


TABLE 11: PHENOTYPE OF UNRESOLVED GENOTYPES

PS*: pancreatic sufficient*

ICM *intestinal current measurement Classification of ICM results as described in chapter 2.2.2.2.1, page 57:* **normal***: all values within range of non-CF controls.* **CF-res***: carbachol response shows residual component.*

NPD *nasal potential difference measurement Classification of NPD results as described in chapter2.2.2.1, page 54:* **normal***: all values within range of non-CF controls.* **CF**: *no residual secretion.* **CF-res**: *spontaneous residual secretion*

	AGE			*CFTR* GENOTYPE		WFH%	FEV1 %pred	FVC %pred	Sweat Cl^- [mval/l]	ICM	NPD
O	8 y	f	PS	IVS8-6(5T)	4000-2 G⇨ A	88	82	83	20	normal	normal
P	28 y	m	PS	IVS8-6(5T)	[1]	103	101	102	55	normal	normal
Q	32 y	f	PS	1898+3 A⇨ G	—	84	18	40	7	normal	CF res
R	37 y	m	PS	—	—	131	69	104	59	CF-res	CF
S	24 y	m	PS	—	—	127	83	87	97	normal	normal

[1] *:* 2752-22A⇨ G – 2789+32A⇨ G

Patients O and P

CLINICAL HISTORY OF PATIENT O Patient O presented a history of allergy, wheezing and bronchiolar obstruction. She was referred to the CF clinic Hannover for differential diagnosis of asthma/neurodermitis vs CF with pulmonary manifestations. *CFTR* mutation analysis had revealed the splice mutation 4000-2 G⇨ A and the variant IVS8-6(5T) on the two chromosomes, respectively.

CLINICAL HISTORY OF PATIENT P Patient P was born as a child of unrelated parents of German descent. Respiratory symptoms were known for paternal members of a three generation pedigree: the grandfather of patient P suffered from chronic bronchitis. The brother of the patient's grandfather died of tuberculosis before the age of 10 years. The father of patient P suffered from a chronic cough with sputum production. On the occasion of his son's diagnosis, a sweat test of 80 mval/l Cl^- was reported. The patient's father was underweight and small in childhood, but gained normal weight in adolescence. Patient P was diagnosed at the age of 14 years by sweat electrolytes[8] of 74 mval/l Na^+ and 76 mval/l Cl^-. The clinical course during infancy and childhood was characterised by growth retardation with poor nutritional status: by the age of 7 years, patient P had the height and weight of a 5-year old child. Recurrent pulmonary infections were treated by three to four courses of antibiotics per year. By the age of 24 years, he presented normal weight, height and lung function parameters and a spermiogram was within the normal range.

GENETIC ANALYSIS OF PATIENT P The *CFTR* gene was screened for mutations in all exons and flanking intron sites. No disease-causing lesion was identified. The patient carried intronic sequence variants surrounding exon 14b on the maternal chromosome (figure 17, page 63). His paternal *CFTR* gene carries the IVS8-6(5T) variant. He is heterozygous at M470V and T854T.

8 The value reported in table 11 was obtained on the day of investigation at the age of 28 years. Sweat tests of patient P have been reported Na^+ 74 mval/l and Cl^- 76 mval/l; Na^+ 62 mval/l and Cl^- 65 mval/l; Cl^- 55mval/l

BASIC DEFECT OF PATIENTS O AND P The investigation of the basic defect did not provide evidence for CFTR dysfunction in patients O and P: with ICM, normal ion secretion was seen upon addition of carbachol and histamine. Results of NPD were consistent with non-CF based on the basal PD (-34 mV and – 14 mV for patients O and P, respectively) and hyperpolarization (ΔPD 19 and 15 mV for patients O and P, respectively) upon superfusion with chloride-free solution in the presence of amiloride. An additional chloride secretory response was stimulated with isoproterenol (ΔPD 8 and 3 mV for patients O and P, respectively).

Patient Q [*case report published in Thorax 54: 278-281(1999)*]

CLINICAL HISTORY OF PATIENT Q Patient Q was the third child born to healthy German parents who are first generation cousins. Both elder siblings of patient Q are deceased: one sibling was born preterm with meconium ileus and died at the age of 10 days. The other sibling died before the age of 6 months. CF was suspected from the autopsy. Patient Q had a history of pulmonary complications with recurrent lower airway infections, chronic cough and sputum production. Chronic nasal polyposis was surgically treated by 13 polypectomies between 5 and 23 years of age. She underwent a lung biopsy at the age of 21 which revealed bronchiectasis, localised purulent bronchitis and surrounding fibrosis. The diagnosis of CF was discarded due to repetitive sweat electrolyte values within the normal range (7-32 mval/l). Nine out of ten throat swabs were positive for *S. aureus*. Antibody titer against *P. aeruginosa* oprF were elevated though no *P. aeruginosa* were cultured. Bronchodilators and antibiotics were prescribed, but the patient generally discontinued medication after a few days. At the age of 32, the patient presented severe clubbing and required continuous oxygen. The patient has been underweight since adolescence, but pancreatic sufficient as determined by normal chymotrypsin, stool elastase and ultrasound of the pancreas. She did not receive enzyme or vitamin supplements and had normal vitamin A and E levels in serum. The differential diagnosis of allergy was excluded by normal specific IgE levels in serum and skin test. The microscopic appearance of a nasal biopsy allowed exclusion of immotile cilia syndrome.

GENETIC ANALYSIS OF PATIENT Q Mutation analysis of the *CFTR* gene revealed heterozygosity for the splice site consensus transition 1898+3 A⇨G. Mutation analysis of the *CFTR* gene did not reveal other disease-causing lesions within the exons or the non coding region up to 4 kb 5' in front of the gene. No evidence for a large genomic deletion was found by Southern blot analysis of an *Apa*I digest hybridised with a CFTR cDNA probe encompassing exons 7 to 24. The patient is homozygous for the $(TG)_mT_n$-M470V-T854T haplotype $(TG)_{11}T_7$-2-1.

BASIC DEFECT OF PATIENT Q The analysis of the basic defect in rectal tissue showed no evidence for defective CFTR as demonstrated by normal ion currents after stimulation with carbachol and histamine in ICM. However, in nasal tissue CF conditions were detected by NPD: patient Q had a basal of –52 mV. Evidence for residual chloride secretion was provided by a hyperpolarization of 3 mV upon superfusion with chloride-free solution in the presence of amiloride. An additional chloride secretory response of 4 mV was stimulated with isoproterenol.

Patient R

CLINICAL HISTORY OF PATIENT R Patient R was diagnosed as having CF during adulthood at the age of 28 based on pulmonary symptoms. He suffered from pneumonia in childhood and always had nasal polyps. Sweat tests have been reported to be borderline or within the lower pathological range. On the day of clinical investigation at 37 years of age, he had an impaired lung function and clubbing was noticed. *Pseudomonas aeruginosa* was detected in sputum culture and the antibody titer against *P. aeruginosa* oprF was elevated. He was treated with inhalation of β_2 adrenergic agonists, oral n-acetylcystein and antibiotics. Patient R was pancreatic sufficient. He received vitamin supplements. Serum vitamin A and E levels have been within the normal range. He did not show any signs of liver disease. Patient R is the father of two healthy children.

GENETIC ANALYSIS OF PATIENT R Mutation analysis of the *CFTR* gene did not provide evidence for disease-causing lesions within the exons or the non coding region up to 4kb 5' in front of the gene. The *CFTR* gene was investigated for deletions by pulsed field gel electrophoresis. Detection of restriction fragments after an *Apa*I digest (fragment size 30 kb and 65 kb), *Fsp*I digest and *Sal*I digest (fragment size ≈ 500 kb) did not show any deviation compared to control samples when hybridized to a *CFTR* cDNA probe containing exons 7 to 24. Patient R is homozygous for the M470V-T854T haplotype 2-1.

BASIC DEFECT OF PATIENT R The investigation of the basic defect confirmed the diagnosis of CF in patient R. In rectal tissue, carbachol and histamine stimulated a residual current without the reversed component typical for CF. However, the magnitude of the responses was only a fifth of the size encountered in non-CF controls (figure 18, page 68). DIDS did not have an effect on the residual currents provoked by carbachol or histamine. In nasal tissue, a basal PD of –39 mV was seen. In the presence of amiloride, neither the superfusion with a chloride-free solution nor the stimulation with isoproterenol provoked a hyperpolarization.

Patient S [*case report published in Hum Genet 102: 582-586 (1998)*]

CLINICAL HISTORY OF PATIENT S Patient S had a history of respiratory symptoms: from the age of 6 years, he has suffered from recurrent bronchitis and chronic cough with excessive sputum production. At the age of 17 years he was diagnosed as having CF on the occasion of a severe pneumonia by highly elevated sweat Cl^- of 102 mval/l. Since then he has been treated by daily inhalation of salbutamol and oral n-acetylcystein and antibiotics. The patient had three to four two-week courses of antibiotic therapy per year. *S. aureus* was detected occasionally in throat swabs or sputum cultures. He is pancreatic sufficient and received no vitamin supplements. Reduced fertility was diagnosed by spermiogram. The differential diagnoses of pseudohypoaldosteronism and Liddle's disease were excluded by normal blood pressure and serum electrolytes as well as plasma renin and aldosterone levels prior to and after administration of furosemide.

GENETIC ANALYSIS OF PATIENT S AND HIS SISTER Mutation analysis of the *CFTR* gene did not reveal any disease-causing lesions. Furthermore, the *CFTR* genes of the patient and his healthy sister are identical by descent as demonstrated by haplotype analysis of polymorphisms within the *CFTR* gene including the fully informative microsatellite marker IVS17bTA. No

recombination event was detected within a 40 cM region encompassing the *CFTR* gene, ranging from the markers PON2 to D7S495 (figure 19). The patient is carrying the $(TG)_mT_n$-M470V-T854T alleles $(TG)_{11}T_7$-2-1 and $(TG)_{10}T_7$-2-1.

BASIC DEFECT OF PATIENT S AND HIS SISTER The investigation of the basic defect did not provide evidence for CFTR dysfunction in patient S or the patient's sister in intestinal or nasal tissue: with ICM, normal ion secretion was seen upon addition of carbachol and histamine. Results of NPD were consistent with non-CF based on a low basal PD (-21 mV and –12 mV for patient S and the patient's sister, respectively) and hyperpolarization (ΔPD 28 and 21 mV for patient S and the patient's sister, respectively) upon superfusion with chloride-free solution in the presence of amiloride.

FIGURE 19: HAPLOTYPES OF PATIENT S AND HIS SISTER

(a) Map showing 40 cM region encompassing the CFTR gene investigated in family of patient S.

(b) 16 marker haplotype for patient S (indicated by arrow) and his sister.

— : non-conclusive (PCR-failure)

Informative phase: alleles identical by descent are marked in red.

Non-informative phase: alleles identical by state are printed in black.

*Alleles at dimorphic marker loci (*PON2, MetH, XV-2c, KM.19, J44, GATT, M470V, T854T, TUB20, J3.11*) are named according to standard nomenclature (see figure 5 on page 39).*

*Alleles at polymorphic loci are named according to their size in repeat units (*GATT, IVS8CA, IVS17bCA, IVS17bTA*) or in arbitrary repeat units as defined in chapter 2.1.2.2.4 (*D7S525, D7S514, D7S495*).*

The alleles of patient S and his sister at J3.11 *are identical by descent unless a double recombination event flanking* J3.11, *with breakpoints between* TUB20 *and* D7S514, *is assumed.*

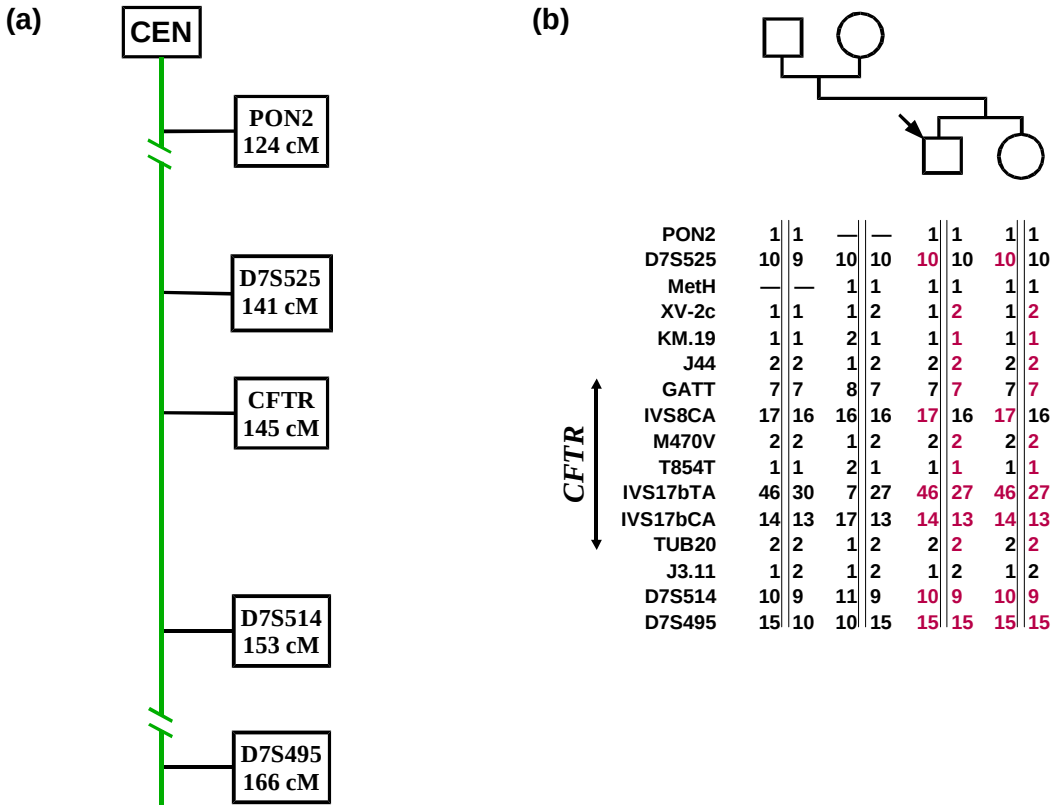


FIGURE 20: RESIDUAL CHLORIDE SECRETION IN CF AND ATYPICAL CF

Residual chloride secretion in patients **A-S** *as assessed by nasal potential difference (PD) measurement, sweat test and intestinal current measurement (ICM).*

NPD: *response to superfusion with chloride-free solution in the presence of amiloride given as* ΔPD[mV]

■ : *CF- typical basal potential difference* |PD| ≤ 40mV

■ : *non-CF or borderline basal PD* |PD| ≥ 40mV

Patient B probably had a cold on the day of investigation and hence a low basal PD.

Sweat chloride:

■ : *sweat chloride concentration given as* [10 × mval/l]

(< 40 mval/l: non-CF range, 40-60 mval/l: borderline range , > 60 mval/l: CF range)

ICM: ❒ : *tracing non conclusive*

classification of carbachol-response:

■, ■, ■, ■ : *% reversed component of carbachol induced signal*

maximum 100% (only reversed response), patient **B, C, D, I, L, M**

minimum 0% (only residual response), patient **F, K, O, P, Q, R, S**

amplitude of chloride secretory component of the carbachol-response (Veeze *et al.* 1991, 1994):

■ : *carbachol-response within non-CF range*

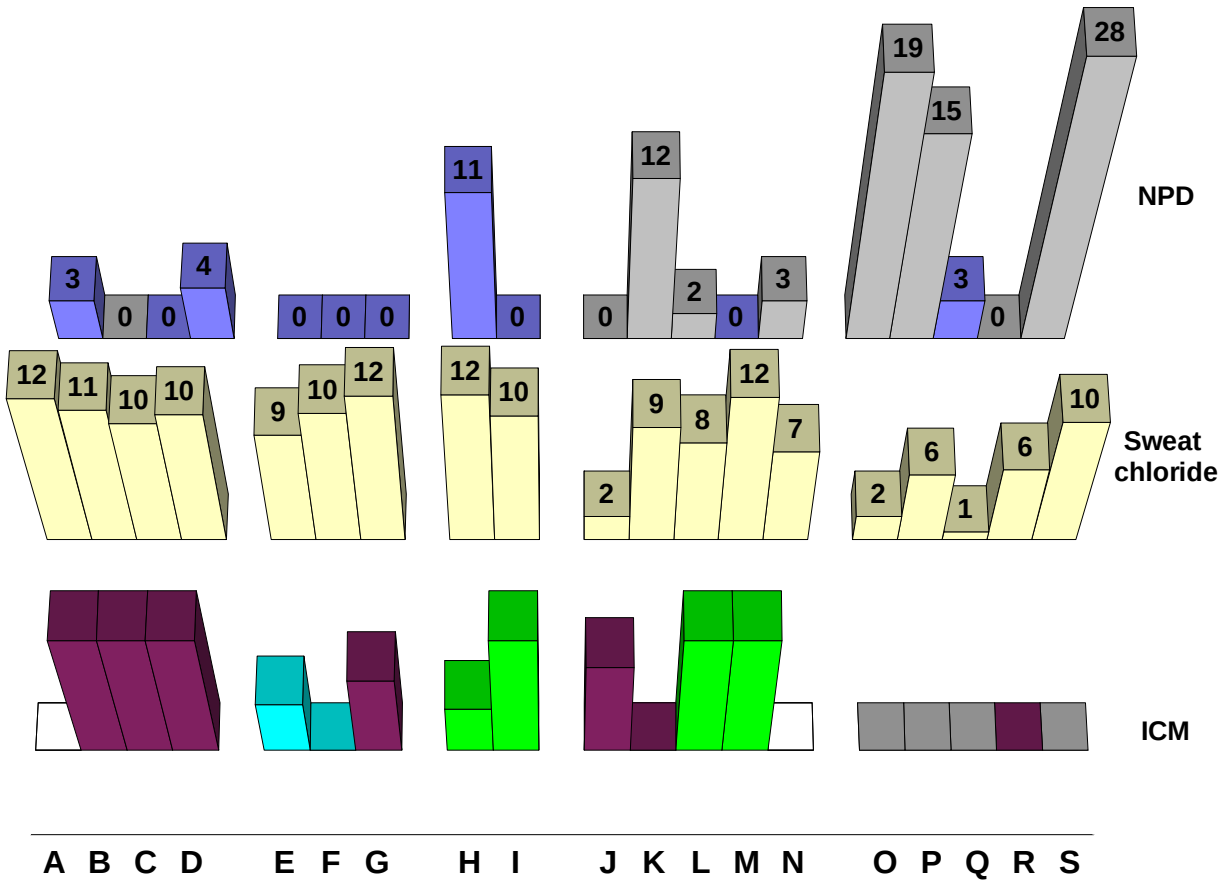
■, ■, ■ : *carbachol-response within CF range*

type of residual chloride secretion:

■ : *no classification of residual component or no residual component observed*

■ : *DIDS-sensitive residual chloride secretion*

■ : *cAMP-sensitive chloride secretion*



3.1.3 Discussion of genotype-phenotype relations

The basic defect in cystic fibrosis is defined by an absence of CFTR mediated chloride conductance in CFTR expressing epithelia. Within the study "*rare genotypes and atypical cystic fibrosis*", patients have been investigated with respect to their basic defect expressed in the epithelium of the lower nasal turbinate (considered representative for the respiratory epithelium), of the sweat gland and of rectal tissue (considered representative for the intestinal epithelium). The results obtained are studied in two approaches: chapter 3.1.3.1. aims at correlating the basic defect in the three tissues examined, regardless of the underlying *CFTR* genotype. In chapter 3.1.3.2, the patient's *CFTR* genotype is taken into account: the basic defect phenotype is evaluated on the basis of the CFTR genotype and the clinical phenotype is evaluated on the basis of the basic defect results.

3.1.3.1 Correlation of the basic defect in three different tissues

Given the hypothesis that the basic defect in cystic fibrosis relies solely on the inability of mutant CFTR to conduct chloride, CFTR expressing epithelia from different tissues but from the same patient should express a comparable picture as the underlying *CFTR* mutation is identical for all tissues of one CF patient. Alternatively, the defect in ion conductance observed in one tissue would not be predictive for the expression of the basic defect in other CFTR expressing tissues from the same patient due to tissue specific, but *CFTR* gene independent factors. This chapter confronts this hypothesis in three sections: first, cases of typical CF are discussed. Then, patients with atypical CF are evaluated. Finally, a brief survey of published data on CFTR associated ion secretion in epithelial cells is presented as a lack of correlation between the impairment of ion conductance of the sweat gland, the intestinal and the respiratory tissue cannot be based on the *CFTR* mutation genotype.

3.1.3.1.1 Cystic fibrosis patients A-N

Given the hypothesis that the defective ion conductance is mainly determined by the *CFTR* genotype, the expression of the basic defect is supposed to be concordant in the three examined tissues taken from the same patient. Figure 20 clearly demonstrates that this is not the case: residual chloride secretion detected in the rectal tissue *(patient E, F, G, H, I, J, K, L, M)* was not always indicative of residual secretion in the nasal tissue *(no residual seen in NPD in patients E, F, G)*. Neither was the evidence for residual secretion in nasal tissue *(patients A, D, H., K, L, N)* indicative in all cases of residual secretion in the rectal tissue *(no residual seen in ICM in patient D)*. Sweat chloride concentrations in the borderline range on the day of clinical investigation *(patient J)* were not matched by a higher content of residual chloride secretion in nasal tissue compared to patients with higher sweat chloride concentrations *(no residual secretion seen in NPD of patient J, the low basal*

PD displayed by patient J was likewise observed in patients K, L and N[9] *which all showed sweat chloride concentrations in the pathological range)*. Two patients *(patients G and K)* were reported to have increasing sweat chloride concentrations with increasing age. This indication of residual chloride secretion is concordant with the residual secretion detected by NPD and ICM in one of the two patients only *(patient K, no evidence for chloride secretion was seen by NPD of patient G)*. In conclusion, when reduced to the presence or absence of residual chloride secretion, the basic defect in the three tissues examined seems to be independently expressed. Hence, other chloride conductances than those mediated by CFTR itself must contribute to the effects observed in a tissue specific manner. Evidence for different sources of residual chloride secretion have been provided by ICM (figure 18[10]): three chloride secretory currents, differing with respect to their sensitivity to DIDS and cAMP, could be discriminated in CF patients.

3.1.3.1.2 Atypical cystic fibrosis patients Q-S

Among the five patients with unresolved *CFTR* genetics (patients O-S) investigated in the study *"rare genotypes and atypical CF"*, the diagnosis of cystic fibrosis was rejected for two patients based on normal ICM and NPD values (see chapter 2.2.2.1 and 2.2.2.2 for details) and sweat chloride concentration in the normal (patient O) or lower pathological range (patient P).

Complete discordance of the basic defect between the two organs affected mainly in CF was observed in patient Q: no evidence for CFTR dysfunction was seen in rectal tissue, but the nasal tissue displayed the CF-typical condition. Residual secretion was detected by NPD. The sweat test of patient Q was non-pathological. In patient R, the basic defect was expressed in all three tissues investigated, but while no evidence for residual chloride secretion was seen in nasal tissue, the rectal tissue showed the chloride secretory component. Patient R had a sweat test compatible with the diagnosis of CF. Patient S had sweat chloride concentrations in the pathological range, but neither NPD nor ICM indicated dysfunctional CFTR in nasal or rectal tissue.

In summary, the tissue-specific expression of the basic defect that was observed for patients with a clear diagnosis of CF (patients A-N) was seen more distinctly for patients with atypical CF. This is consistent with the definition of atypical CF as based on the condition CBAVD, wherein the manifestation of the disease is associated with molecular lesions in the *CFTR* gene (Dumur *et al.* 1990, Chillon *et al.* 1995) and leads to symptoms within one organ only.

9 Patient B also had a low basal, presumably due to a cold.
10 Figure 18 is shown in chapter 3.1.2.5 on page 68.

3.1.3.1.3 Ion transport in epithelial cells

The ion transport in CFTR expressing epithelial cells (reviewed by Anderson 1992) is complex on the level of the regulation pathways and with respect to the variety of the molecular entities involved: different chloride secretory responses have been attributed to Ca^{2+} mediated and cAMP mediated pathways in HT29 cells (Bajnath *et al.* 1992), T84 cells (Vajanaphanich *et al.* 1994) and cultured colonic epithelia (MacVinish *et al.* 1993). Both intracellular mediators are considered to be linked (Vajanaphanich *et al.* 1995). Molecular entities, contributing to the chloride transport of epithelial cells, have been described as outwardly rectifying chloride channel ORCC (Egan *et al.* 1992, Gabriel *et al.* 1993) and, distinguishable from the ORCC (Fischer *et al.* 1992), the outwardly-rectifying depolarisation induced channel ORDIC (reviewed by Thinnes and Reymann 1997). The transport of counterions, as mediated by the amiloride-sensitive Na^+ channel and the basolateral K+ channel, has been demonstrated to be controlled by CFTR (Ismailov *et al.* 1996, Stutts *et al.* 1997, Loussouarn *et al.* 1996, McNicholas *et al.* 1997). The activation of CFTR by extracellular ATP in airway cells (Stutts *et al.* 1995) is considered to be mediated by P2Y2 receptors and CFTR-associated ATP channels (Sugita *et al.* 1998, Watt *et al.* 1998). The cytoskeleton is considered to regulate the interaction of CFTR and the ENaC (Ismailov *et al.* 1997) and the regulation of CFTR mediated chloride transport itself (Naren *et al.* 1998). Recently, the PDZ protein motif, known to bind factors interacting with for example the β2- adrenergic receptor, the purinergic receptor, components of the cytoskeleton (Short *et al.* 1998) or proteins directly involved in ion transport such as the Na^+/H^+ exchanger, has been identified in CFTR (Hall *et al.* 1998) and thus provides evidence for the link of CFTR to other regulatory pathways.

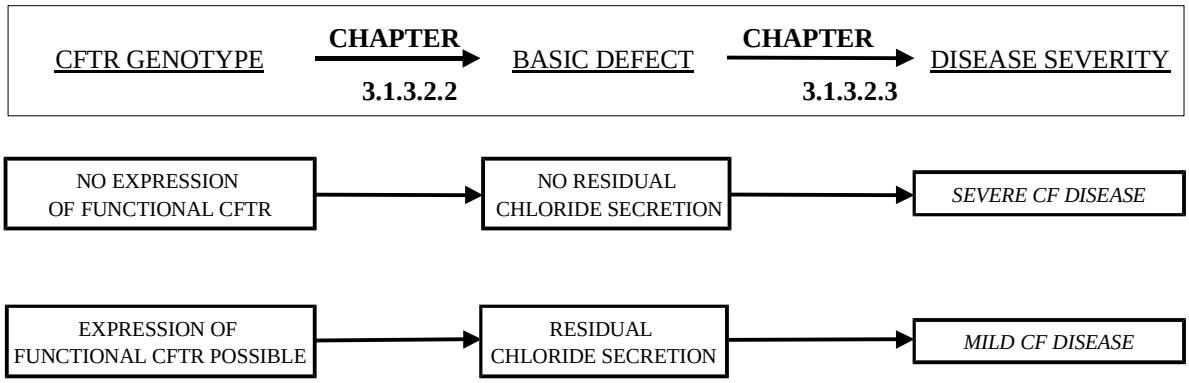
Regarding the complexity of this network, the lack of correlation between the basic defect in the tissues examined seems obligatory: nasal tissue, intestinal tissue and the sweat gland represent different organs and consequently, the regulation of ion transport can be expected to reflect the individual characteristics of the differentiated tissues. Furthermore, while CFTR is expressed in all three tissues examined, this is most likely not the case for other ion pathways. In conclusion, while non-functional CFTR may be regarded as the basis of the basic defect of defective ion conductance in sweat gland, nasal and rectal tissue, various tissue-specific elements can be expected to modulate the ion conductance properties of CFTR expressing epithelia.

3.1.3.2 Correlation of CFTR genotype, basic defect and disease severity

Since the molecular analysis of the *CFTR* gene has allowed the identification of more than 700 sequence variations with presumed functional consequences for the CFTR protein, correlation of *CFTR* mutation genotype and disease severity has been attempted (reviewed by Zielenski and Tsui 1995). Figure 21 illustrates the concept employed: the prediction of the molecular phenotype for a given *CFTR* allele is used to assess whether functional CFTR protein can be expressed. A severe disease phenotype is then assumed to result from non-functional CFTR mutants while *CFTR* alleles allowing the generation of wild type protein are generally expected to result in milder phenotypes. In chapter 3.1.3.2.1, data from the literature on the mutations displayed by patients A to P is summarised. The patient's genotypes are grouped according to whether their *CFTR* genotype is expected to result in no CFTR protein, some mutant CFTR protein or reduced amounts of wild-type CFTR protein. The genotype-phenotype correlation for the patients investigated in the study "*rare genotypes and atypical cystic fibrosis*" is discussed in two steps: the relation between the *CFTR* mutation genotype and the basic defect is discussed in chapter 3.1.3.2.2 and the correlation of basic defect and disease phenotype is investigated in chapter 3.1.3.2.3.

FIGURE 21: *CFTR* GENOTYPE — CF PHENOTYPE CORRELATION: A CONCEPT

CF disease severity viewed as a consequence of the CFTR genotype. Severe CF disease is considered to be the result of absent CFTR *function, while mild CF disease is surmised to result from residual* CFTR *activity.*



3.1.3.2.1 The molecular phenotype of mutant CFTR alleles

Generally, no functional CFTR protein is expected to arise from frameshift mutations, as for CFTRdelEx2,3 investigated in patient D. Most nonsense mutations – such as G542X, R553X, Y1092X, studied in patients A, B and C – have been associated with mRNA reduction and/or exon skipping (Hamosh *et al.* 1991, Hamosh *et al.* 1992b, Hull *et al.* 1994b, Will *et al.* 1995). However,

not all nonsense mutations can be considered to result in null phenotypes: wild type mRNA levels, possibly translated into a truncated CFTR protein, have been detected for R1162X (Rolfini *et al.* 1993, Will *et al.* 1995).

Like most *CFTR* alleles, ΔF508-CFTR has been characterised as a protein with defective maturation (Cheng *et al.* 1990). Secondary missense mutation R553Q or R553M or R555K have been shown to revert this maturation defect (Teem *et al.* 1993, Teem *et al.* 1996). The Patients E, F and G carried one ΔF508 allele with a second missense mutation.

Rare *CFTR* missense mutations have been investigated in heterologous expression systems with respect to the processing of the protein and its electrophysiological properties. Most investigated missense mutations have been shown to result in a maturation defect or altered chloride conductance properties (Smit *et al.* 1995, Seibert *et al.* 1996a, Seibert *et al.* 1996b, Seibert *et al.* 1997, Vankeerberghen *et al.* 1998a, Vankeerberghen *et al.* 1998b). Milder CF phenotypes, characterised for example by pancreatic sufficiency, were assumed to correlate with residual activity (Champigny *et al.* 1995, Sheppard *et al.* 1995), correct processing (Seibert *et al.* 1996a) or unimpaired interaction with the CFTR regulated ORCC (Fulmer *et al.* 1995). However, not all mutations classified as correlating with pancreatic sufficiency show cAMP stimulated chloride conductance (Fanen *et al.* 1997). The relation between defective maturation and chloride conductance phenotype is unresolved, as most mutant CFTR proteins are processed inefficiently and most CFTR mutants exhibit chloride conductance to some degree when localised correctly. Different functional properties for CFTR proteins underlying the same degree of maturation impairment have been shown for CFTR-R1066H, CFTR-R1066C and CFTR-R1066L (Cotton *et al.* 1996). Mutations expected to belong in the category of mutant CFTR proteins with altered conductance or processing were investigated in patient H and I in the homozygous condition.

CFTR mutations which generate or delete mRNA splicing consensus sequences have been shown to give rise to incorrectly spliced mRNA (Hull *et al.* 1993, Zielenski *et al.* 1993). However, reduced levels of correctly spliced mRNA have been shown for some nucleotide substitutions affecting less conserved positions distinct to the splice site consensus sequence (Highsmith *et al.* 1994, Highsmith *et al.* 1997). Consequently, wild type CFTR protein might be formed in those cases, associated with a milder CF phenotype. Patients J, K, L, M and N are carrying mutations which are known (J, K, L) or expected (M, N) to yield reduced amounts of wild type mRNA.

Defective splicing, leading to the insertion or omission of single or multiple exons for a part of the CFTR mRNA, has also been observed in non-CF tissues (Chu *et al.* 1991, Slomski *et al.* 1992, Melo *et al.* 1993, Yoshimura *et al.* 1993, Hull *et al.* 1994a). The most prominent alternatively spliced transcript results from exclusion of exon 9 (Chu *et al.* 1991), which gives rise to non-functional CFTR protein (Strong *et al.* 1993). The amount of misspliced mRNA was shown to correlate with the length of a preceding $(TG)_mT_n$-repeat (Chu *et al.* 1993, Cuppens *et al.* 1998). The condition $(TG)_mT_5$, termed IVS8-6(5T) (Chillon *et al.* 1995), was shown to be overrepresented in several disease conditions exhibiting symptoms of CF or atypical CF (Chillon *et al.* 1995, Pignatti *et*

al. 1996, Friedman *et al.* 1997, Sharer *et al.* 1998) and hence, the allele IVS8-6(5T) is considered as a risk allele. Two patients (O, P) bearing the IVS8-6(5T) allele on one chromosome, were investigated.

3.1.3.2.2 Correlation of CFTR mutation genotype and basic defect phenotype

The relation of *CFTR* mutation and basic defect phenotype is discussed in four subsections in which patients are grouped according to whether their CFTR genotype is expected to result in no CFTR protein, some mutant CFTR protein or reduced amounts of wild-type CFTR protein. Thereafter, the few cases where at least one CFTR allele appeared to be functional according to the *CFTR* mutation analysis on the genomic level, are evaluated on a case-to-case basis.

3.1.3.2.2.1 No CFTR protein expected: patients A, B, C, D

Consistent with the hypothesis of nonsense-mutations resulting in functional knock-outs, all patients homozygous for stop-mutations had highly elevated sweat chloride concentrations and none showed residual chloride secretion in the intestinal tissue. However, in three out of the four cases evidence for residual secretion was seen by NPD in respiratory tissue. This points to the existence of non-CFTR chloride channels in airway tissue in these patients. Residual chloride conductances mediated by non-CFTR channels in patients homozygous for a nonsense mutation have been described before: two patients homozygous for G542X have been reported as showing residual secretion in rectal tissue (Veeze *et al.* 1994).

3.1.3.2.2.2 Expression of mutant CFTR protein possible: patients E, F, G, H, I

Three patients, bearing a second missense mutation on the ΔF508 allele, all showed sweat chloride concentrations within the pathological range. No evidence for residual chloride secretion was detected in nasal tissue, but clear evidence for residual secretion was detected by ICM (figure 18). However, the chloride secretory component was sensitive to DIDS, indicating that this current is not CFTR mediated but due to alternative, Ca^{2+} activated channels. DIDS has been shown to inhibit CFTR only when having access to the intracellular, cytosolic domains of the protein (Linsdell *et al.* 1996). For patients H and I, who are homozygous for a missense mutation or an in-frame deletion, CFTR activity cannot be excluded on the basis of the *CFTR* genotype. However, both patients had highly elevated sweat chloride concentrations indicative of a lack of CFTR mediated chloride transport in the sweat duct. Residual chloride secretion activity was seen in both patients, either by NPD and ICM (patient H) or only in rectal tissue (patient I) upon stimulation with cAMP. This discrepancy might reflect the opposing roles of CFTR in the sweat gland (reabsorption of chloride) compared with nasal and rectal tissue (chloride secretion), indicating that alterations like E92K in patient H impair the role of CFTR in reabsorption, but not in secretion of chloride. Alternatively, the

residual secretion observed by ICM or NPD might be mediated by an alternative channel not expressed in the sweat gland. For patient I, a cAMP activated chloride conductance was seen in the absence, but not in the presence, of DIDS indicating its origin by alternative, non-CFTR chloride channels.

3.1.3.2.2.3 Low amounts of wild type CFTR expected: patients J, K, L, M and N

All patients with splice mutations showed residual chloride secretion in nasal or rectal tissue: in NPD, four out of the five patients showed a basal PD in the borderline or non-CF range, indicating residual chloride secretion. Furthermore, a response to superfusion with chloride-free solution was seen in three of the four patients. Apart from the elevated sweat chloride concentrations in patients K, L and N, the data obtained by NPD and ICM for the patients J, K, L and N are consistent with the hypothesis of CFTR-mediated residual chloride conductance. For patient M, a non-CFTR mediated chloride conductance can be surmised: the highly elevated sweat chloride concentration, the CF-typical basal PD and the lack of residual secretion in nasal tissue points to defective CFTR in this patient, being compound heterozygous for 3850-3 T⇨G. Even though this mutation is affecting the not obligatory conserved –3 position and thus has been presumed to yield more wild type CFTR transcript than alleles bearing mutations at the highly conserved positions –2 and –1 (Dörk *et al.* 1993), no evidence for CFTR mediated chloride conductance was found in patient M. The cAMP stimulated chloride conductance in rectal tissue of patient M was seen in the absence, but not in the presence, of DIDS and hence is most likely not CFTR-mediated. An inactivation of a –3 splice site has also been shown for 297-3 C⇨T (Bienvenue *et al.* 1994).

3.1.3.2.2.4 Normal amounts of wild type CFTR expected: patients O, P, Q, R and S

Five patients suspected of carrying at least one wild type *CFTR* allele have been investigated in three tissues with respect to the basic defect. In two out of the five cases, none of the tissues examined showed a pattern typical for CF (patients O and P). A pathological condition restricted to one of the three tissues investigated was observed in two patients (Q and S). Only for one out of the five cases with unresolved *CFTR* genotypes were conditions consistent with the diagnosis CF seen in nasal tissue, rectal tissue and the sweat gland (patient R).

Heterozygosity for IVS8-6(5T) and non-CF condition in three tissues (patients O and P)

Based on the basic defect measurements, the diagnosis of CF was rejected in patients O and P even though patient O was compound heterozygous for a mutation affecting a splice site at the conserved –2 position and the IVS8-6(5T) variant. Patient P, whose *CFTR* alleles were screened throughout the complete coding sequence (table 10[11]), likewise carried the IVS8-6(5T) variant but except for two nucleotide exchanges encompassing exon 14b (described in detail in chapter 3.1.1.2), patient P

11 Table 10 is shown in chapter 3.1.2.5 on page 67.

did not carry any sequence variations. Presumably, the alterations of introns 14a and 14b were without consequence for the transcription of the *CFTR* gene as they did not introduce a novel splicing signal nor did they destroy the splice site consensus sequences of exon 14a. Hence, when comparing the *CFTR* alleles of patients O and P under the assumption that all sequence variations have been detected, patient P can be considered as heterozygous for a wild type allele and the IVS8-6(5T) variant and patient O can be considered as heterozygous for a disease-causing lesion and the IVS8-6(5T) variant. Though IVS8-6(5T) is established as a sequence alteration modulating the penetrance of a mutant *CFTR* allele (Kiesewetter *et al.* 1993) and can be considered as a risk allele for pulmonary disease (Pignatti *et al.* 1996, Friedman *et al.* 1997), CBAVD (Chillon *et al.* 1995) and pancreatitis (Sharer *et al.* 1998), neither the isolated IVS8-6(5T) condition nor compound heterozygosity with a *CFTR* disease-causing lesion was sufficient to transmit the basic defect typical for CF to nasal or rectal tissue, as demonstrated by patients P and O, respectively.

Carriership for a disease-causing lesion and CF condition in nasal tissue (patient Q)

Patient Q showed a CF-typical NPD but no evidence for defective CFTR in rectal tissue or the sweat gland. All *CFTR* exons and flanking intron sequences were screened by SSCP (table 10) and the *CFTR* gene was investigated for large genomic alterations by PFGE, but only the splice site consensus transition 1898+3 A⇨G was identified on one chromosome. This sequence variation can be considered to be a disease-causing lesion since patient N, being homozygous for this allele, exhibited clinical symptoms and showed an impaired chloride conductance in nasal tissue and the sweat gland compatible with CF. The carrier status of patient Q was sustained by the family history: a sibling died presumably of meconium ileus and CF and thus, the *CFTR* alleles within the family of patient Q were most likely associated with pancreatic insufficiency. However, patient Q was pancreatic sufficient and hence, carriership of patient Q seems more likely than the condition of her carrying two *CFTR* mutant alleles causing a severe gastrointestinal phenotype.

The tissue-dependency of the CF typical chloride conductance defect seen in patient Q can be explained assuming an impaired regulation of CFTR expression in the respiratory epithelium, mediated by tissue specific regulatory elements. The manifestation of the basic defect is probably eased by the consanguineous background yielding an excess of homozygous genotypes at many loci throughout the genome, including sequence variants with an unfavourable effect.

No disease-causing lesion identified and CF condition in sweat gland (patient S)

Patient S showed highly elevated sweat electrolytes but no evidence for defective chloride conductance was seen in ICM or NPD. Screening of all *CFTR* exons and flanking intron sites did not reveal any disease-causing lesion (table 10[12]) and consequently, patient S was suspected of carrying two wild type *CFTR* alleles. Furthermore, patient S shares the *CFTR* spanning haplotype with his asymptomatic sister and hence, any disease-causing lesion missed by screening the *CFTR* gene of patient S should be present in the patient's sister unless a de novo mutation has occurred.

12 Table 10 is shown in chapter 3.1.2.5 on page 67.

As for patient Q, the chloride conductance defect observed in the the sweat gland in patient S, but not by ICM or NPD, could be explained by tissue specific regulatory elements being defective in patient S but unimpaired in his sister. Another hypothesis involves a defect in the transport of the counterion Na^+ via the amiloride-sensitive Na^+ channel ENaC: as in CF, the condition pseudohypoaldosteronism type I can lead to elevated sweat chloride and sodium values (Hanakoglu *et al.* 1994, Marthinsen *et al.* 1998). The ENaC coding genes SCNN1A, SCNN1B and SCNN1G have been shown to cause pseudohypoaldosteronism (Chang *et al.* 1996, Strautnieks *et al.* 1996a). Within the family of patient S, the polymorphisms D12S889 — located near SCNN1A (Baens *et al.* 1995) — and β-ENaCGT — located in in SCNN1B (Shimkets *et al.* 1994) and near SCNN1G (Voilley *et al.* 1995) — have been investigated. Patient S and his sister share at both loci only one allele. Thus, it is possible that different alleles for the α, β and γ subunits of the ENaC were inherited by the siblings, presumably related to different Na^+ conductance properties of the resulting channel.

No disease causing lesion identified and CF condition in three tissues (patient R)

The investigation of patient R clearly demonstrated CF conditions in the rectal and nasal tissue and a sweat test consistent with CF. Thus, an impairment of the CFTR protein has to be assumed and accordingly, a defect of the *CFTR* gene, a regulatory element in *cis* or a regulating element acting in *trans* has to be inferred. Both *CFTR* alleles of patient R have been screened extensively by PFGE for large genomic alterations and by SSCP for small sequence variations in all *CFTR* exons and flanking intron sites (table 10) including 4 kb of sequence preceding exon 1. The 4 kb sequence exceeds the characterised promoter region (Chu *et al.* 1991, Koh *et al.* 1993). Though sequence alterations within the *CFTR* promoter sequence have been described (Bienvenue *et al.* 1995), they seem to be rare: extensive screening of interspecies conserved regions of the *CFTR* promoter in 205 patients with CF, atypical CF or CF-like conditions conditions without an identified mutation on both *CFTR* alleles failed to reveal disease-causing lesions (Verlingue *et al.* 1998). Hence, unless both *CFTR* mutations were missed in patient R because the respective sequence was not screened for — as in the case of a sequence alteration generating a cryptic splice site as described for 3849+10kbC⇨T (Highsmith *et al.* 1994) or in case of a regulatory element located within an intron as described for intron 1 (Smith *et al.* 1996) — the molecular lesions causing the lack of CFTR related conductance measured in patient R have to be located in *trans*, presumably mediating the expression of the *CFTR* gene through a regulatory element. A defective regulation of CFTR expression of at least one allele can be considered equivalent to a splice site mutation that allows the generation of minor proportions of wild type transcript and thus should lead to a basic defect pattern comparable to that displayed by patients J, K, L, M and N. Indeed, patient R showed residual chloride secretion in rectal tissue similar to patient K (homozygous for 3894+10kbC⇨T) and had a moderately elevated sweat chloride concentration.

3.1.3.2.3 Correlation of basic defect phenotype and severity of CF disease

Out of the 19 patients who participated in the study "*rare genotypes and atypical cystic fibrosis*", a non-conclusive ICM result was obtained for patients A and N. Patients O, P and S are not discussed within this chapter due to their non-CF condition displayed in nasal and rectal tissue. The remaining 14 patients were grouped with respect to their basic defect: no residual chloride secretion was seen in both, ICM and NPD, for one patient. Discordance when comparing nasal and rectal tissue was observed for 2 patients (residual in NPD only) and 8 patients (residual in ICM only). 3 patients displayed residual chloride secretion in both tissues. Within this chapter, three attempts were made to relate the clinical phenotype – as accessible by pancreatic status, wfh% and FEVPerc – of these 14 patients to the results of the basic defect measurements. Chapter 3.1.3.2.3.1 examines if discordant basic defect phenotype is reflected by the clinical status. In chapter 3.1.3.2.3.2, pancreatic sufficient and pancreatic insufficient patients are compared with respect to their ICM results. Chapter 3.1.3.2.3.3 discusses whether the patient's pulmonary status is reflected by the presence or absence of residual chloride secretion as measured by ICM or NPD.

3.1.3.2.3.1 Disease severity and expression of the the basic defect in 14 CF patients

No residual chloride secretion in nasal and rectal tissue

Among the 14 patients investigated with a conclusive result for ICM and NPD, only one patient did not show any evidence for residual chloride secretion in both tissues investigated with respect to the basic defect. Patient B was pancreatic insufficient, had a normal wfh%, but an impaired lung function.

Residual chloride secretion in nasal, but not in rectal tissue

In two patients, (C and D), residual chloride secretion was detected by NPD, but not by ICM. These patients all had a normal wfh%, an impaired pulmonary status and were pancreatic insufficient.

Residual chloride secretion in rectal, but not in nasal tissue

In seven patients, (E, F, G, I, J, M, and R), residual chloride secretion was seen in rectal tissue, but not in nasal tissue. Patient Q was likewise discordant comparing ICM and NPD: while the CF-condition was detected in nasal tissue, the rectal tissue displayed a non-CF condition. Patients E, F, G, I, J, M and R all had a normal wfh%. Patients E, F, G, I and M were pancreatic insufficient while patients J and R were pancreatic sufficient. An impaired pulmonary status was seen in patients F, G, J and M (severely affected) and patient R (moderately affected), but not in patients E and I. Patient Q was pancreatic sufficient, had a reduced wfh% and the pulmonary status was severely impaired.

Residual chloride secretion in nasal and rectal tissue

Three patients (H, K, L) showed residual chloride secretion by NPD and ICM. The patients were all pancreatic sufficient and had a normal wfh%. An unimpaired pulmonary status was seen in patient H, patient K received a lung transplant and patient L had a severely affected pulmonary status.

3.1.3.2.3.2 Pancreatic status and residual chloride secretion in rectal tissue

6 of the 14 patients discussed here were pancreatic sufficient (H, J, K, L, R and Q). In all cases, chloride secretion was detected by ICM. Considering the rectal tissue as representative of the gastrointestinal tract, the correlation between pancreatic status and residual chloride conductance is evident as described by Veeze *et al.* (1994). However, of the remaining 8 pancreatic insufficient patients (B, C, D, E, F, G, I and M), residual chloride secretion was detected by ICM in 5 patients. Interestingly, in four of these five cases, evidence was seen for a non-CFTR origin of the detected chloride conductance (patients E, F, I and M). In contrast, a chloride secretory pattern in ICM consistent with the hypothesis of CFTR mediated conductance was seen in patients H, J, K, L, R and Q. Hence, it is tempting to speculate that pancreatic sufficiency is not solely determined by any residual chloride secretion, but based on residual CFTR activity. This finding suggests that alternative chloride channels, detected by NPD and ICM, are not expressed in pancreatic tissue and hence do not protect against pancreatic disease. Sustaining this hypothesis, it was recently demonstrated that CLCA1, the Ca^{2+} activated chloride channel, is not expressed in pancreatic tissue (Gruber 1998b).

3.1.3.2.3.3 Pulmonary status and residual chloride secretion in nasal tissue

Various degrees of impairment of pulmonary function were observed regardless of the presence or absence of residual chloride conductance in nasal or rectal tissue (chapter 3.1.3.2.3.1). In six cases among the 14 patients discussed here, residual secretion was detected by NPD (patients C, D, H, K, L and Q). As shown in table 12, no correlation between FEV1%pred and presence of residual chloride conductance was observed. Likewise, no correlation of residual chloride secretion measured in rectal tissue and FEV1%pred was detected (table 13). This observation is consistent with the concept of high variability of the pulmonary status, described for patient groups carrying the same *CFTR* genotype (Kerem *et al.* 1990, Gasparini *et al.* 1992, Rozen *et al.* 1995, de Braekeleer *et al.* 1997).

The patients examined within the study "*rare genotypes and atypical cystic fibrosis*" represent a small selected panel of elderly patients. Thus, even though no correlation between FEV1%pred and the presence of residual chloride secretion in nasal or intestinal was noticed, the advanced age of most patients investigated indicates an advantage of residual secretion for the survival. Sustaining this hypothesis, the five patients (K, L, M, Q and R) who were older than 30 years on the day of investigation all showed residual chloride secretion detectable by ICM.

The complexity of CF lung disease is the subject of other chapters within this thesis wherein the results of the "*European CF Twin and Sibling study*" are presented: the influence of genetic vs epigenetic factors on CF lung disease is discussed in chapter 3.2.1. Modulation of CF lung disease by the candidate gene locus D12S889, close to the SCNN1A gene encoding the α subunit of the ENaC, is discussed chapter 3.2.3.8.1. The modulation of disease severity due to modulating factors within the *CFTR* linkage group is discussed in chapter 3.2.3.8.2.

TABLE 12: RESIDUAL SECRETION IN NASAL TISSUE AND PULMONARY STATUS

Pulmonary status, judged by FEV1%pred and the patient's age, and residual chloride secretion in ***nasal tissue.***

no residual chloride secretion in NPD PATIENT (age in years)
residual chloride secretion in NPD PATIENT (age in years)

FEV1%pred	PATIENT (age in years)				
> 100%	H (12 y)	I (16 y)			
80%-100%	E (22 y)	K (‡)			
60%-80%	B (23 y)	R (37 y)			
40%-60%	C (25 y)	D (18 y)	F (24 y)	G (17 y)	J (20 y)
< 40%	L (58 y)	M (36 y)	Q (32 y)		

‡ patient K received a lung transplant

Table 13: RESIDUAL SECRETION IN ICM AND PULMONARY STATUS

Pulmonary status, judged by FEV1%pred and the patient's age, and residual chloride secretion in ***rectal tissue.***

no residual chloride secretion in ICM PATIENT (age in years)
residual chloride secretion in ICM PATIENT (age in years)

FEV1%pred	PATIENT (age in years)				
> 100%	H (12 y)	I (16 y)			
80%-100%	E (22 y)	K (‡)			
60%-80%	B (23 y)	R (37 y)			
40%-60%	C (25 y)	D (18 y)	F (24 y)	G (17 y)	J (20 y)
< 40%	L (58 y)	M (36 y)	Q (32 y)†		

‡ patient K received a lung transplant

3.2 The European CF Twin and Sibling study

Within this chapter, the analysis of the clinical data (chapter 3.2.1), of the basic defect as accessible by intestinal current measurement (chapter 3.2.2) and of the genotype at selected candidate gene loci are presented (chapter 3.2.3).

3.2.1 Analysis of clinical phenotype

The data collected on twins and sibs were used to analyse the effect of sex, age difference, *CFTR* genotype, pancreatic status, monozygosity and colonisation with *Pseudomonas aeruginosa* on the severity of the disease (chapter 3.2.1.1). The parameters employed were weight for height (wfh%), percentiles for FEV1%pred (FEVPerc) and, derived thereof by combination, the quantity DfO[13].

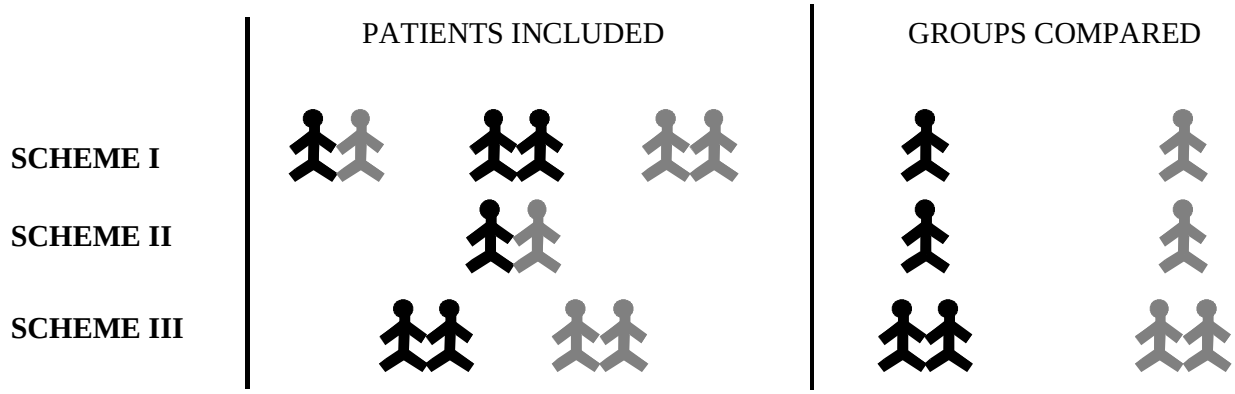
FIGURE 22: SCHEMES FOR STATISTICAL TESTS

Patients are symbolised as grey figure or black figure to denote the attribute analysed, e.g. gender (comparison of males and females) or pancreatic status (comparison of pancreatic sufficient and pancreatic insufficient patients).

PATIENTS INCLUDED*: pairs used for evaluation in scheme I, II and III*

GROUPS COMPARED*: patients compared using scheme I, II and III*

SCHEME I*: comparison of individual patients with different attributes regardless of kinship*
SCHEME II*: comparison of individual patients with different attributes within one pair*
SCHEME III*: comparison of pairs with different attributes*



Three different approaches were used to calculate the influence of an attribute (e.g. monozygosity status or pancreatic status) on the parameters (e.g. wfh% or FEVPerc): ignoring the relationship of the patients within their pairs, patients were sorted into two groups with different attributes (figure 22, scheme I). To evaluate the effect of the attribute on the parameter, a Mann-Whitney rank test was applied. Alternatively, only pairs composed of two sibs differing in the attribute studied were considered (figure 22, scheme II). A Wilcoxon rank test was used to evaluate the effect of the

13 Defined in figure 10 on page 48 as distance from origin of a data point representing a patient in the diagram of rank number of FEVPerc plotted against rank number of wfh%.

attribute on the parameter in this intrapair comparison. When grouping pairs composed of two patients with identical attribute (figure 22, scheme III), a parameter reflecting the intrapair variance was studied (e.g. the intrapair difference of wfh%). A Mann-Whitney rank test was employed for statistical analysis of the effect of a pair's characteristic on the intrapair difference in a parameter. In chapter 3.2.1.2, an approach is shown to estimate the number of modulating factors involved in shaping the disease phenotype. Non-parametric analysis was carried out according to:

E. Weber "Grundriss der biologischen Statistik" 8th ed. Fischer Verlag, Stuttgart; 1980

3.2.1.1 Non parametric analysis of disease severity and discordance

GENDER

Comparing all males and females from sib pairs regardless of their kinship (figure 22, scheme I), males and females did not differ in their wfh% (273 males, 270 females; $p = 0.46$) or FEVPerc (284 males, 278 females; $p = 0.34$). Within the group of monozygous twins (figure 22, scheme I), males had a better wfh (18 males, 18 females; $p = 0.05$) and a slightly better FEVPerc (18 males, 18 females; $p = 0.16$). The combined parameter DfO reached significance (18 males, 18 females; $p = 0.04$). An intrapair comparison of the disease severity for males and females was made (figure 22, scheme II) for all dizygous patient pairs composed of a male and a female sibling. Males and females were significantly different in wfh% and FEVPerc, but not in the combined parameter DfO. While females were significantly better in wfh% (203 pairs, $p = 0.05$), males were significantly better in FEVPerc (143 pairs, $p = 0.05$). Upon combining the parameters to DfO this crosswise relationship resulted in nonsignificance (137 pairs, $p = 0.8$). Table 14 summarises the findings on the influence of the gender in CF twins and sibs.

TABLE 14: INTERPAIR- AND INTRAPAIR INFLUENCE OF GENDER
*p-values classified as: $p > 0.05 \Rightarrow$ non-significant (*n.s.*); $p < 0.05 \Rightarrow$ significant (*sig*)*

SCHEME (figure 22)	GROUP	COMPARISON OF PATIENTS	*p*	PARAMETER	BETTER PHENOTYPE
I	ALL SIB PAIRS	INTERPAIR	n.s.	FEVPerc, wfh%	—
I	MONOZYGOUS TWINS	INTERPAIR	sig	DfO, wfh%	males
	— // —	INTERPAIR	n.s.	FEVPerc	males
II	SIB PAIRS	INTRAPAIR	sig	wfh%	females
	— // —	INTRAPAIR	sig	FEVPerc	males
	— // —	INTRAPAIR	n.s.	DfO	—

The influence of the patient's gender on the nutritional status, as represented by wfh%, seems inconsistent comparing the results of the intrapair- and interpair analysis: while males and females did not differ in wfh% when all twins and sibs were compared, irrespective of their kinship (interpair

comparison; figure 22, scheme I), females were significantly better in wfh% than males when only sibpairs of brother and sister were compared (intrapair comparison; figure 22, scheme II). In order to assess the influence of age, four classes of patients and patient pairs were defined (age below 7 years, age 7 to 12 years, age 12 to 16 years, age 16 to 21 years) and the statistical tests were repeated for each age class to allow a cross-sectional analysis.

FIGURE 23: CROSS SECTIONAL ANALYSIS OF GENDER INFLUENCE

The influence of gender on nutritional status (wfh%, top row) and pulmonary status (FEVPerc, bottom row) is quantified by the parameter **u** *obtained from the non-parametric test statistics.* **u** *depends on the sample size analysed and the number of patients or patient pairs used for the test are given in the tables on the right. Significance ($p<0.05$) is reached for* $|\mathbf{u}| < 1.65$ *(one-tailed test).*

Interpair-comparisons (employing scheme I, figure 22):
Intrapair-comparisons (employing scheme III, figure 22):

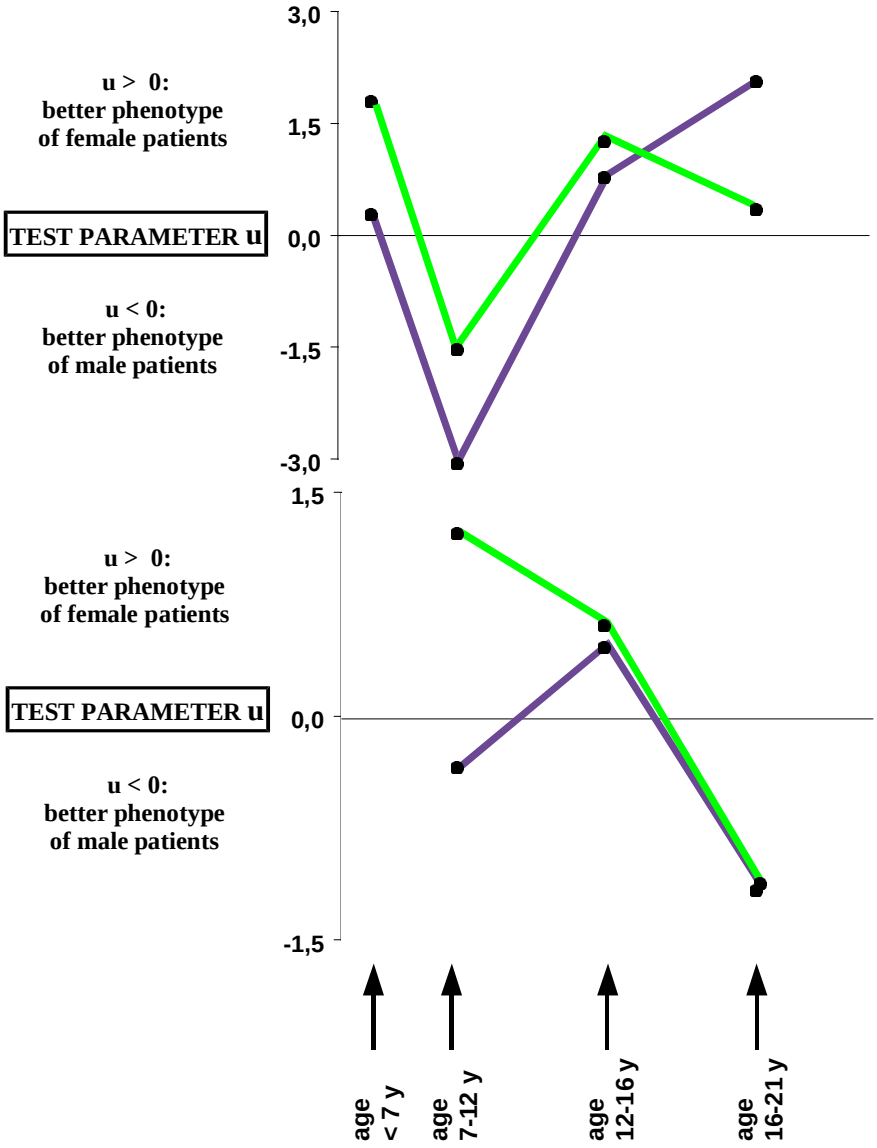
wfh%

age	no of patients		no of pairs
<7 y	m	112	58
	f	96	
7-12 y	m	111	45
	f	103	
12-16 y	m	80	30
	f	65	
16-21 y	m	82	36
	f	76	

FEVPerc

age	no of patients		no of pairs
7-12 y	m	67	39
	f	74	
12-16 y	m	67	26
	f	57	
16-21 y	m	80	34
	f	72	

The x-axis represents the age analysed. The following age-classes were defined:
<7y *(wfh% only)*
7-12 y
12-16 y
16-21 y



Results are displayed in figure 23: while males had a better wfh% than females within the age group of 7 to 12 years, the situation was reversed among patients with an age of 16 to 21 years. This observation is consistent with the data published by Morison *et al.* (1997), reporting an increased mortality among females in the age of 5 to 15 years, but a n increased stability of the weight of grown-up females compared to grown-up males. As displayed in table 14, both effects compensated each other when the complete group of patients (all ages) was analysed.

As observed for wfh%, the influence of gender on FEVPerc changed with the patient's age. A discrepancy between intrapair and interpair comparison of females in the age of 7 to 12 years was calculated. As the measurement of pulmonary function requires compliance and practice of the patient, the deviation might be due to the fact that siblings are trained for lung function at the same CF center and hence, brother and sister are performing equally while the results of unrelated males and females are subject to different errors.

It is a generally accepted fact that females with cystic fibrosis have a worse prognosis than males (Corey *et al.* 1996) characterised by poorer survival and poorer nutritional status (Corey *et al.* 1988). This effect could not be attributed to a different calorie uptake or energy management comparing CF males and females, but it was hypothesised that the onset of menarche and the change in the hormonal balance interferes with "immunological defence mechanisms in the lung of CF females facilitating progressive pulmonary involvement and subsequent mortality" (Corey *et al.* 1988). Even though the data presented here was not sampled to allow a longitudinal study, the cross sectional analysis of wfh% and FEVPerc in age groups provides evidence for this hypothesis: the influence of the gender on the nutritional status changes dramatically during puberty and correlates with a slightly delayed change in gender influence on FEVPerc, leading to an overall worse pulmonary status for female CF patients. This hypothesis is sustained by the impairment of CFTR function by progesterone and estradiol (Sweezy *et al.* 1996).

CFTR GENOTYPE

The influence of the *CFTR* genotype on the disease severity was evaluated by comparing ΔF508 homozygotes to carriers of other genotypes (table 15). When comparing all patients (figure 22, scheme I), ΔF508 homozygous patients were significantly worse in wfh% (933 patients, 332 thereof ΔF508 homozygous; $p = 0.01$) and FEVPerc (745 patients, 264 thereof ΔF508 homozygous; $p < 0.0001$). When testing the same hypothesis for monozygous twins, no difference between ΔF508 homozygotes and patients with other genotypes was observed (36 patients, 20 thereof ΔF508 homozygous; wfh: $p = 0.18$; FEVPerc: $p = 0.13$). The intrapair difference of wfh% and FEVPerc did not differ between ΔF508 homozygous pairs and pairs with other genotypes (figure 22, scheme III; wfh%: 404 pairs, 166 thereof ΔF508 homozygous; p = 0.11; FEVPerc: 356 patient pairs, 114 thereof ΔF508 homozygous; p = 0.22).

TABLE 15: INFLUENCE OF ΔF508 HOMOZYGOSITY
*p-values classified as: p > 0.05 ⇒ non-significant (*n.s.*); p < 0.05 ⇒ significant (*sig*)*

SCHEME (figure 22)	GROUP	COMPARISON	*p*	PARAMETER
I	SIB PAIRS	PATIENTS	sig	FEVPerc, wfh%
I	MONOZYGOUS TWINS	PATIENTS	n.s.	FEVPerc, wfh%
IIIa	ALL PAIRS	PAIRS	n.s.	INTRAPAIR DIFFERENCE IN FEVPerc, wfh%

AGE DIFFERENCE IN SIB PAIRS

All sib pairs were tested for differences of disease severity caused by the age difference of the two sibs (figure 22, scheme II). The older and the younger sib were not significantly different in wfh% (396 pairs, $p = 0.13$) or FEVPerc (282 pairs, $p = 0.80$).

PANCREATIC STATUS

An effect of the pancreatic status on the disease severity but not on the intrapair variation thereof was observed: pancreas insufficient patients were significantly worse in wfh% (figure 22, scheme I; 933 patients, 787 thereof pancreas insufficient; $p < 0.0001$) and FEVPerc (656 patients, 538 thereof pancreas insufficient; $p < 0.0001$). The intrapair differences of wfh% and FEVPerc did not differ statistically between pairs with two pancreas sufficient sibs and pairs with two pancreas insufficient sibs (figure 22, scheme III; wfh%: 426 pairs, 63 pairs thereof composed of two pancreas sufficient sibs; p = 0.33; FEVPerc: 307 pairs, 52 pairs thereof composed of two pancreas sufficient sibs; p = 0.23).

MONOZYGOSITY

The disease severity did not differ comparing patients derived from monozygous twin pairs and patients derived from dizygous patient pairs (figure 22, scheme I; wfh%: 933 patients, 83 thereof from monozygous twin pairs, $p = 0.27$; FEVPerc: 657 patients, 62 thereof from monozygous twin pairs, $p = 0.27$). When comparing monozygous and dizygous patient pairs, the intrapair difference was significantly lower for wfh% and almost significantly lower for FEVPerc (figure 22, scheme III; wfh%: 464 pairs, 27 thereof monozygous twins, $p < 0.0001$; FEVPerc: 328 pairs, 18 thereof monozygous twins, $p = 0.06$).

COLONISATION WITH Pseudomonas aeruginosa

Patients colonised with *Pseudomonas aeruginosa* had a significantly lower wfh% and FEVPerc than those without P.a. (figure 22, scheme I; wfh%: 933 patients, thereof 504 *P.a.* colonised, $p < 0.0001$; FEVPerc: 657 patients, 403 thereof *P.a.* colonised, $p < 0.0001$). This was not observed within the subgroup of monozygous twins where wfh% and FEVPerc were not significantly different between colonised and non-colonised patients (figure 22, scheme I; 36 patients, thereof 25 *P.a.* colonised; wfh%: $p = 0.20$, FEVPerc $p = 0.25$).

3.2.1.2 Estimation of the number of disease modulating factors

Within chapter 3.2.1.1, the following results can be interpreted as evidence for the influence of the genetic background on the CF disease phenotype:

(I) Monozygous twins are significantly more concordant than dizygous patient pairs. This is especially of interest because only half of the monozygous twin pairs are ΔF508 homozygous whereas the other half of the pairs displays other *CFTR* genotypes. Hence, a group of patient pairs, non-homogeneous with respect to the major disease-causing gene, displayed the highest degree of concordance among all groups of pairs analysed.

(II) Comparing groups of pairs sorted according to their *CFTR* genotype, no higher degree of concordance was observed when a group of ΔF508 homozygous patient pairs (being homogeneous with respect to their *CFTR* genotype) was compared to all other patient pairs (being non-homogeneous with respect to their *CFTR* genotype). Likewise, no higher degree of concordance was observed comparing pairs composed of two pancreatic insufficient patients (displaying predominantly severe *CFTR* mutation genotypes) to pairs composed of two pancreatic sufficient patients (displaying mostly mild *CFTR* mutation genotypes).

(III) Within most sib pairs, the time point of diagnosis is the same for both siblings and consequently, the older sibling is diagnosed at a later age. However, no significant difference between the disease severity of the older and the younger sib was observed.

In order to assess the influence of modulating effects on the concordance of CF twins and sibs, a quantitative description of the intrapair-concordance was defined: the difference of rank numbers between sibling A and sibling B of a patient pair was calculated for wfh% and FEVPerc. The resulting distributions of these **i**ntrapair-**r**ank **n**umber **d**ifferences (**IRND**), obtained for all patient pairs (non-homogeneous *CFTR* mutation genotype) or the subgroup of ΔF508 homozygous patient pairs (homogeneous *CFTR* mutation genotype) was than compared to the distribution of IRND expected for a group of unrelated couples.

The IRND distribution expected for a group of unrelated couples was deviated based on the following considerations: the probability of obtaining a specific rank number difference depends on the magnitude of the rank number difference "m" and the total number of pairs "n/2". n/2 pairs result in the rank numbers 1, 2, 3, 4, ·····, n-1, n. The rank number difference m=1 can be obtained by occupying rank numbers 1 and 2 or similarly rank numbers 2 and 3 or similarly rank numbers n-1 and n. This results in 2×(n-1) probable constellations to obtain the rank number n=1. In contrast, there is only one constellation to obtain the maximal rank number difference n-1, by occupying rank numbers 1 and n, respectively. The mean expected frequency f for any rank number difference m is therefore given by the normalised expression f = (n-m)/(n-1).

For comparison of expected and observed IRND distribution, classes of equally sized IRND were estimated based on the expected IRND distribution (minimal size of the IRND classes: expectancy value E = 10 couples per IRND class). The occupation of these IRND classes was than

calculated for the observed IRND distribution (CF twin and sibling pairs). Random fluctuations of the IRND class occupation were minimised by raising the size of the IRND classes (E = 20, 30, 40 couples expected per IRND class). Statistical comparison was done with a χ^2 test.

FIGURE 24: INTRAPAIR-RANK NUMBER DIFFERENCES FOR CF TWINS AND SIBS

The distribution of intrapair - rank number differences for CF twins and sibs was compared to a distribution of intrapair - rank number differences for a group of unrelated couples. See chapter 3.2.1.2 for details.

Parameter: i = *number of intrapair - rank number difference classes used for calculation*
E = *number of pairs expected within each of the n classes*

Deviation from expectancy value *: difference between the number of pairs expected within a class and the number of pairs observed within each class expressed as % of* E

intrapair rank number difference *was expressed as % of maximal rank number difference*

(a) *Distribution of intrapair - rank number differences in wfh% for all CF twins and sibs*
——— : i = 47, E = 10 ——— : i = 9, E = 50

(b) *Distribution of intrapair - rank number differences in FEVPerc for all CF twins and sibs*
——— : i = 32, E = 10 ——— : i = 7, E = 50

(c) *Distribution of intrapair - rank number differences in wfh% for all ΔF508 homozygous twins and sibs*
——— : i = 11, E = 10 ——— : i = 4, E = 30

(d) *Distribution of intrapair - rank number differences in FEVPerc for all ΔF508 homozygous twins and sibs*
——— : i = 11, E = 10 ——— : i = 4, E = 30

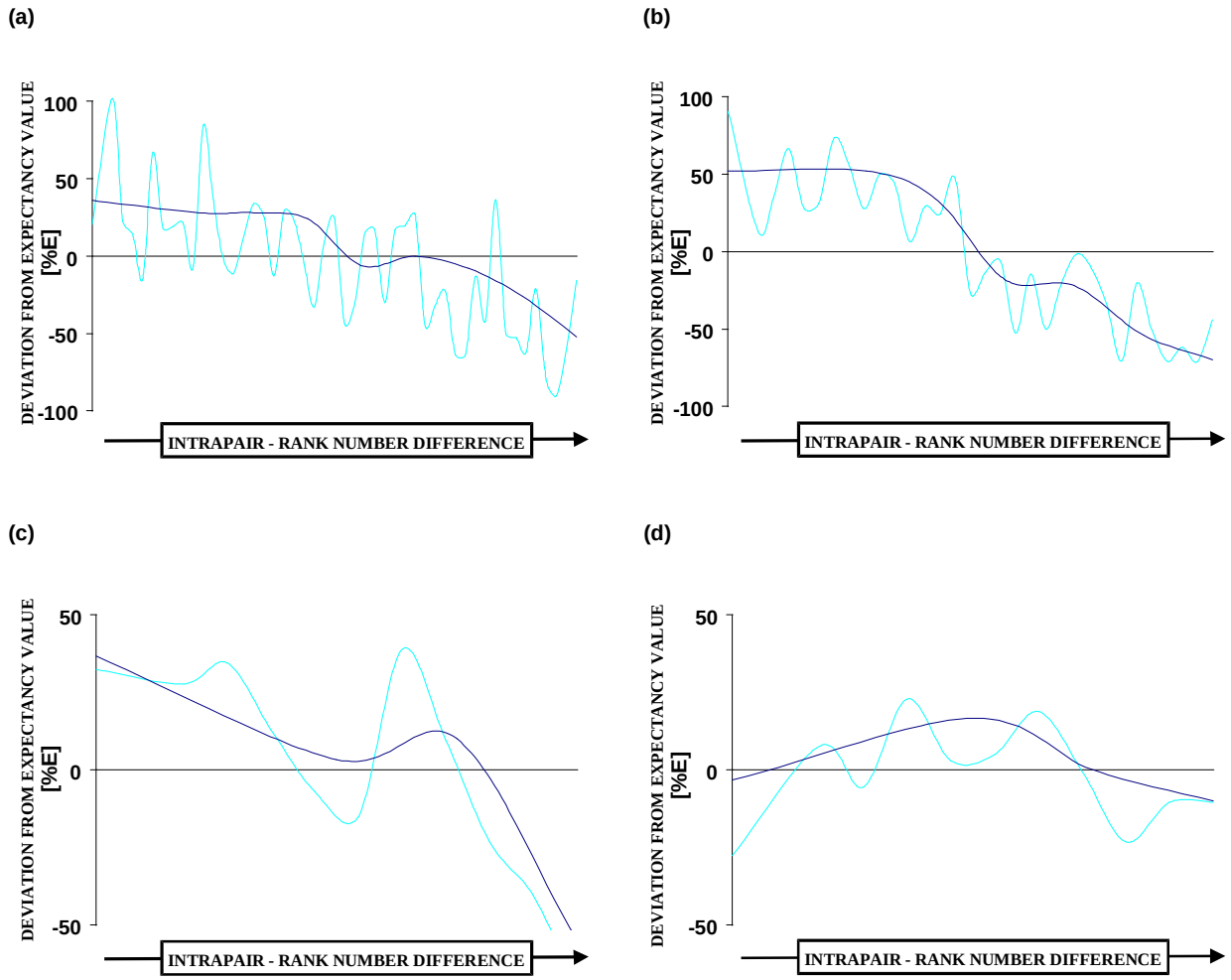


Figure 24 displays the results in normalised form: the x-axis represents the IRND, expressed as % of the maximal IRND n-1. The y axis represents the occupation of the IRND class, expressed as % of the expected occupation E. The IRND distribution for wfh% and FEVPerc rank numbers differs significantly when comparing all CF twins and sibs to a group of unrelated couples (wfh%: 9 IRND classes, p << 0.001, figure 24a; FEVPerc: 7 IRND classes, p << 0.001, figure 24b). Analysing the subgroup of ΔF508 homozygous patient pairs only, a significant difference between observed IRND distribution compared to a group of unrelated couples was obtained for wfh% (4 IRND classes, p = 0.025, figure 24c), but not for FEVPerc (4 IRND classes, p = 0.9, figure 24d).

CF twins and sibs carry identical mutations within the *CFTR* gene, live in a similar environment and share sequences from the genetic background. Hence, it can be expected that the patient pairs do not resemble a set of unrelated couples. When analysing pairs representing all *CFTR* genotypes, the difference between wfh% and FEVPerc IRND distributions of CF twins and siblings compared to a group of unrelated couples demonstrates the influence of the major affected gene on the intrapair concordance. In contrast, the *CFTR* gene cannot be held responsible for the deviation of IRND distributions when ΔF508 homozygous patient pairs are analysed. Any deviation of observed from expected IRND distribution is caused by shared – i.e. inherited – factors besides the *CFTR* gene itself.

Considering the results obtained on ΔF508 homozygous twins and siblings, the number of modulating factors detectable by this approach must be taken into account: when a class shows an excess of pairs compared to the expected number of couples, another class must in a likewise manner display less pairs than expected as both, the group of unrelated couples and the analysed group of patient pairs are equally sized by definition. Thus, the maximal number of independent modulating factors detectable due to an occupation of an IRND class deviating from the expectancy value is given by half of the number of IRND classes analysed. Hence, the nutritional status, represented by wfh%, is most likely to be regulated by a small number of modulating factors besides the *CFTR* gene – otherwise, no difference between the IRND distribution of ΔF508 homozygous twins and siblings compared to a group of unrelated pairs would have been observed. For the pulmonary status, a large number of modulating factors besides the *CFTR* gene can be estimated, as the IRND distribution of ΔF508 homozygous twins and siblings resembles the group of unrelated couples. While the concordance of monozygous twins in FEVPerc points to the existence of inherited factors, modulation of lung disease by epigenetic factors cannot be excluded. It seems most likely that inherited factors which determine the ability of an individual to deal with the environment are the basis of the variation observed.

In conclusion, the analysis of the intrapair variance of wfh% and FEVPerc has demonstrated an influence of shared factors on the nutritional status in CF while the modulation of the pulmonary status appears to be dominated by individual factors.

3.2.2 Intestinal current measurement

The basic defect in CF, i.e. the impaired chloride conductance of affected tissues, is accessible by short circuit current measurement of rectal suction biopsies (Veeze *et al.* 1991). In addition, residual chloride secretion has been observed among CF patients, predominantly when genotypes associated with a milder course of the disease have been analysed (Veeze *et al.* 1994). Within the "*European CF twin and sibling study*", patients homozygous for ΔF508 have been selected for investigation if they express an extreme phenotype with respect to their nutritional status, their pulmonary status or the intrapair concordance of one or both parameters (see chapter 2.2.1 for details on the selection of patient pairs).

Carbachol has been demonstrated to stimulate the basolateral and apical K+ channels through an increase of the intracellular Ca^{2+} concentration (Schultheiss and Diener 1997). Cl^- secretion is then supported by the hyperpolarization of the cell membrane. In CF patients, the lack of CFTR mediated chloride secretion was shown to result in a net K^+ efflux ("reversed response"), which can be followed by a small current compatible with Cl^- secretion ("residual response", Veeze *et al.* 1994). Figure 25 displays the intrapair variation of carbachol-induced signals for the investigated patient pairs. Signals are given as PDH% (defined in figure 15, page 59) representing the reversed proportion of the response with PDH% = 100 corresponding to a signal composed of only the K+ secretory response. In the plot each data point represents one patient pair, by assigning the PDH% value obtained for one sib on the x axis and by assigning the PDH% value obtained for the other sib on the other axis.

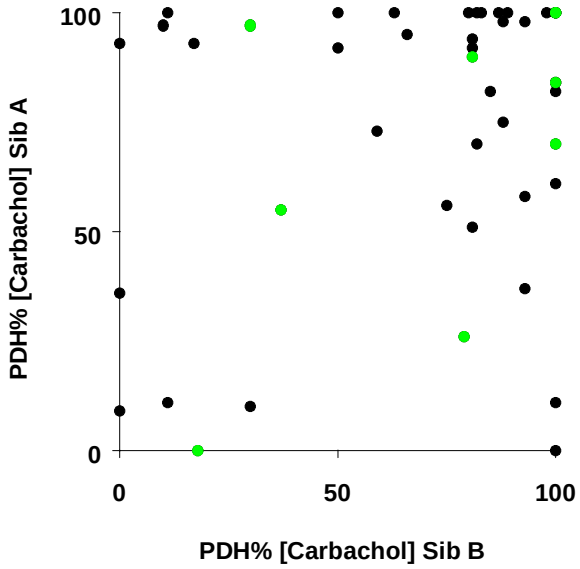
Pairs who are concordant with respect to their secretory response are to be found near the 45° diagonal axis in figure 25. Monozygous twins, being genetically identical, are expected to be concordant unless environmental factors or the experimental variance obscure the result. Estimating an error of ± 20% PDH, monozygous twins are found in 6 out of 8 pairs in a concordant status. For the remaining two cases, the influence of epigenetic factors may have to be taken into account. Indeed, one of the two monozygous twins discordant with respect to ICM is likewise discordant in clinical state and general appearance as these twins do not look monozygous, but resemble rather a sib pair.

Non-monozygous patient pairs demonstrated various combinations of a residual component and the intrapair variation thereof. Interestingly, for $^1/_3$ of the ΔF508 homozygous patients, residual secretion could be observed. This is consistent with the analysis of the shape types (table 4[14]), likewise indicating the presence of residual chloride secretion in $^1/_3$ of the patients.

14 Table 4 is shown in chapter 2.2.2.2.2 on page 58.

FIGURE 25: INTRAPAIR VARIATION OF CARBACHOL-INDUCED SIGNALS

PDH% *as defined in figure 15, page 59 for Sib A and Sib B of a pair (●). ● : monozygous twin pair.*

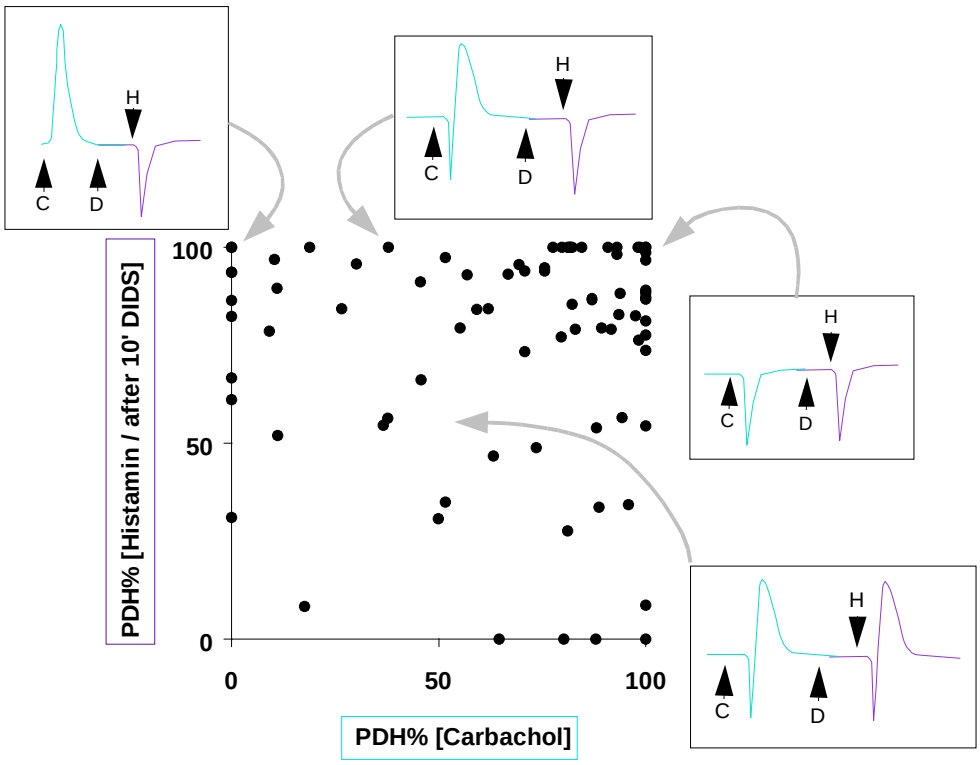


In cases of discordant residual secretion within a twin or sib pair, factors other than the *CFTR* gene – shared by both sibs – have to attribute for the modulation of the basic defect. Nevertheless, discordance might be assigned to the CFTR protein, as a modulation of CFTR expression mediated by a factor acting in *trans* could lead to a discordant CFTR expression when comparing two sibs.
As described in chapter 2.2.2.2, the protocol employed makes an attempt to discriminate between residual currents with respect to their DIDS-sensitivity. CFTR is considered DIDS-insensitive unless DIDS is added to the intracellular side (Lindsdell and Hanrahan 1996). In contrast, the Ca^{2+} regulated chloride channel CLCA1 was shown to be sensitive to DIDS (Gruber *et al.* 1998b). To allow comparison of secretory currents in the presence and the absence of DIDS within one tissue sample, the signals obtained upon stimulation of the biopsy by carbachol (without DIDS) and histamine (after an incubation with DIDS) are compared. However, carbachol and histamine are considered to act through non-identical pathways: both carbachol and histamine have been demonstrated to raise the intracellular Ca^{2+} levels. But, in contrast to histamine, the stimulation with carbachol could not be repeated immediately due to elevated levels of an inositol tetrakis phosphate species with an inhibitory effect (Vajanaphanich *et al.* 1994).

In figure 26, the effect of DIDS on the residual current is shown. Within the plot, each data point represents a patient. No effect of DIDS on the residual component was seen for all cases plotted near the 45° diagonal axis. However, several points within the upper left corner demonstrated a Cl^- conductance sensitive to DIDS indicating the presence of a non-CFTR mediated conductance in these patients.

FIGURE 26: DIDS SENSITIVITY OF RESIDUAL CURRENTS

Comparison of reversed proportion of carbachol- and histamine stimulated signals after 10' incubation with DIDS. PDH% as defined in figure 15, page 59. The pictograms show idealised tracings representing the data points indicated by the arrows. C*: carbachol,* D*: DIDS;* H*: histamine*



In summary, residual Cl^- secretion was detected among ΔF508 homozygous patients. This conductance was sensitive to DIDS in a subset of these patients indicating the presence of a non-CFTR mediated conductance, possibly based on the Ca^{2+} regulated chloride channel CLCA1.

3.2.3 Analysis of selected candidate genes

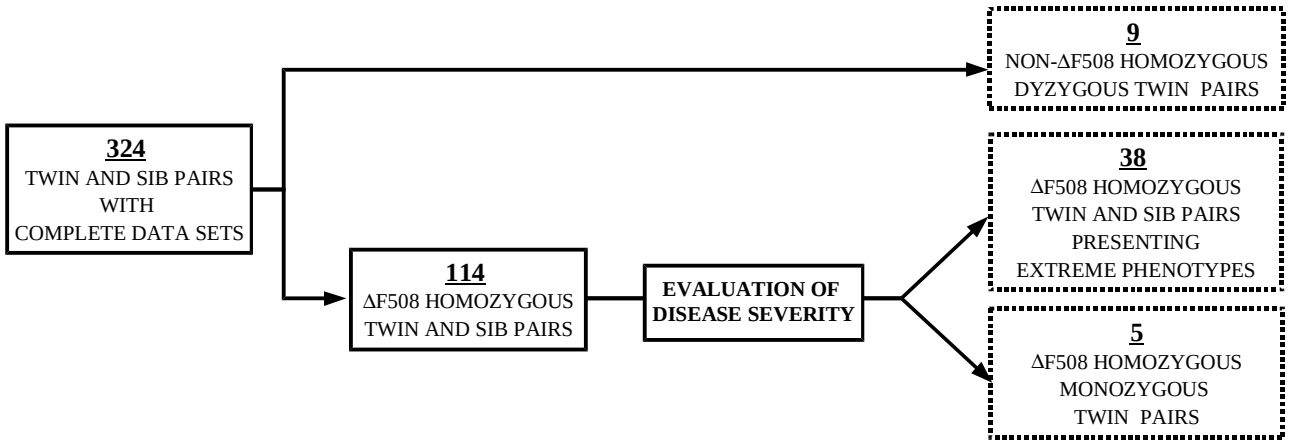
3.2.3.1 Selection of patient pairs for genotyping

The phenotype of ΔF508 homozygous patient pairs was analysed in order to identify pairs with an extreme course of the disease. Patients were ranked according to their nutritional status – represented by wfh% – and their pulmonary status – represented by percentiles for FEV1%pred – and pairs were selected based on non-average properties for the nutritional status, the pulmonary status or the intrapair discordance of one or both parameters (chapter 2.2.1).

The quantitative description of both the intrapair discordance and the disease severity enabled the ranking of pairs in sequences expressing the trait *"discordance"* (DIS) or the combined traits

"concordant/severe disease" (CON-) and *"concordant/mild disease"* (CON+). 114 ΔF508 homozygous patient pairs were ordered in each sequence and all pairs with rank numbers lower than 20 (concordant pairs) or lower than 25 (discordant pairs) were selected for genetic analysis[15]. Pairs ranked simultaneously in two conflicting categories were categorised separately as "discordant and concordant mild disease" (DC) and "non-discordant" (ND) as described in chapter 2.2.1.4. In total, 38 non-monozygous ΔF508 homozygous patient pairs, 19 thereof with both parents sampled, were selected. In addition, 5 monozygous ΔF508 homozygous twin pairs and 9 dizygous non-ΔF508 homozygous twin pairs, 3 thereof with both parents sampled, were typed (figure 27).

FIGURE 27: SELECTION OF PAIRS FOR GENOTYPING



3.2.3.2 Selection of markers near candidate gene loci

With an increasing number of genes mapped during the process of the human genome project, a candidate gene approach becomes feasible in accessing a gene's role in shaping the course of a disease. For CF, several groups of genes with examples listed in table 16 can be considered:

(a) ion channels having the capability of altering the CF basic defect
(b) gene products which might be involved in processing of CFTR protein as most CFTR mutants have been described to be defective in maturation in *in vitro* systems
(c) gene products involved in determining the major clinical manifestations
(d) gene products which maintain the integrity of polarised, CFTR expressing epithelial cells
(e) gene products which mediate the response to therapeutic drugs.

Within this thesis, results on the candidate gene ENaC are presented (chapter 3.2.3.4). The role of the obesity-related gene LEP on chromosome 7 is addressed by investigating a 40 cM region encompassing *CFTR* (chapter 3.2.3.5).

15 based on DNA availability by December 1997

TABLE 16: EXAMPLES FOR CANDIDATE GENES

- ***gene products with the capability of altering the basic defect***
  - alternative Cl^- channels
    - outwardly rectifying chloride channel ORCC
    - outwardly rectifying depolarisation induced chloride channel (ORDIC)
    - Ca^{2+} activated chloride channel (CLCA1)
  - channels interacting with CFTR
    - amiloride-sensitive Na^+ channel ENaC
- ***gene products involved in processing maturing CFTR proteins***
  - chaperones: HSP70, HSP90, calnexin
- ***gene products involved in determining the major clinical manifestations***
  - modulation of nutritional status / obesity related genes, e.g. LEP
  - modulation of pulmonary status
    - gene products involved in host defence
      - against chemicals / xenobiotics
        - glutathione S transferases (GST-Genes)
        - cytochrome P450 Enzymes (CYP-Genes)
      - against bacterial pathogens: HLA
      - antiproteases
- ***gene products maintaining integrity of the epithelium***
  - cell differentiation, e.g. hepatocyte growth factor HGF
- ***gene products mediating response to therapeutic drugs***
  - e.g. response to bronchodilators: β_2 adrenergic receptor

TABLE 17: LOCI NEAR CANDIDATE GENES ANALYSED WITHIN THIS THESIS

TEST OF POLYMORPHIC LOCUS FOR ALLELIC ASSOCIATION

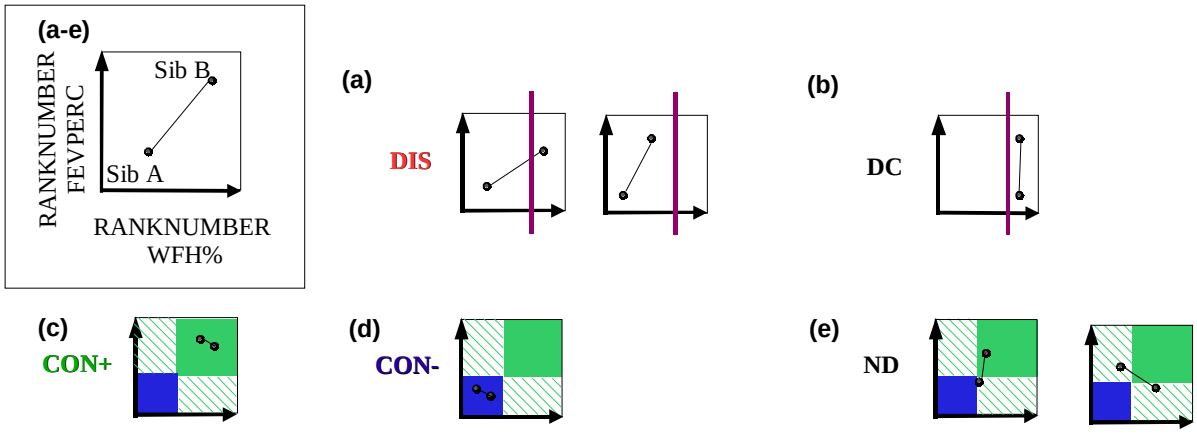
	Marker	**Candidate Gene**
CHAPTER 3.2.3.4	D12S889 βENaCGT	ENaC
CHAPTER 3.2.3.5	PON2 D7S525 MetH J3.11 D7S514 D7S495	Linkage group/obesity
CHAPTER 3.2.3.6	D2S1788	POMC/obesity

CORRELATION OF GENOTYPE TO CLINICAL PHENOTYPE

CHAPTER 3.2.3.7.1	D12S889
CHAPTER 3.2.3.7.2	D7S525, MetH, J3.11, D7S514, D7S495
CHAPTER 3.2.3.7.3	D2S1788

FIGURE 28: PAIRS SELECTED FOR GENETIC ANALYSIS: INTRAPAIR DIFFERENCES

Within this figure, each patient pair is represented by a plot of the ranknumber for wfh% on one axis and ranknumber for FEVPerc on the other axis. This diagram was used to calculate the parameters DfO and DELTA quantifying the disease severity and the intrapair-discordance, respectively (see figure 10c, page 48). The categories DIS, CON+, CON-, ND and DC were defined based on the rank numbers for the respective pairs which were displayed in table 1, page 52 (see also in figure 12 on page 53 or refer to chapter 2.2.1.4 for details on the ranking procedure). The following pictograms are representative for the five categories:



PATIENT PAIRS SELECTED FOR GENOTYPING

Colour coding is used to emphasise the pairs which were ranked discordant (DIS), concordant/mild disease (CON+) and concordant/severe disease (CON-) as introduced in figure 2 on page 26.

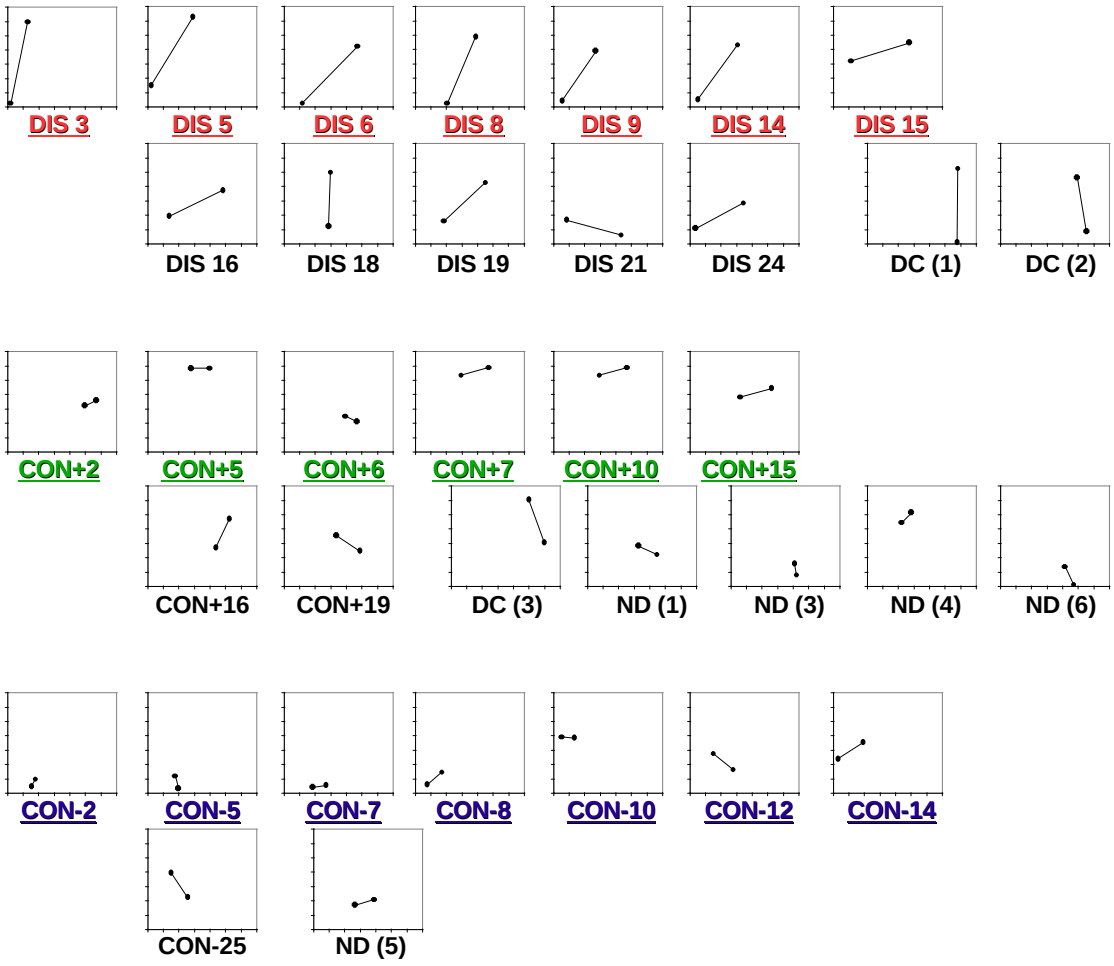


FIGURE 29: PAIRS SELECTED FOR GENETIC ANALYSIS: CLINICAL DATA

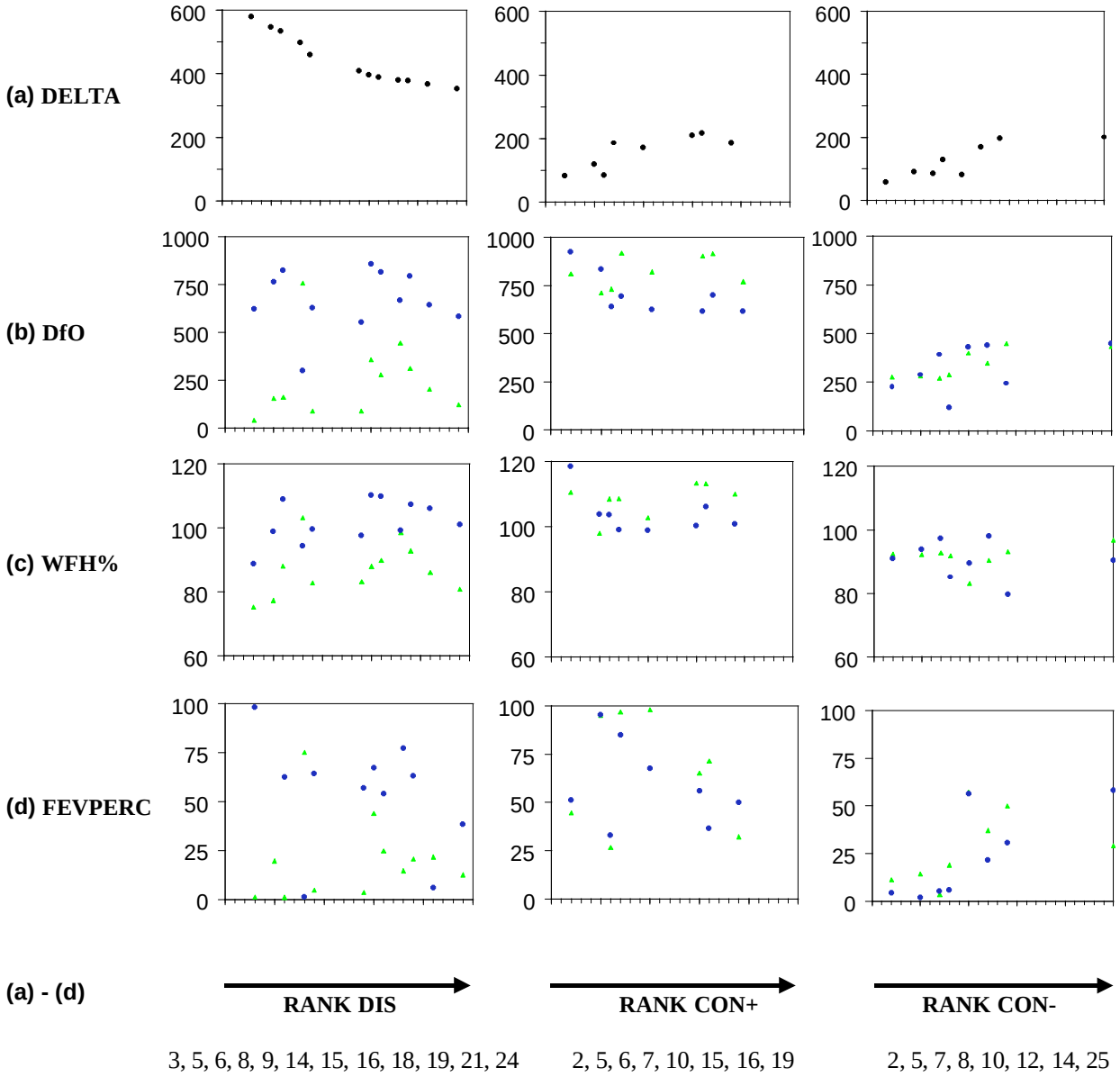
Figures (a) - (d) display DELTA (the composite parameter quantifying discordance, see figure 10c, page 48), DfO (the composite parameter quantifying disease severity, see figure 10c, page 48), wfh% (representing the nutritional status) and FEVPerc (representing the pulmonary status) for the patient pairs ranked DIS, CON+ and CON-. Within each diagram, rank number decreases from left to right, hence the pairs fulfilling the criteria best are found closest to the y-axis. Clinical data for the pairs with the following rank numbers are displayed:

DIS: 3, 5, 6, 8, 9, 14, 15, 16, 18, 19, 21, 24

CON+: 2, 5, 6, 7, 10, 15, 16, 19

CON-: 2, 5, 7, 8, 10, 12, 14, 25

(a), (b) composite parameters as defined in chapter 2.2.1.3
(a) DELTA describing the discordance (b) distance from origin (DfO) describing the severity of disease
(c) weight for height (wfh%) (d) percentiles for FEV1%pred (FEVPerc)
(b), (c), (d): ● Sibling A of pair; ▲ Sibling B of pair



3.2.3.3 Evaluation of genotype data of pairs in ranked in cohorts

Rank numbers for all selected non-monozygous ΔF508 homozygous patient pairs were shown in table 1[16]. The diagrams in figure 28 represent each pair used in candidate gene analysis by displaying the clinical parameters wfh% and FEVperc as defined for evaluation of the composite parameters DELTA (describing the intrapair discordance by the distance of both patients to each other within that diagram, figure 10c[17]) and DfO (describing the disease severity as the distance from origin in that diagram; figure 10c). Within each of the categories DIS, CON+ and CON-, pairs are arranged by decreasing rank number corresponding to a decline in the representation of *intrapair discordance* (sequence DIS), *disease mildness and intrapair concordance* (sequence CON+) and *disease severity and intrapair concordance* (sequence CON-). The decrease in DELTA within the sequences of pairs ranked DIS and the increase in DELTA within the sequences of pairs ranked CON+ and CON- is shown in figure 29a. As expected, siblings from discordant pairs (DIS) differ markedly in wfh%, FEVPerc and DfO while siblings from concordant pairs (CON+ and CON-) are comparable in all three parameters (figure 29b, c, d). The groups of pairs scored "concordant" (CON+ and CON-) and "discordant" (DIS) pairs differ completely with respect to the parameter DELTA : none of the selected pairs ranked CON+ or CON- have values of DELTA larger than 300 while all selected pairs ranked discordant have values of DELTA larger than 300. Likewise, the groups of pairs with mild disease (CON+) and severe disease (CON-) differ completely with respect to DfO: none of the siblings from pairs ranked CON- have values of DfO larger 500 than while all of the siblings from pairs ranked CON+ have values of DfO larger than 500. Hence, the selected patient pairs compose three groups with extreme and non-overlapping phenotypes.

Data on DNA polymorphisms was analysed by an association study: alleles were counted for the extreme ranking pairs within each group and the allele frequencies were compared. For 114 ΔF508 homozygous pairs, the 10% most extreme pairs correspond to 11 couples, defined by a rank number border of 11. However, DNA for less than 6 pairs was available if that cut-off border was used resulting in numbers too small to be evaluated statistically. Hence, for an initial comparison of genotype data the rank number 15 – enabling typing of 6 or more pairs for all markers studied – was employed. Upon detection of an allelic imbalance for any of the three groups, the complete cohort typed was evaluated involving the following sequences (pairs defined as extremes are underlined)[18]:

14 discordant pairs
— rank number DIS 3, 5, 6, 8, 9, 14, 15, 16, 18, 19, 21, 24
— 2 pairs from combined categories DC(1), DC(2)

16 Shown in chapter 2.2.1.4 on page 52.
17 Shown in chapter 2.2.1.3 on page 48.
18 dizygosity status of DC(4) and ND(2) not confirmed by July 1998, these pairs were excluded from evaluation

13 concordant/mild disease pairs
— rank numbers CON+ 2, 5, 6, 7, 10, 15, 16, 19
— 5 pairs from combined categories ND(1), ND(3), ND(4), ND(6), DC(3)

9 concordant/severe disease pairs
— rank numbers CON- 2, 5, 7, 8, 10, 12, 14, 25
— 1 pair from combined category ND(5)

Analysis of the data for complete cohorts was carried out as follows: starting with the 6 most extreme ranking pairs, allele frequencies were determined. Subsequent inclusion of the next ranking pair with genotype data lead to groups of cumulatively pooled couples with a size of 7, 8, 9 pairs etc. up to 14 pairs (all discordant pairs if all patients successfully typed), up to 13 pairs (all concordant / mild disease pairs if all patients successfully typed) or up to 9 pairs (all concordant / severe disease pairs if all patients successfully typed).

When data on expectancy values for alleles were available, the observed allele frequencies were compared to the values expected for the cohort of the respective size [frequencies for ΔF508 chromosomes were published for MetH and J3.11 (Tümmler *et al.* 1990), XV-2c, KM19 and TUB20 (Dörk *et al.* 1992)]. For chromosome 7 markers PON2 (Adkins *et al.* 1993), D7S525, D7S514 and D7S495 (Dib *et al.* 1996), no allele frequencies were available for ΔF508 chromosomes. For these markers all equally sized cohorts were compared to detect an allelic imbalance between pairs ranked DIS, CON+ or CON-. A similar approach was necessary for the markers D12S889 (Baens *et al.* 1995) and βENaCGT (Shimkets *et al.* 1994): these markers were selected due to their position relative to the candidate genes but are not integrated into common marker sets and, consequently, no expected allele frequencies are known.

Results on polymorphic markers are reported in the following format:

locus name - size in arbitrary repeat units, e.g. D7S514 - 13

3.2.3.4 Results on ion channel ENaC

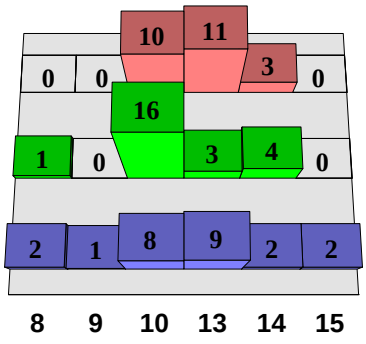
The amiloride sensitive sodium channel ENaC is related to CF disease in two ways: the channel was reported to be regulated by CFTR (Ismailov *et al.* 1996, Stutts *et al.* 1997) and mutations within the genes encoding for ENaC subunits have been demonstrated to cause pseudohypoaldosteronism (Chang *et al.* 1996, Strautnieks *et al.* 1996a), a disorder described as mimicking the clinical appearance of mild CF lung disease (Hanakoglu *et al.* 1994, Marthinsen *et al.* 1998). The ENaC is composed of three subunits. SCNN1A, coding for the α subunit, was mapped to chromosome 12p13. A complex repeat polymorphism, D12S889, has been mapped within a 40kb distance to SCNN1A on a cosmid clone (Baens *et al.* 1995). The satellite is composed of a $(CA)_n(GT)_m$ core repeat. Five alleles were identified among 10 unrelated Caucasians (Baens *et al.* 1995). The genes SCNN1B and SCNN1G, encoding the β and γ subunits of the ENaC, are both located within a 400kb fragment on chromosome 16p12-p13 (Voilley *et al.* 1995). Within SCNN1B, the dinucleotide repeat polymorphism β-ENaCGT was described (Shimkets *et al.* 1994). Five different alleles have been typed at β-ENaCGT in 50 unrelated Caucasians.

FIGURE 30: ALLELES NEAR SCNN1A, SCNN1B AND SCNN1G

D12S889 - near SCNN1A

Left: *allele distributions for 33 typed pairs (132 chromosomes)*
Right: *alleles displayed by the 6 most extreme pairs ranked*
DIS CON+ CON-

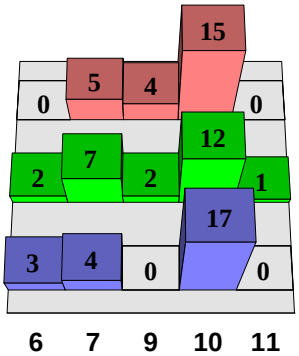
ARU	8	9	10	12	13	14	15
DIS (13)	—	—	19	2	23	8	—
CON+ (11)	2	—	27	—	10	5	—
CON- (9)	2	1	10	3	14	4	2
Σ	*4*	*1*	*56*	*5*	*47*	*17*	*2*



βENaCGT - near SCNN1B and SCNN1G

Left: *allele distributions for 35 typed pairs (140 chromosomes)*
Right: *alleles displayed by the 6 most extreme pairs ranked*
DIS CON+ CON-

ARU	6	7	9	10	11
DIS (13)	2	13	4	33	—
CON+ (13)	2	17	3	29	1
CON- (9)	3	5	—	28	—
Σ	*7*	*35*	*7*	*90*	*1*



3.2.3.4.1 Allelic association at D12S889 and βENaCGT

For all CF pairs typed, 7 different alleles were detected at D12S889 and 5 different alleles were detected at β-ENaCGT. The most frequent alleles were D12S889-10, on 38% of chromosomes typed, and β-ENaCGT-10, on 63% of chromosomes typed. At D12S889, for 24 out of 44 pairs a different genotype was observed between sib A and sib B. For β-ENaCGT, only 16 out of 46 pairs displayed non-identical genotypes for both sibs. For both, D12S889 and β-ENaCGT, no noticeable differences were observed comparing the number of pairs sharing one or more alleles identical by state within the groups of extreme pairs DIS, CON+ and CON-.

At D12S889, an association of the allele D12S889-10 with the phenotype *"concordant / mild disease"* was observed (figure 30). On all ΔF508 chromosomes, the two alleles D12S889-10 and D12S889-13 were presented on equal proportions of chromosomes (40% D12S889-10 and 38% D12S889-13, respectively). Within the extreme DIS and CON- pairs, both alleles were distributed as described for the complete cohort, but CON+ pairs ranking 15 or lower displayed D12S889-10 on 16 out of 24 chromosomes and D12S889-13 on 3 out of 24 chromosomes. The allele distribution at D12S889 for the 6 pairs with lowest rank numbers differed most noticeable between groups DIS and CON+.

The marker β-ENaCGT showed no evidence of an allelic imbalance as judged by comparing the groups CON+, CON- or DIS (figure 30): within each group of extreme pairs ranking 15 or lower, the allele β-ENaCGT-10 was found on equivalent proportions of typed chromosomes (15 out of 24 extreme DIS, 12 out of 24 extreme CON+ and 21 out of 28 extreme CON- pairs, respectively).

3.2.3.5 Results on the linkage group on chromosome 7

The high frequency of carriers for CF causing alleles – 1 in 22 healthy persons, as deduced from an incidence of 1 affected person in 2000 newborns (Romeo 1989) – has been the subject of speculation as to the underlying mechanisms leading to the prevalence of the CF causing alleles. The maintenance of a high frequency of mutant *CFTR* alleles was attributed to genetic drift or a reproductive advantage of the heterozygous state, the latter being either transmitted through the *CFTR* gene itself or by genes within linkage disequilibrium to *CFTR*. Evidence for so-called "hitch-hiking genes", detectable through markers localised in the *CFTR* gene region, has been reported by three authors: Santis *et al.* (1990b) has reported an association of J3.11-2 alleles with milder pulmonary disease among ΔF508 homozygotes. Tümmler *et al.* (1990) has described an association of MetH and J3.11 alleles with growth parameters in CF patients. Macek *et al.* (1997) has reported an increased proportion of CS.7/*HhaI*-2 alleles in elderly female individuals within the Czech population. Within this chapter, allele frequencies at loci PON2[19], MetH/*MspI*, XV-2c/*TaqI*, KM.19/*PstI*, TUB20/*PvuII*, J3.11/*MspI*, D7S514 and D7S495 – localised in a 40 cM interval surrounding the *CFTR* gene as showed in figure 31 – are reported for CF twin and sib pairs with selected phenotypes.

19 Sequence variant Arg191Gln in PON2 as shown in figure 5 in chapter 2.1.2.2.4 on page39.

FIGURE 31: INTEGRATED MAP OF 40 cM REGION AROUND *CFTR*

The relative positions of the map elements are consistent with mapping data generated by Tsui et al. 1985, Rommens et al. 1989, Green et al. 1995, Höglund et al. 1995, Bouffard et al. 1997, Fisher et al. 1998, Li et al. 1998. Positions in cM are given as published by Bouffard et al. 1997. Colours refer to positions on the genetic map or the physical

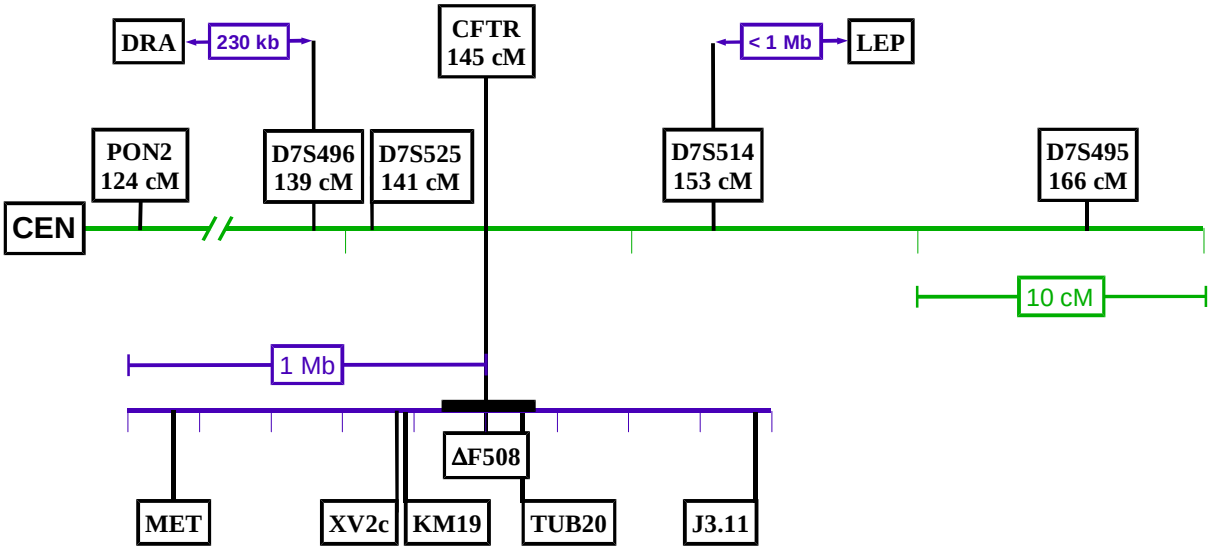


TABLE 18: ALLELE FREQUENCIES FOR D7S525, D7S514 AND D7S495

Expected frequencies are derived from the GDB for 56 Caucasian chromosomes (CEPH).

D7S525 ***(34 ΔF508 homozygous pairs and 8 non-ΔF508 homozygous dizygous twin pairs typed)***

Size [bp]	219	221	223	227	229	231	233	235
Frequency [%] non-CF chromosomes (CEPH)	**2**	**2**	**7**	**10**	**30**	**43**	**4**	**2**
Size [ARU]	4	5	6	8	9	10	11	12
Frequency [%] ΔF508 homozygous pairs	**—**	**1**	**7**	**4**	**26**	**57**	**4**	**—**
Number of alleles (8 Non-ΔF508 homozygous pairs)	—	—	3	—	20	9	—	—

D7S514 ***(34 ΔF508 homozygous pairs and 6 non-ΔF508 homozygous dizygous twin pairs typed)***

Size [bp]	147	149	151	153	155	157	159
Frequency [%] non-CF chromosomes (CEPH)	**3**	**3**	**17**	**41**	**28**	**5**	**—**
Size [ARU]	8	9	10	11	12	13	14
Frequency [%] ΔF508 homozygous pairs	**—**	**15**	**6**	**42**	**35**	**1**	**1**
Number of alleles (6 Non-ΔF508 homozygous pairs)	—	3	4	13	4	—	—

D7S495 ***(36 ΔF508 homozygous pairs and 8 non-ΔF508 homozygous dizygous twin pairs typed)***

Size [bp]	150	152	154	158	160	162	164	166	168
Frequency [%] non-CF chromosomes (CEPH)	**7**	**—**	**14**	**18**	**—**	**20**	**27**	**5**	**5**
Size [ARU]	8	9	10	12	13	14	15	16	17
Frequency [%] ΔF508 homozygous pairs	**5**	**1**	**19**	**23**	**1**	**25**	**21**	**4**	**1**
Number of alleles (8 Non-ΔF508 homozygous pairs)	1	2	5	10	—	10	1	—	3

3.2.3.5.1 Allele distributions at D7S525, D7S514 and D7S495

Allele frequencies for D7S525, D7S514 and D7S495 were published for non-CF families (references see chapter). The allele distributions for ΔF508 chromosomes and non-CF chromosomes matched with regard to the most frequent alleles. Additionally, similar "gaps" in the sequence of D7S495 alleles arranged by size – alleles D7S495-9, D7S495-11 and D7S495-13, respectively – were discernible for both the CEPH panel and the typed ΔF508 chromosomes. However, the most frequent and the second most frequent alleles were observed in different proportions for all polymorphic markers, most pronounced at D7S525 and least pronounced at D7S495. Furthermore, the allele spectrum at D7S525 and D7S514 appeared to be smaller for typed CF chromosomes than for typed non-CF chromosomes: the shortest and the largest D7S525 allele observed on non-CF chromosomes was not observed among ΔF508 homozygous patients and the shortest D7S514 allele on non-CF chromosomes did not occur on typed CF chromosomes. A small set of non-ΔF508 CF chromosomes was analysed for dizygous twin pairs. The differences of the allele distributions between ΔF508 homozygous and non-ΔF508 homozygous pairs were again most striking at D7S525 and nearly unnoticeable at D7S495.

3.2.3.5.2 Allelic association at loci between PON2 and D7S495

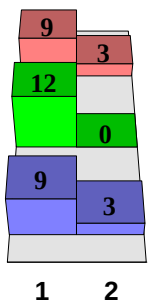
3.2.3.5.2.1 MetH and J3.11: ± 1 Mb near *CFTR*

No difference between the allele distributions among CON+, CON- or DIS pairs have been observed at the loci XV-2c, KM.19 and TUB20. Allele frequencies for the markers XV-2c, KM.19 and TUB20 observed in the groups of CF twins and sibs did not differ from the expected frequencies (Tümmler *et al.* 1990, Dörk *et al.* 1992) for ΔF508 chromosomes. Genotypes at XV-2c and TUB20 were identical for both sibs within a pair for all patient couples investigated. Non identical genotypes within a pair were observed for one pair at KM.19.

FIGURE 32: ALLELE DISTRIBUTIONS AT MetH AND J3.11
Left: *allele distributions for 32 typed pairs (64 alleles) at MetH and 34 typed pairs (68 alleles) at J3.11*
Right:: *alleles displayed by the 6 most extreme ranking pairs* DIS CON+ CON-

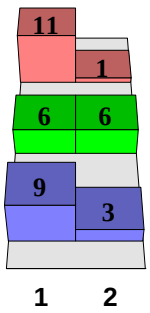
MetH

Allele	1	2
DIS (12)	18	6
CON+ (11)	20	2
CON- (9)	13	5
Σ	*51*	*13*



J3.11

Allele	1	2
DIS (13)	17	9
CON+ (12)	14	10
CON- (9)	14	4
Σ	*45*	*23*



At MetH, an association of MetH-1 with the phenotype *"concordant / mild disease"* was detected (table 19 and figure 32). The deviation of observed allele frequency to the frequency expected for ΔF508 chromosomes was significant for the extreme CON+ pairs ranked 15 or lower (χ^2 = 3,81 with p = 0.05 at χ^2 = 3.84). The allele frequencies for CON- pairs and DIS pairs were as expected for ΔF508 chromosomes. In one pair, sib A and sib B carried different genotypes at MetH.

At J3.11, an association of allele J3.11-1 with the phenotype *"discordant"* was noticed (table 20 and figure 32). Among DIS pairs the deviation from the allele frequency expected for ΔF508 chromosomes was significant for the 6 pairs ranked 14 and lower (χ^2 = 3.81 with p = 0.05 at χ^2 = 3.84). The allele frequencies for the CON+ pairs and the CON- pairs at J3.11 were as expected for ΔF508 chromosomes. For all patient pairs studied, the same genotype was observed for sib A and sib B at J3.11.

TABLE 19: GENOTYPES AT METH FOR CONCORDANT / MILD DISEASE PAIRS

pair ND(1) showed a recombination event: genotype sib A 1-2, sib B 2-2
expected allele frequencies for ΔF508 chromosomes: MetH allele 1 0.7, allele 2 0.3

Pair:	CON+2	CON+5	CON+6	CON+7	CON+10	CON+15	CON+16	CON+19	ND(1)	ND(3)	ND(6)
MetH	1—1	1—1	1—1	1—1	1—1	1—1	1—1	1—2	rec	1—1	1—1

OBSERVED IN	CON+ 2-15	CON+ 2-16	CON+ 2-19	ALL CON+
ALLELE 1	12	14	15	19.5
ALLELE 2	0	0	1	2.5
EXPECTED FOR	n=12	n=14	n=16	n=24
ALLELE 1	8.4	9.8	11.2	15.4
ALLELE 2	3.6	4.2	4.8	6.6
χ^2	3.8	4.7	3.2	4.4

TABLE 20: GENOTYPES AT J3.11 FOR DISCORDANT PAIRS

expected allele frequencies for ΔF508 chromosomes: J3.11 allele 1 0.6, allele 2 0.4

Pair:	DIS3	DIS5	DIS6	DIS8	DIS9	DIS14	DIS16	DIS18	DIS19	DIS21	DIS24	DC(1)	DC(2)
J3.11	1—1	1—1	1—1	1—1	1—1	1—2	1—2	1—2	1—2	1—2	2—2	1—2	1—2

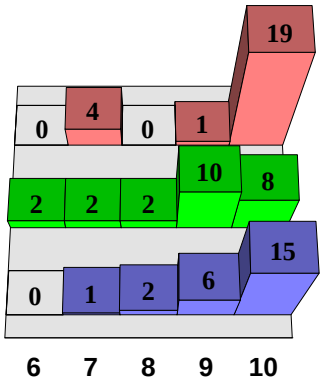
OBSERVED IN	DIS 3-14	DIS 3-16	DIS 3-18	ALL DIS
ALLELE 1	11	12	13	17
ALLELE 2	1	2	3	9
EXPECTED FOR	n=12	n=14	n=16	n=26
ALLELE 1	7.2	8.4	9.6	15.6
ALLELE 2	4.8	5.6	6.4	10.4
χ^2	3.8	2.9	2.2	0.1

FIGURE 33: ALLELE DISTRIBUTIONS AT D7S525 AND D7S514

D7S525

Left: *allele distributions for 32 typed pairs (128 chromosomes)*
Right: *alleles displayed by the 6 most extreme pairs ranked*
DIS CON+ CON-

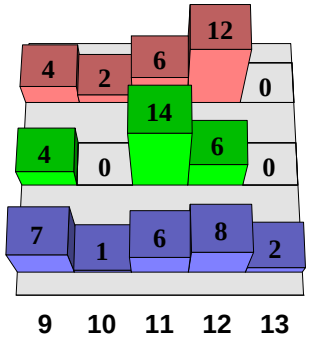
ARU	6	7	8	9	10	11
DIS (12)	—	6	2	6	30	4
CON+ (11)	2	2	2	18	19	1
CON- (9)	—	1	2	6	27	—
Σ	*2*	*9*	*6*	*30*	*76*	*5*



D7S514

Left: *allele distributions for 32 typed pairs (128 chromosomes)*
Right: *alleles displayed by the 6 most extreme pairs ranked*
DIS CON+ CON-

ARU	9	10	11	12	13
DIS (13)	8	3	18	23	—
CON+ (10)	4	—	22	14	—
CON- (9)	9	3	14	8	2
Σ	*21*	*6*	*54*	*45*	*2*



3.2.3.5.2.2 D7S525 and D7S514: 10 cM and closer to *CFTR*

Localised further towards the centromer than MetH, D7S525 was studied. 6 alleles were identified on all chromosomes typed, with allele D7S525-10 being most frequent on ΔF508 chromosomes (57% of alleles were D7S525-10 and 26% were D7S525-9) while allele D7S525-9 was most frequent for non-ΔF508 homozygous dizygous twins tested (D7S525-9 was found on 20 out of 36 chromosomes, and D7S525-10 was observed on 9 out of 36 chromosomes). For 9 out of 43 pairs, a non-identical genotype was observed between both sibs. Association of a D7S525 allele with the phenotype *"concordant / mild disease"* was noticed (figure 33): chromosomes from pairs ranking 15 or lower were typed D7S525-10 on 19 out of 24 alleles for DIS pairs, on 20 out of 28 alleles for CON- pairs, but only on 8 out of 24 alleles on CON+ pairs. The allele distribution at D7S525 was different comparing six pairs with lowest rank numbers between CON+ and DIS or between CON+ and CON-. Allele distributions between DIS and CON- pairs did not differ at D7S525.

Located further towards the telomer than J3.11, D7S514 was investigated. On all chromosomes typed, 6 alleles were identified. On ΔF508 chromosomes, alleles D7S514-11 and D7S514-12 were most frequent (42% and 35%, respectively). The proportion of both alleles was different for non-

ΔF508 homozygous dizygous twins typed: D7S514-11 was found on 13 out of 36 chromosomes and allele D7S514-12 was observed on 4 out of 36 chromosomes. In 3 out of 42 pairs, the genotype at D7S514 was different for sib A and sib B. Allelic imbalances at D7S514 were found comparing all three groups of pairs, DIS, CON+ and CON- (figure 33). The differences in allele distributions were most pronounced between DIS and CON+ pairs: for the six pairs with lowest rank numbers, 6 alleles showed D7S514-11 and 12 alleles showed D7S514-12 on DIS pairs, while on CON+ pairs 14 alleles carried D7S514-11 and 6 alleles carried D7S514-12. Interestingly, an association between heterozygosity at D7S514 and the phenotype *"concordant/severe disease"* was observed: for 5 out of 12 pairs ranked CON+ or DIS with rank number 15 or lower, both sibs were homozygous for a D7S514 genotype. In contrast, 8 out of 9 pairs ranked CON- showed a heterozygous D7S514 genotype for both sibs (table 21).

TABLE 21: HETEROZYGOSITY AT D7S514

NUMBER OF SIBLINGS WITHIN A PAIR WITH HETEROZYGOUS GENOTYPES AT D7S514

Extreme DIS, CON+, CON- pairs	0	1	2	*All DIS, CON+, CON- pairs*	0	1	2
DIS (6 pairs)	2	—	4	DIS (13 pairs)	4	1	8
CON+ (6 pairs)	3	—	3	CON+ (10 pairs)	5	—	5
CON- (6 pairs)	—	—	6	CON- (9 pairs)	1	—	8

3.2.3.5.2.3 PON2 and D7S495: further away than 10 cM from *CFTR*

Localised further towards the centromer than D7S525, PON2 was analysed. No difference was observed when comparing allele frequencies between pairs ranked DIS, CON+ or CON-. Comparison of PON2 allele frequencies expected for non CF chromosomes to those observed for CON+, CON- or DIS pairs showed a significant deviation in all three cases, indicating the expected linkage disequilibrium for the *CFTR* region (Eiberg *et al.* 1985). In 4 out of 36 typed pairs, a different genotype for sib A and sib B was observed at PON2.

Located further towards the telomer than D7S514, D7S495 was analysed. 10 different alleles were observed at D7S495. The allele spectrum differed between pairs ranked DIS, pairs ranked CON+ and pairs ranked CON- (figure 34): for pairs ranking 15 or lower, the most frequent allele was D7S495-15 for DIS pairs, D7S495-12 for CON+ pairs and D7S495-14 for CON- pairs (observed on 8 out of 24, 9 out of 24 and 10 out of 28 chromosomes, respectively). The difference between allele frequencies was pronounced when comparing DIS and CON+ pairs or CON+ and CON- pairs. Non-identical D7S495 genotypes for both sibs within a pair were observed for 15 out of 46 pairs. For 3 out of 11 pairs ranked DIS, no identical D7S495 allele was found when comparing sib A and sib B. This condition was not observed among 8 pairs ranked CON+ or 8 pairs ranked CON-. For 13 out of 22 pairs ranked concordant (CON+ or CON- or ND), both alleles of both siblings at D7S495 were identical by state (table 22). In three cases, a double recombination event resulted in a different D7S495 allele on both chromosomes for both sibs of a pair (0 identical alleles between sib A and sib B). These three pairs were ranked discordant and this condition was not observed on any concordant pair.

FIGURE 34: ALLELE DISTRIBUTIONS AT D7S495

D7S495

Left: *allele distributions for 34 typed pairs (136 chromosomes)*
Right: *alleles displayed by the 6 most extreme pairs ranked*
DIS CON+ CON-

ARU	8	9	10	12	13	14	15	16	17
DIS (13)	5	—	9	6	2	8	21	1	—
CON+ (12)	2	1	17	16	—	9	3	—	—
CON- (9)	—	—	—	11	—	16	3	4	2
Σ	*7*	*1*	*26*	*33*	*2*	*33*	*27*	*5*	*2*

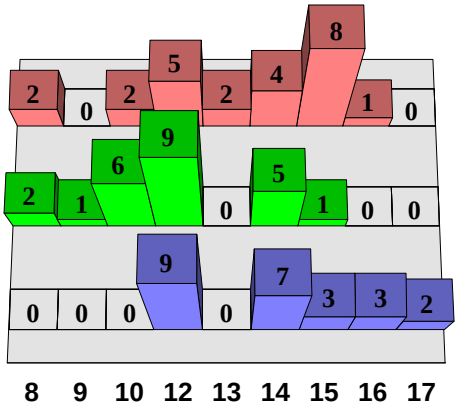


TABLE 22: ALLELES IDENTICAL BY STATE AT D7S495

NUMBER OF ALLELES IDENTICAL BY STATE WITHIN A PAIR

Extreme DIS, CON+, CON- pairs		0	1	2	*All DIS, CON+, CON- pairs*		0	1	2
DIS	(6 pairs)	3	—	3	DIS	(13 pairs)	3	4	6
CON+	(6 pairs)	—	2	4	CON+	(12 pairs)	—	4	8
CON-	(6 pairs)	—	4	2	CON-	(9 pairs)	—	5	4

3.2.3.5.2.4 Recombination events at loci between PON2 and D7S495

The probability of observing a recombination event increases with the physical distance between the loci studied as well as with the number of alleles displayed by the locus investigated. Consequently, more non-identical genotypes were observed at polymorphic markers D7S525, D7S514 and D7S495 compared to PON2. As shown in table 23, the most recombination events were observed at D7S495: a third of all pairs did not display two alleles identical by state.

When comparing the genetic map of the region encompassing the *CFTR* gene and the observed frequency of recombination events, a crosswise relationship between the genetic distances of polymorphic markers to *CFTR* and the number of recombination events among the typed pairs were observed. Regarding the positions on the genetic map of D7S525 and D7S514, displaying a similar allele spectrum, D7S514 is located at twice the distance to *CFTR* compared with D7S525. However, a fifth of the pairs typed had non-identical genotypes at D7S525 but less than a tenth of the pairs typed at D7S514 provided evidence for recombination events between the *CFTR* gene and D7S514 (table 23).

TABLE 23: RECOMBINATION AT LOCI BETWEEN PON2 AND D7S495

(a) number of pairs displaying less than 2 alleles identical by state when genotypes of sib A and sib B were compared.

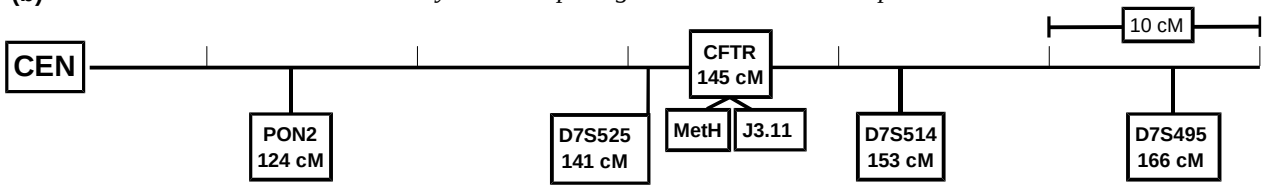
(b) Graphical representation of non-proportional relationship, observed at D7S525 and D7S514, between the distance on genetic map and number of recombination events observed among selected twins and sibs with extreme phenotypes.

PON2 is a diallelic marker, thus less recombination events are observable at PON2 compared to polymorphic marker loci. In contrast, D7S525 and D7S514 are comparable with respect to the allele distribution (see figure 17 on page 63). However, only a minority of patient pairs showed recombination events at D7S514 while a fifth of the typed pairs displayed less than two alleles identical by state at D7S525. This is unanticipated when taking the genetic map into account: D7S525 is located closer to the CFTR gene than D7S514 (Bouffard et al 1997) indicating that on chromosomes of the general population, a lower recombination frequency has been observed between CFTR and D7S525 compared to CFTR and D7S514.

(a)	PON2	D7S525	MET	J3.11	D7S514	D7S495
PAIRS TYPED	36	43	38	31	42	46
NO OF PAIRS ≠ 2 IBS	4	9	1	0	3	15

≠2 IBS: less than two alleles identical by state comparing sib A to sib B within a pair

(b)



	PON2 - CFTR	D7S525 - CFTR	CFTR - D7S514	CFTR - D7S495
DISTANCE ON GENETIC MAP	21 cM	4 cM	8 cM	21 cM
% of pairs ≠ 2 alleles IBS OBSERVED AMONG SELECTED PATIENT PAIRS	12 %	21 %	7 %	33 %

3.2.3.6 Results on D2S1788

D2S1788 was evaluated due to the results gained on markers J3.11 and D7S514: located further towards the telomer than *CFTR*, obesity related traits have been mapped linking D7S514 to extremity skinfolds and D7S495 to waist circumference (Duggirala *et al.* 1996). The human homologue of the murine obesity gene, the leptin gene LEP, maps within 1 Mb of D7S514 (Green *et al.* 1995) and is a candidate gene for the modulation of CF disease via the nutritional status. The results gained on J3.11 and D7S514 show an allelic imbalance in pairs ranked discordant which points to a *trans* conducted influence on a gene located further towards the telomer than *CFTR*: CF twins and sibs have inherited the same *CFTR* gene region up to D7S514, demonstrated by the fact that two D7S514 alleles are identical by state for all but 3 out of 42 pairs. Hence, any influence at a locus close to D7S514 evoking discordance within a CF sib pair cannot be caused by the gene transmitted with D7S514 itself, but by a gene for which both discordant siblings carry different alleles. A candidate region influencing serum leptin levels has been identified near D2S1788 (Comuzzie *et al.* 1997), close to the *POMC* gene encoding for proopiomelanocorticoin, which gives rise to hormones inducing

glucocorticoid secretion. Glucocorticoids have been shown to regulate the expression of leptin (De Vos *et al.* 1995). An influence on the expression levels of a gene qualifies as a *trans* influence and accordingly, D2S1788 was investigated in the panel of CF twins and sibs.

3.2.3.6.1 Allelic association at D2S1788

At D2S1788, a total of 13 alleles have been observed on 34 typed ΔF508 homozygous pairs. The most frequent alleles were D2S1788-12, D2S1788-19, D2S1788-20 and D2S1788-22 on 11%, 13%, 14% and 21% of chromosomes, respectively. Comparable proportions of these alleles were observed for subgroups of CON+ and CON- pairs, but among DIS pairs, other alleles were dominant (figure 35): D2S1788-18 was typed on 8 alleles out of 14 pairs ranked DIS while D2S1788-18 was found on no chromosome from 12 CON+ or 8 CON- pairs. D2S1788-21 was identified on 13 chromosomes from 14 pairs ranked DIS but only 3 chromosomes from 12 pairs ranked CON+ and no chromosome from 8 pairs ranked CON- were typed D2S1788-21. For a total of 10 out of 14 pairs categorised as DIS, at least one sib carried one of the alleles D2S1788-18 or D2S1788-21 on either one or both chromosomes. Based on the observation of the association of D2S1788-18 and D2S1788-21 alleles with the phenotypic trait "discordance", pairs sharing one or both of these alleles identical by state were excluded from the analysis of allele sharing among concordant and discordant pairs[20] (table 24, row DIS*: all pairs with one or both alleles D2S1788-18 or D2S1788-21 identical by state excluded). 7 pairs ranked discordant and 20 pairs ranked concordant remained. A fifth of the concordant pairs had no identical allele comparing genotypes of sib A and sib B. This condition was observed in 4 out of 7 DIS* pairs. In a third of the pairs ranked concordant, both alleles were identical by state. This was observed for one out of 7 DIS* pairs.

TABLE 24: ALLELES IDENTICAL BY STATE AT D2S1788

NUMBER OF ALLELES IDENTICAL BY STATE WITHIN A PAIR

Extreme DIS, CON+, CON- pairs		0	1	2	*All DIS, CON+, CON- pairs*		0	1	2
DIS – all genotypes	(6 pairs)	3	1	2	DIS	(14 pairs)	4	4	6
DIS*	(4 pairs)	3	1	—	DIS*	(7 pairs)	4	2	1
CON+	(6 pairs)	2	2	2	CON+	(12 pairs)	3	6	3
CON-	(6 pairs)	1	2	3	CON-	(8 pairs)	1	4	3
CON+ and CON-	(12 pairs)	3	4	5	CON+ and CON-	(20 pairs)	4	10	6

**7 out of 14 pairs ranked DIS with one or two alleles D2S1788-18 or D2S1788-21 IBS excluded; this condition was not observed for pairs ranked CON+ or CON-*

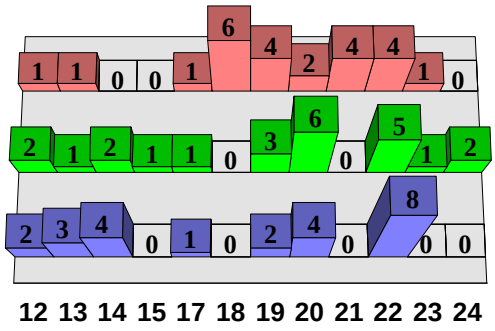
20 The hypotheses tested by (a) comparison of the number of alleles identical by state between pairs ranked concordant and pairs ranked discordant or (b) analysing allelic association with the phenotypes concordant and discordant are incompatible when an allelic association with the phenotype "discordance" is detected: when analysing a gene modulating the phenotype, discordance is expected when both sibs of a pair carry non-identical genotypes at the gene locus. If a marker locus analysed is linked to the gene, non-identical genotypes at the marker locus indicate non-identical genotypes at the gene locus. Overrepresentation of one allele in the subgroup of discordant pairs points to the hypothesis that an allele at the gene locus is itself causing the phenotype "discordance", even if —or, caused by the fact that — both sibs are carrying that allele identical by descent.

FIGURE 35: ALLELE DISTRIBUTIONS AT D2S1788

D2S1788

Left: *allele distributions for 34 typed pairs (136 chromosomes)*
Right: *alleles displayed by the 6 most extreme pairs ranked*
DIS CON+ CON-

ARU	11	12	13	14	15	17	18	19	20	21	22	23	24
DIS (14)	1	4	1	—	—	1	8	9	4	13	12	2	1
CON+ (12)	—	6	3	2	1	1	—	8	8	3	10	4	2
CON- (8)	—	4	3	5	—	1	—	2	7	—	9	1	—
Σ	*1*	*14*	*7*	*7*	*1*	*3*	*8*	*19*	*19*	*16*	*31*	*7*	*3*



3.2.3.7 Correlation of genotype and clinical phenotype

In chapters 3.2.3.4, 3.2.3.5 and 3.2.3.6, the overrepresentation of one allele was described at loci D12S889, D7S525, MetH, D7S514, D7S495 for pairs ranked "*concordant / mild disease*" (CON+), at locus D7S495 for pairs ranked "*concordant / severe disease*" (CON-) and at loci J3.11, D7S514, D7S495 and D2S1788 for pairs ranked "*discordant*" (DIS). The effects are summarised in table 25. This chapter aims at correlating a pair's genotype and the phenotypic trait suspected to be influenced by the allele studied.

TABLE 25: SUMMARY OF OBSERVED ALLELIC ASSOCIATIONS

LOCUS	ALLELE	OVERREPRESENTED FOR PAIRS RANKED
D7S525	D7S525-9	CON+
MetH	MetH-1	CON+
J3.11	J3.11-1	DIS
D7S514	D7S514-11	CON+
D7S514	D7S514-12	DIS
D7S495	D7S495-10	CON+
D7S495	D7S495-14	CON-
D7S495	D7S495-15	DIS
D12S889	D12S889-10	CON+
D2S1788	D2S1788-18	DIS

For all alleles overrepresented in one of the three patient groups DIS, CON+ and CON-, a direct relationship of intrapair-discordance (allelic imbalance observed for DIS pairs) or disease severity (allelic imbalance observed for CON+ or CON- pairs) was conjectured. The maximal effect was presumed for the condition when all four chromosomes of both patients of a pair were carrying the allele that was suspected as modulating the phenotypic appearance. The similarity of a pair's genotype to this genotype with inferred maximal influence was estimated by counting the number of chromosomes of a pair carrying the allele that was overrepresented in one patient pair group DIS, CON+ or CON-. In detail, the following genotypes of a pair are possible when assessing the influence of an allele **a** at a locus displaying more than two alleles:

GENOTYPE Sib A	GENOTYPE Sib B	NUMBER OF ALLELES **a** OBSERVED IN PAIR	NUMBER OF ALLELES IDENTICAL BY STATE
a — **a**	**a** — **a**	4	2
a — **a**	**a** — other	3	1
a — **a**	other — other	2	0
a — other	**a** — other	2	1 or 2
a — other	other — other	1	0, 1 or 2
other — other	other — other	0	0, 1, or 2

The condition "two alleles **a** observed within pair" does not discriminate between pairs where one sib carries both alleles **a** and pairs where each sib carries one allele **a**.

The allelic imbalances were detected within groups of patient pairs which were categorised employing parameters that combine nutritional and pulmonary status. In order to analyse the influence of an allele on the nutritional and the pulmonary status, the genotypes were correlated to parameters describing either disease severity or the discordance based on wfh%, FEVPerc and the combined parameters DfO and DELTA[21]. To describe the disease severity, the sum of rank numbers of both sibs within a pair (ΣRANK) was calculated for wfh%, FEVPerc and DfO respectively. The intrapair discordance was assessed using the intrapair rank number difference (ΔRANK) of wfh% and FEVPerc as well as the rank number for DELTA. Statistical analysis was carried out with a Mann-Whitney rank test.

3.2.3.7.1 Near SCNN1A: D12S889

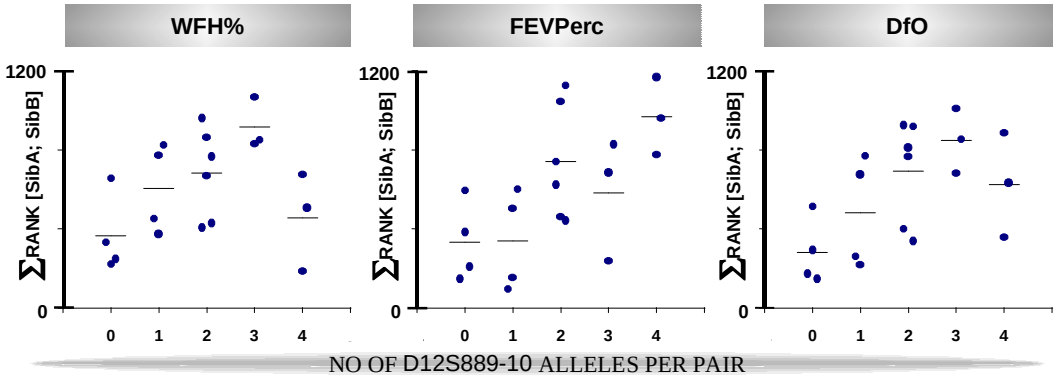
Figure 36 displays the similarity of a pair's genotype to the condition of homozygosity for the allele D12S889-10, found associated with the phenotype concordant / mild disease (chapter 3.2.3.4). An increasing wfh%, FEVPerc and DfO was seen with an increasing number of D12S889-10 alleles. The effect was significant comparing pairs where each sib carries at least one D12S889-10 allele to pairs where one or no sib carries a D12S889-10 allele ($p < 0.01$ for FEVPerc and $p < 0.05$ for wfh%).

21 Defined in figure 10 on page 48.

FIGURE 36: MODULATION OF DISEASE SEVERITY AT D12S889

ΣRANKwfh%, ΣRANKFEVPerc *and* ΣRANKDfO *vs number of* D12S889-10 *alleles per pair*

No of D12S889-10 alleles per pair	0	1	2	3	4
No of pairs ranked CON+ or CON-	4	4	6	3	3



3.2.3.7.2 40 cM region encompassing the CFTR gene

At two loci located further towards the centromer than *CFTR* and at three loci located further towards the telomer than *CFTR* an allelic imbalance was observed for pairs ranked concordant. Figure 37 and figure 38 display the relationship between alleles D7S525-9, MetH-1, D7S514-11, D7S495-10 and D7S495-14 and the disease severity for all pairs ranked CON+ or CON-. The influence of alleles J3.11-1, D7S514-12 and D7S495-15 on the intrapair discordance is shown in figure 39. The effects are summarised in chapter 3.2.3.8.2.

3.2.3.7.2.1 Modulation of disease severity among concordant pairs

The association observed for MetH-1 and disease mildness was mediated by the nutritional status as demonstrated by increasing ΣRANKwfh% comparing 1 or 2 alleles MetH-1 for a patients genotype ($p < 0.01$). ΣRANKFEVPerc and MetH-1 showed no correlation. The combined parameter ΣRANKDfO displayed a less pronounced relationship to the number of MetH-1 alleles than ΣRANKwfh% (figure 37b). At D7S525, located further towards the centromer than MetH, a similar trend was observed: with an increasing number of D7S525-9 alleles, ΣRANKwfh% but not ΣRANKFEVPerc was higher. Pairs with no D7S525-9 alleles displayed the broadest spectrum of ΣRANKwfh%, ΣRANKFEVPerc and ΣRANKDfO (figure 37a).

Further towards the telomer than *CFTR*, the relationship of the genotype to the clinical severity was observed at D7S495: while pairs carrying D7S495-10 alleles had higher ΣRANKwfh% and ΣRANKDfO than pairs carrying no D7S495-10 allele ($p < 0.01$), pairs with D7S495-14 alleles had a lower ΣRANKwfh% and ΣRANKDfO than pairs carrying no D7S495-14 allele ($p < 0.01$). The number

of D7S495-10 or D7S495-14 alleles did not correlate with ΣRANKFEVPerc. Again, the broadest spectrum of ΣRANKwfh% and ΣRANKDfO was observed for pairs without the alleles detected in allelic imbalance at locus D7S495 (figure 38b, c). At D7S514, the influence of D7S514-11 alleles on disease severity was week: the 5 pairs displaying four D7S514-11 alleles had a higher ΣRANKwfh% than most pairs with other genotypes. The mean ΣRANKFEVPerc and ΣRANKDfO increased with the number of D7S514-11 alleles (figure 38a). At D7S514 and D7S495, both alleles found associated with CON+ pairs occurred nearly exclusively on either 0 or 2 or 4, but not on 1 or 3 chromosomes of pairs ranked concordant. In four cases, D7S495-14 – observed most frequently among CON- pairs – was found on one out of four chromosomes within a pair. All cases with 2 alleles D7S514-11 or D7S495-10 or D7S495-14 were composed of genotypes wherein each sib carried one of the alleles.

FIGURE 37:MODULATION OF DISEASE SEVERITY (I): CENTROMER

(a) ΣRANKwfh%, ΣRANKFEVPerc *and* ΣRANKDfO *vs number of* D7S525-9 *alleles per pair*

No of D7S525-9 alleles per pair	0	1	2	3	4
No of pairs ranked CON+ or CON-	11	2	2	3	2

(b) ΣRANKwfh%, ΣRANKFEVPerc *and* ΣRANKDfO *vs number of* MetH-1 *alleles per sib*

No of MetH-1 alleles per pair	0	1	2
No of pairs ranked CON+ or CON-	0	7	13

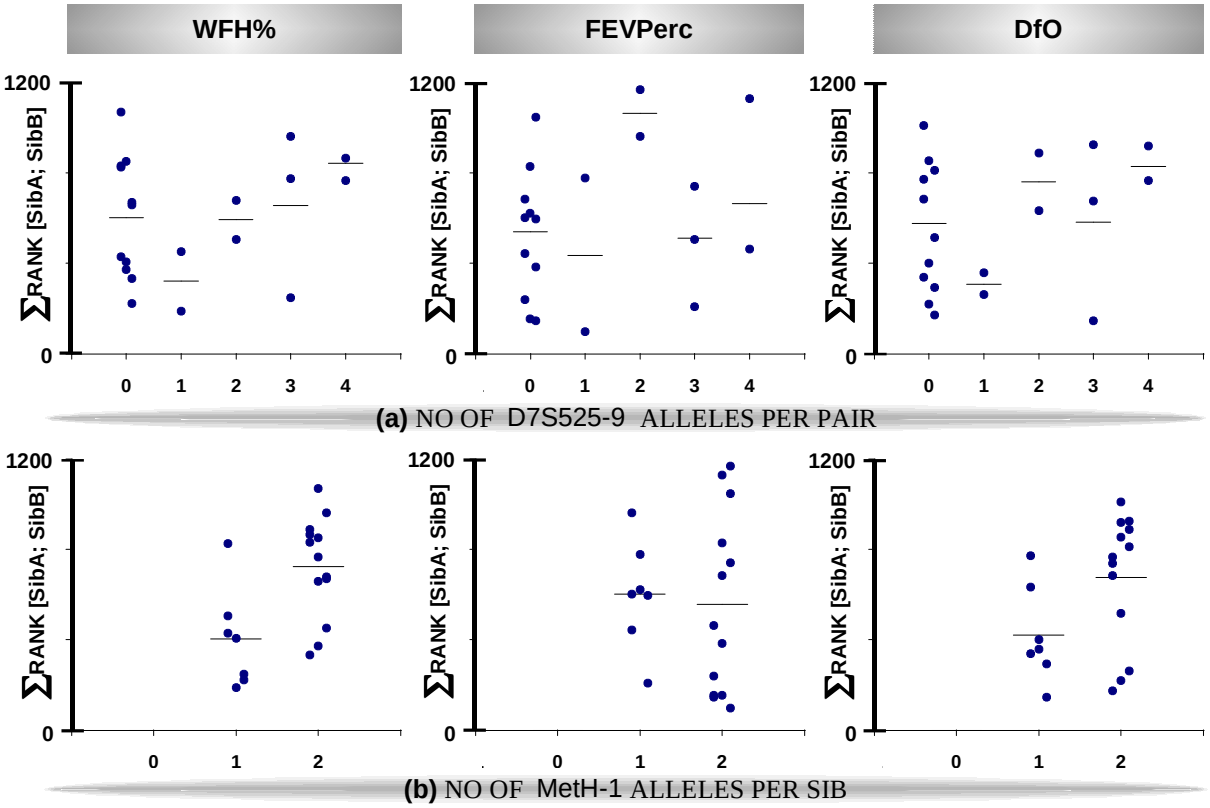


FIGURE 38: MODULATION OF DISEASE SEVERITY (II): TELOMER

(a) ΣRANKwfh%, ΣRANKFEVPerc *and* ΣRANKDfO *vs number of* D7S514-11 *alleles per pair*

No of D7S514-11 alleles per pair	0	1	2	3	4
No of pairs ranked CON+ or CON-	6	—	8	—	5

(b) ΣRANKwfh%, ΣRANKFEVPerc *and* ΣRANKDfO *vs number of* D7S495-10 *alleles per pair*

No of D7S495-10 alleles per pair	0	1	2	3	4
No of pairs ranked CON+ or CON-	14	1	4	—	2

(c) ΣRANKwfh%, ΣRANKFEVPerc *and* ΣRANKDfO *vs number of* D7S495-14 *alleles per pair*

No of D7S495-14 alleles per pair	0	1	2	3	4
No of pairs ranked CON+ or CON-	8	4	7	1	1

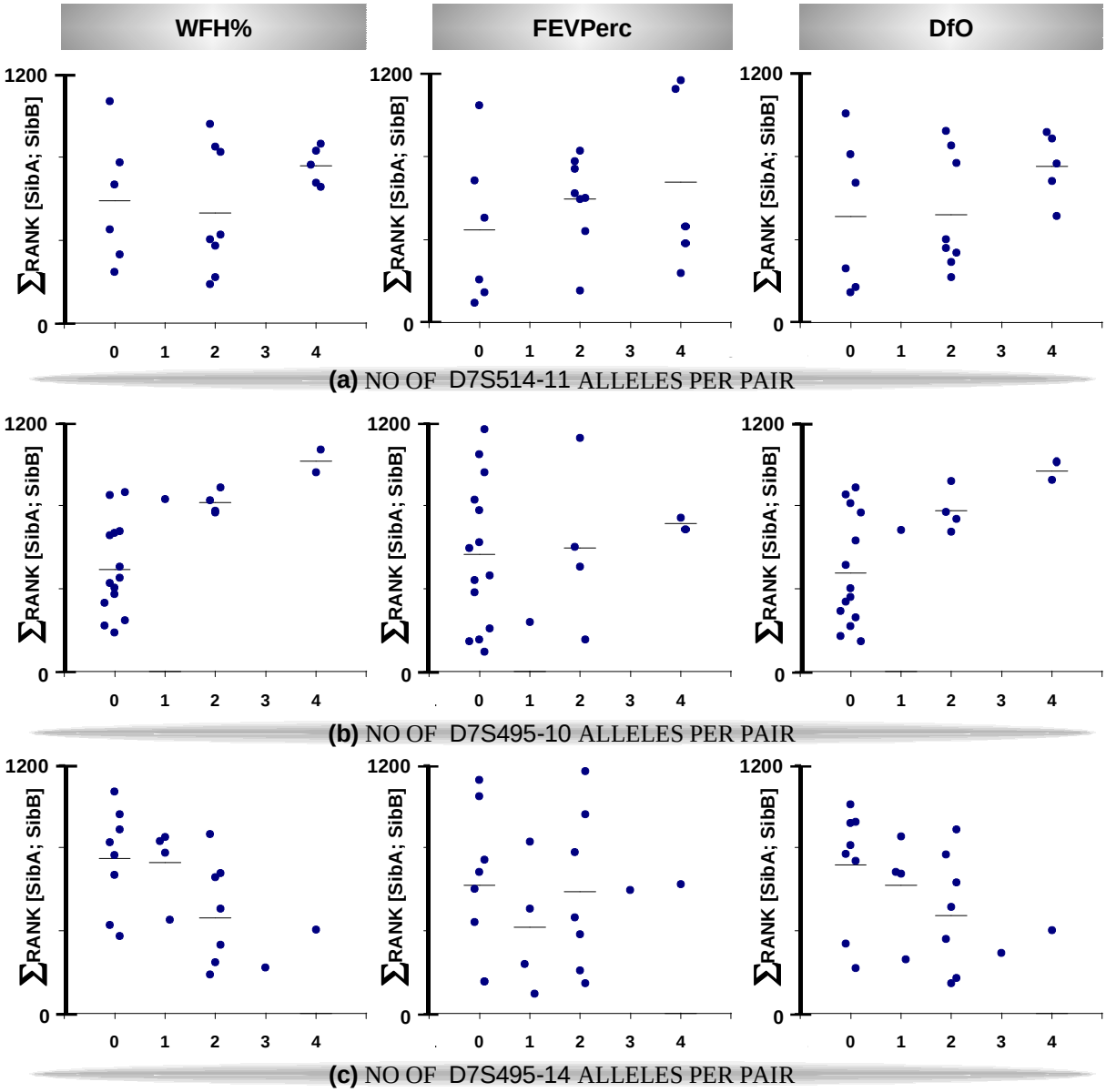


FIGURE 39: ALLELES OBSERVED PREDOMINANTLY FOR PAIRS RANKED DIS

(a) ΔRANKwfh%, ΔRANKFEVPerc *and* DELTA *vs number of* J3.11-1 *alleles per sib*

No of J3.11-1 alleles per sib	0	1	2
No of pairs ranked DIS	2	7	5

(b) DELTA *vs number of* D7S514-12 *alleles per pair*

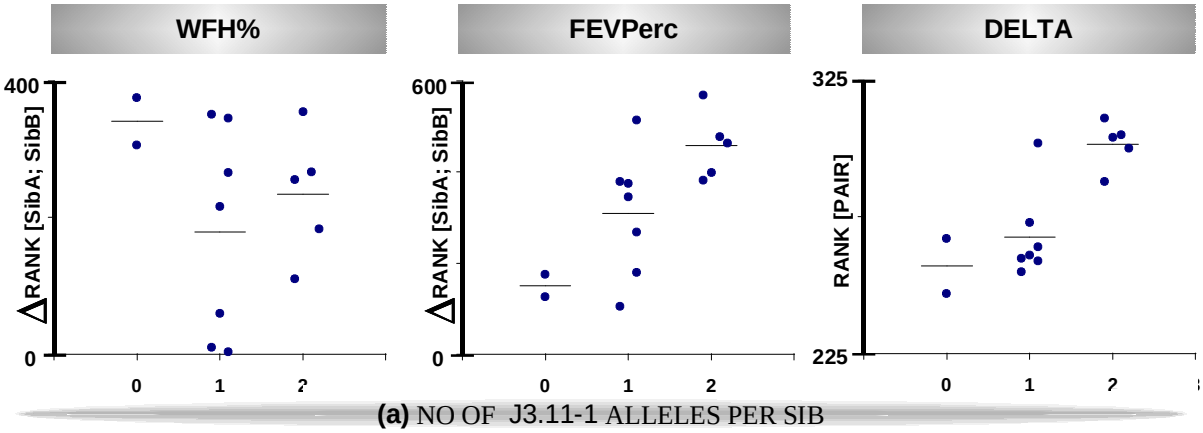
No of D7S514-12 alleles per pair	0	1	2	3	4
No of pairs ranked DIS	5	1	3	—	4

(c) DELTA *vs number of* D7S495-15 *alleles per pair*

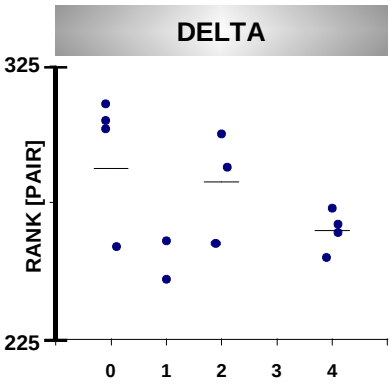
No of D7S495-15 alleles per pair	0	1	2	3	4
No of pairs ranked DIS	3	5	2	—	3

(d) Alleles IBS at D7S495 (pairs with 1 or 2 alleles D7S495-15 IBS are excluded as explained in footnote on page 119)

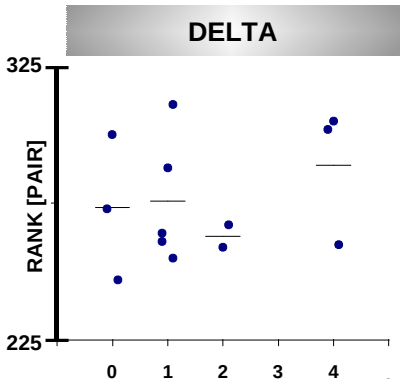
No of alleles IBS per pair	0	1	2
No of pairs CON+, CON- or DIS*	3	13	12



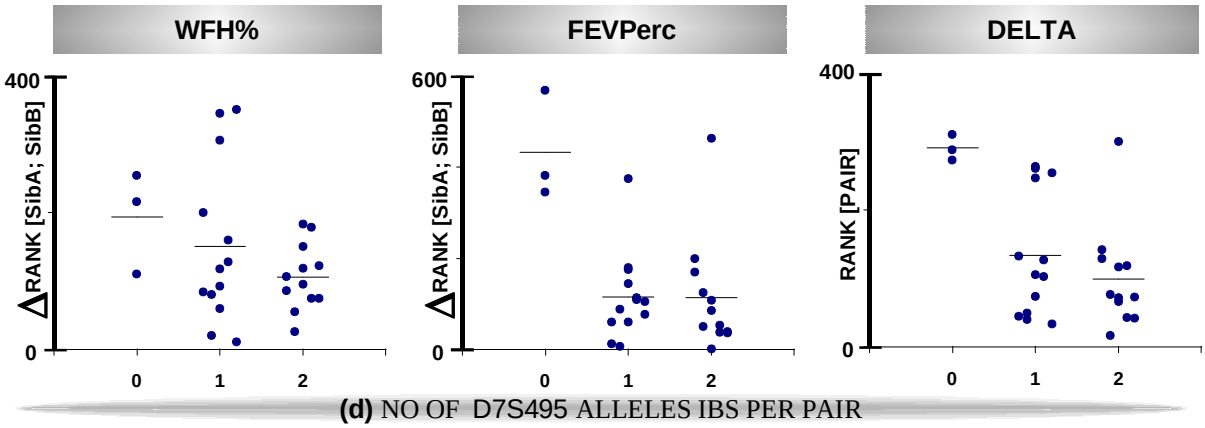
(a) NO OF J3.11-1 ALLELES PER SIB



(b) NO OF D7S514-12 ALLELES PER PAIR



(c) NO OF D7S495-15 ALLELES PER PAIR



(d) NO OF D7S495 ALLELES IBS PER PAIR

3.2.3.7.2.2 Modulation of intrapair discordance

The association found for J3.11-1 was seen to have a larger influence on the discordance of the pulmonary status than on the discordance of the nutritional status. This was demonstrated by higher ΔRANKFEVPerc and DELTA for pairs homozygous for J3.11-1 compared to the heterozygous state ($p < 0.01$). No association between J3.11-1 alleles and ΔRANKwfh% was observed (figure 39a). At D7S514 and D7S495, the most frequent alleles for discordant pairs were D7S514-12 and D7S495-15. However, no correlation between the number of D7S514-12 or D7S495-15 alleles within a pair and ΔRANKwfh%, ΔRANKFEVPerc or DELTA was detected (figure 39b,c) for all pairs ranked DIS.

When analysing the sharing of alleles at D7S495 after exclusion of pairs with one or two alleles D7S495-15 identical by state[22], intrapair discordance in wfh% and FEVPerc was found to decrease with an increasing number of alleles shared by both siblings.

3.2.3.7.2.3 Reconstruction of haplotypes surmised to transmit extreme phenotypes

So far, the genotype-phenotype correlation was investigated for single loci within the *CFTR* linkage group within this chapter. Next, the additive effect of all loci for which alleles have been observed in association with an extreme phenotype is assessed. The resulting allele combination, i.e. the haplotype for the chromosomal region, was considered to transmit the phenotype for which the allelic association was noticed. As in the case of single loci, the maximum effect was presumed to occur when all four chromosomes of both patients of a pair were carrying the haplotype considered to transmit the extreme phenotype.

In order to compare a pair's genotype to these proposed haplotypes, the similarity of a pair's genotype to homozygosity for a haplotype has to be described. In order to do so for a single **m**arker locus "m" exhibiting an association of the **a**llele "a" with a phenotypic trait, the genotype of a pair was construed to the number of alleles "a" encountered within the genotypes of both patients within a pair. In order to quantify the similarity of a pair's genotype to homozygosity for a haplotype "a_1-a_2 -a_3- ... -a_n" composed of multiple marker loci "m_1-m_2-m_3- ... -m_n", a similar approach was taken: at all loci composing the haplotype the number of alleles "a_1", "a_2", "a_3", ... ,"a_n" observed in the genotype of both patients within a pair was counted.

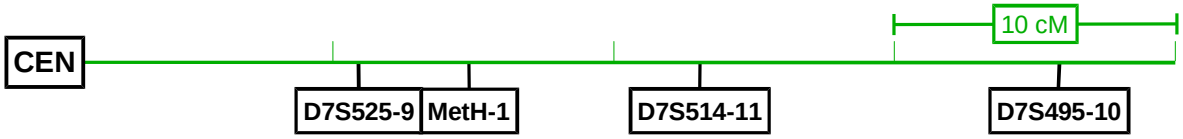
This approach considers each locus composing the haplotype as equal. However, neighbour loci are linked and consequently alleles at adjacent marker loci are not independent of each other. The probability of encountering a definite allele at locus "m_2" depends on the allele observed at locus

22 The hypotheses tested by (a) comparison of the number of alleles identical by state between pairs ranked concordant and pairs ranked discordant or (b) analysing allelic association with the phenotypes concordant and discordant are incompatible when an allelic association with the phenotype "discordance" is detected: when analysing a gene modulating the phenotype, discordance is expected when both sibs of a pair carry non-identical genotypes at the gene locus. If a marker locus analysed is linked to the gene, non-identical genotypes at the marker locus indicate non-identical genotypes at the gene locus. Overrepresentation of one allele in the subgroup of discordant pairs points to the hypothesis that an allele at the gene locus is itself causing the phenotype "discordance", even if —or, caused by the fact that — both sibs are carrying that allele identical by descent.

"m_1" or "m_3". The similarity of a pair's genotype to homozygosity for a given haplotype as defined above ignores this fact: deviations from the definite haplotype are weighed equally, irrespective of the marker position showing the deviation. Furthermore, all loci considered within this chapter are near *CFTR* and consequently, both siblings within a patient pair have a higher probability of sharing their alleles than of carrying non-identical genotypes. Linkage is only taken into account at loci J3.11 and MetH: no recombination was observed at either loci located within a 1 Mb range to the *CFTR* gene (see tables 26 and 27 for details). Finally, the probability of encountering a definite allele at a marker locus depends on the frequency of that allele and on the number of alleles displayed at that locus. Allelic distributions are not taken into account within this chapter.

TABLE 26: D7S525-MetH-D7S514-D7S495 HAPLOTYPE 9-1-11-10

The map displays all markers for which the alleles shown were detected in association with the phenotype "concordant/mild disease" (see chapter 3.2.3.5). The table displays the maximum number of these alleles encountered per genotype of a pair at each locus. For the markers D7S525, D7S514 *and* D7S495, *a maximum of four alleles* D7S525-9, D7S514-11 *and* D7S495-10 *can compose the genotype of the pair. At* MetH, *genotypes were, in all but one case investigated, identical between sib A and sib B of a pair. This lack of recombination was compensated for by arbitrarily counting only the alleles observed per single genotype, yielding a maximum of two alleles* MetH-1 *per pair.*



(1)	Marker	D7S525	MetH	D7S514	D7S495	
(2)	Alleles at loci (1) associated with phenotype "concordant / mild disease"	D7S525-9	MetH-1	D7S514-11	D7S495-10	
	Maximal number of alleles (2) observed per genotype of a pair:	4	2	4	4	**Σ: 14**

Observed number of alleles D7S525-9, MetH-1, D7S514-11 *and* D7S495-10 *for pairs ranked DIS, CON+ and CON-*

n: *number of alleles* 9, 1, 11, 10 *at loci* D7S525, MetH, D7S514, D7S495 *observed per pair*
av. *: average number of alleles* D7S525-9, MetH-1, D7S514-11 *and* D7S495-10 *observed per pair*
av. ={Σ (n × number of pairs carrying n alleles)}/{sum of all pairs}

	n *number of alleles* 9, 1, 11, 10 *at loci* D7S525, MetH, D7S514, D7S495 *observed per pair*												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**av.**
number of pairs ranked CON+ with **n** alleles	—	1	—	1	1	1	1	1	—	1	1	1	**7**
number of pairs ranked CON- with **n** alleles	—	1	4	3	—	1	—	—	—	—	—	—	**4**
number of pairs ranked DIS with **n** alleles	2	2	2	2	1	1	—	1	1	—	—	—	**4**

Based on the results summarised in table 28, haplotypes were reconstructed for the traits "discordant" (involving loci J3.11, D7S514 and D7S495) and "concordant / mild disease" (involving loci D7S525, MetH, D7S514 and D7S495): taking the observed associations of the alleles D7S525-9, MetH-1, D7S514-11 and D7S495-10 with the phenotype "concordant / mild disease" into account, the concluding D7S525-MetH-D7S514-D7S495 haplotype is 9-1-11-10. Likewise, the association of alleles J311.-1, D7S514-12 and D7S495-15 with the phenotype "discordant" leads to the J3.11-D7S514-D7S495 haplotype 1-12-15.

The results are displayed in table 26 for the D7S525-MetH-D7S514-D7S495 haplotype 9-1-11-10 (surmised as transmitting the phenotype "concordant / mild disease"). Homozygosity for the haplotype 9-1-11-10 corresponded to a total of 14 matching alleles encountered at these 4 loci. On the average, a pair ranked CON+ displayed 7, but a pair ranked DIS or CON- only 4 matching alleles. Genotypes with 10, 11 or 12 alleles encountered in the haplotype 9-1-11-10 were found only in pairs ranked CON+. The correlation of the genotype – interpreted as similarity to homozygosity for the D7S525-MetH-D7S514-D7S495 haplotype 9-1-11-10 – to the clinical phenotype is displayed in figure 40 for all pairs ranked CON+ or CON-. An increasing number of matching alleles was found to correlate with a better wfh%, but not FEVPerc.

FIGURE 40: SIMILARITY TO "CON+ HAPLOTYPE" AND DISEASE SEVERITY

Similarity of a pairs genotype to homozygosity for the haplotype 9-1-11-10 (D7S525-MetH-D7S514-D7S495)
All pairs ranked CON+ or CON- with information on all four loci are shown (18 pairs).

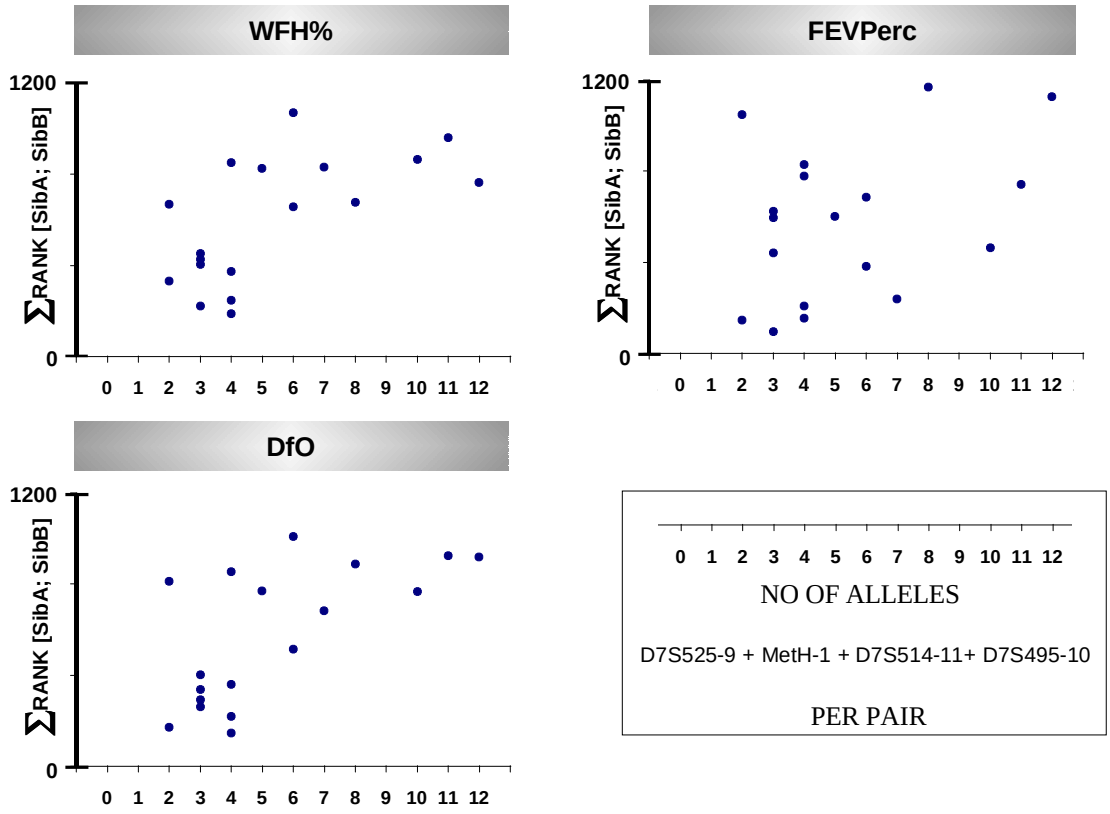
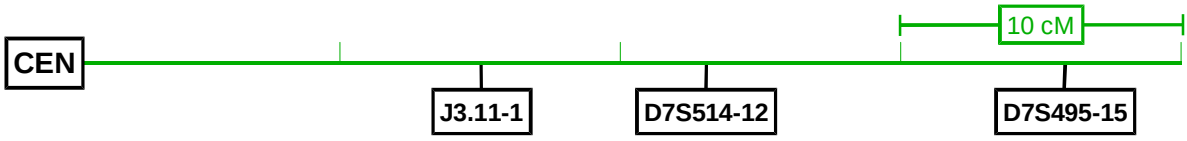


Table 27 displays the results for the J3.11-D7S514-D7S495 haplotype 1-12-15 (suspected as transmitting the phenotype "discordance"). A total of 10 matching alleles corresponded to homozygosity for the haplotype at these 3 loci. On the average, a pair ranked DIS displayed 5, but a pair ranked CON+ or CON- only 3 matching alleles. For pairs ranked concordant, a maximum of 5 matching alleles was observed while pairs ranked DIS also displayed 6 or 7 alleles matching the haplotype. The correlation of the genotype – converted into similarity to homozygosity for the J3.11-D7S514-D7S495 haplotype 1-12-15 – to the intrapair discordance in the clinical parameters is displayed in figure 41 for all pairs. No quantitative relationship was observed between the number of matching alleles and the intrapair discordance in wfh%, FEVPerc or the combined parameter DELTA.

TABLE 27: J3.11-D7S514-D7S495 HAPLOTYPE 1-12-15

The map displays all markers for which the alleles shown were detected in association with the phenotype "discordant" (see chapter 3.2.3.5). The table displays the maximal number of these alleles encountered per genotype of a pair at each locus. For the markers D7S514 *and* D7S495*, a maximum of four alleles* D7S514-12 *and* D7S495-15 *can compose the genotype of the pair. At* J3.11*, genotypes were, in all but one case investigated, identical between sib A and sib B of a pair. This lack of recombination was compensated for by arbitrarily counting only the alleles observed per one genotype, yielding a maximum of two alleles* J3.11-1 *per pair.*



(1)	Marker	J3.11	D7S514	D7S495	
(2)	Alleles at loci (1) associated with phenotype "discordant"	J3.11-1	D7S514-12	D7S495-15	
	Maximal number of alleles (2) observed per genotype of a pair:	2	4	4	**Σ: 10**

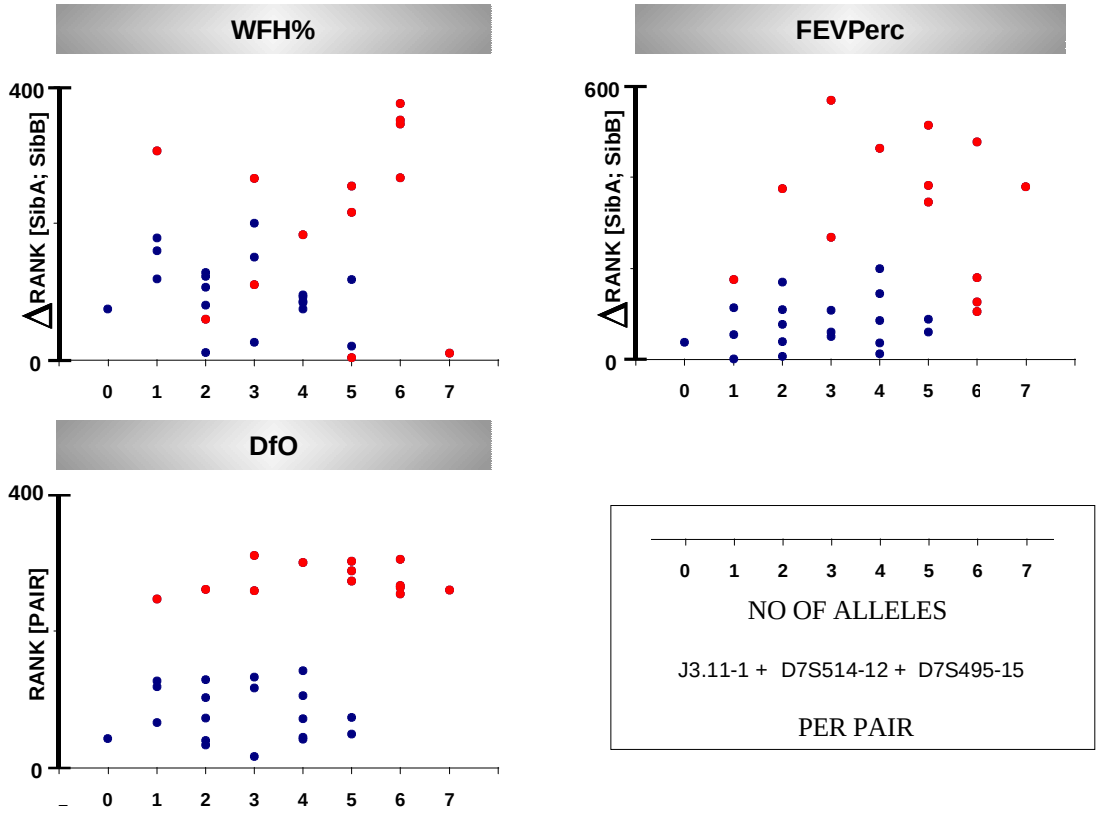
Observed number of alleles J3.11-1, D7S514-12 *and* D7S495-15 *for pairs ranked DIS, CON+ and CON-*

n: *number of alleles* 1, 12, 15 *at loci* J3.11, D7S514, D7S495 *observed per pair*
av. *: average number of alleles* J3.11-1, D7S514-12 *and* D7S495-15 *observed per pair*
av. ={Σ (n × number of pairs carrying n alleles)}/{sum of all pairs}

	n *number of alleles* J3.11-1, D7S514-12 *and* D7S495-15 *observed per pair*								
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**av.**
number of pairs ranked CON+ with **n** alleles	1	2	1	2	3	1	—	—	**3**
number of pairs ranked CON- with **n** alleles	—	1	4	1	2	1	—	—	**3**
number of pairs ranked DIS with **n** alleles	—	1	1	2	1	3	4	1	**5**

FIGURE 41: SIMILARITY TO "DIS HAPLOTYPE" AND DISCORDANCE

Similarity of a pairs genotype to homozygosity for the haplotype 1-12-15 (J3.11-D7S514-D7S495)
● : pairs ranked DIS *(13 pairs)* ● : pairs ranked CON+ or CON- *(19 pairs)*



3.2.3.7.3 Locus D2S1788

At D2S1788, an allelic imbalance was noticed: allele D2S1788-18 was noted only for pairs ranked DIS and allele D2S1788-21 was detected predominantly for pairs ranked DIS. The analysis of genotype-phenotype relations with respect to locus D2S1788 was performed by comparing the number of alleles identical by state with the intrapair discordance in the clinical parameters wfh% and FEVPerc and the combined parameter DELTA. As shown in figure 42, a decrease of the intrapair discordance in wfh% with an increasing number of identical alleles between siblings was noted only when pairs sharing the alleles found associated with the phenotype "discordance" were excluded (marked by red data points in the top row in figure 42 and omitted in the bottom row; $p < 0.01$ comparing pairs with 0 alleles IBS to pairs with 1 or 2 alleles IBS in wfh%).

The hypotheses tested by (a) comparing the number of alleles identical by state between pairs ranked concordant and pairs ranked discordant or (b) analysing allelic association with the

phenotypes concordant and discordant are incompatible when an allelic association with the phenotype "discordance" is detected: when analysing a gene modulating the phenotype, discordance is expected when both sibs of a pair carry non-identical genotypes at the gene locus. If a marker locus analysed is linked to the gene, non-identical genotypes at the marker locus indicate non-identical genotypes at the gene locus. Overrepresentation of one allele in the subgroup of discordant pairs points to the hypothesis that an allele at the gene locus is itself causing the phenotype "discordance", even if —or, caused by the fact that — both sibs are carrying that allele identical by descent. Consequently, the correlation of intrapair discordance and allele sharing was carried out after omitting pairs sharing alleles associated with the phenotype "discordance".

FIGURE 42: IBS AT D2S1788

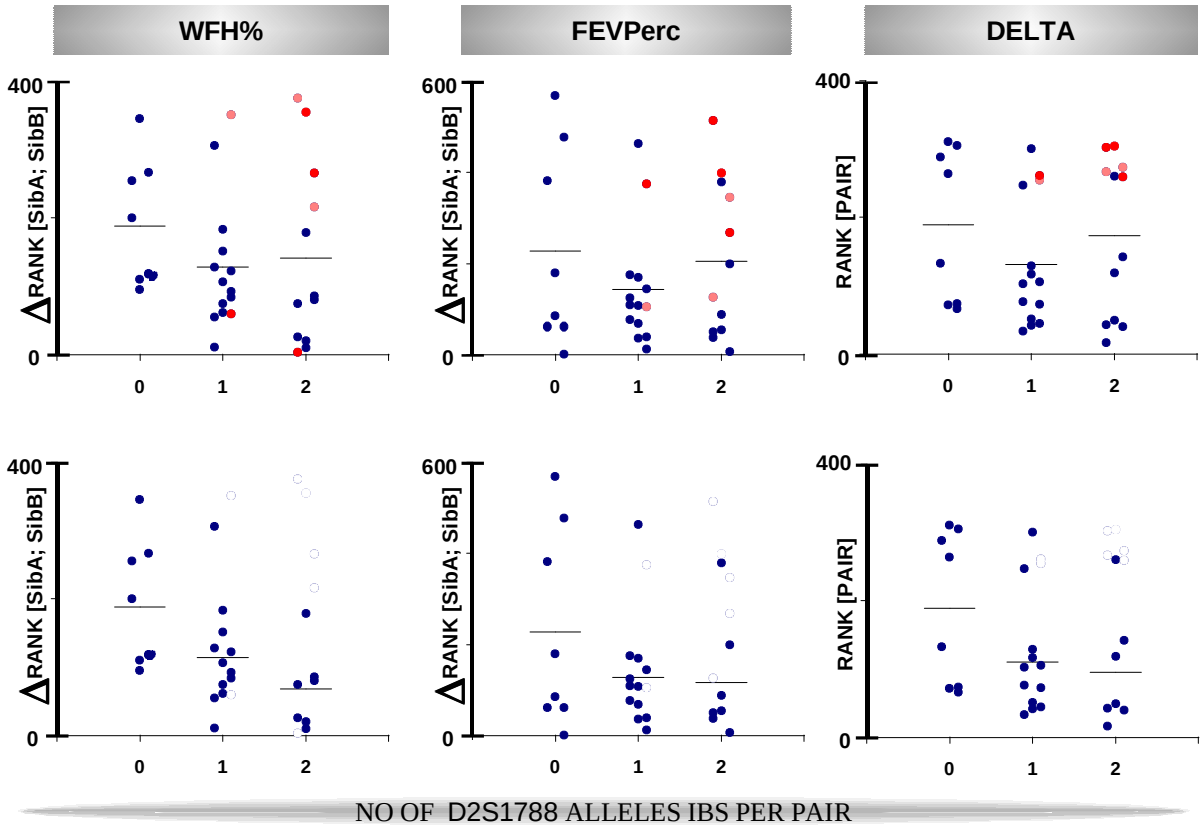
(a), (b) ΔRANKwfh%, ΔRANKFEVPerc *and* DELTA *vs number of* D2S1788 *alleles IBS per pair*

Top row: *all pairs*

No of D2S1788 alleles IBS	0	1	2
No of pairs	8	14	12

: pairs sharing one or two alleles 18 or 21 IBS : other pairs

Bottom row: *pairs sharing one or two alleles D2S1788-18 or D2S1788-21 excluded*



3.2.3.8 Discussion: CF disease severity modulation by non-*CFTR* genetics

3.2.3.8.1 Hypothesis: CF disease severity is modulated by ENaC

As shown in chapter 3.2.3.7.1, a D12S889 allele was observed to be associated with a mild clinical course in CF twins and siblings : ΔF508 homozygous patient pairs had a better wfh% and FEVPerc when both siblings carried at least one allele D12S889-10.

D12S889 is located within a 40 kb distance to SCNN1A (Baens *et al.* 1995), encoding for the α subunit of the amiloride sensitive sodium channel ENaC. Presumably, the observed effect is transmitted by the ENaC itself: the amiloride-sensitive sodium channel was reported to be regulated by CFTR (Ismailov *et al.* 1996, Stutts *et al.* 1997) and thus, sequence variations within the α subunit that convey a different phenotype to the channel might compensate for the regulatory defect in the CF condition. For the ENaC, allelic variants leading to a "gain of function " and allelic variants leading to a "loss of function" have been described: mutations resulting in increased Na^+ reabsorption have been discovered in SCNN1B and SCNN1G in patients with Liddle's disease (Shimkets *et al.* 1994, Hansson *et al.* 1995, Schild *et al.* 1995, Jeunemaitre *et al.* 1997). Sequence variations leading to decreased Na^+ reabsorption have been identified in patients with pseudohypoaldosteronism type 1 in SCNN1A, SCNN1B and SCNN1G (Chang *et al.* 1996, Straunieks *et al.* 1996b).

Apart from a postulated benign effect transmitted by an allelic variant of the α subunit of the ENaC (presumably linked to the allele D12S889-10), the amiloride-sensitive Na^+ channel can also be considered to be a risk factor as patients with pseudohypoaldosteronism type 1 have been described as exhibiting symptoms of lung disease compatible with mild CF (Hanakoglu *et al.* 1994), including colonisation with *Pseudomonas aeruginosa* (Marthinsen *et al.* 1998). Nevertheless, no D12S889 allele was identified in association with a severe phenotype in CF twins and sibs. Finally, an association of a benign phenotype was observed with a D12S889 allele (located close to SCNN1A), but not with a βENaCGT allele (located near SCNN1B and SCNN1G), inviting speculation on a dominant role of the α subunit in modulating CF disease, distinct from the β and γ subunits of the ENaC.

Generally, the detection of an association of an allelic variant at a polymorphic marker locus with a phenotypic character is eased (a) if alleles at the candidate gene locus transmitting the same phenotype are linked to the same allele at the marker locus investigated and (b) if two alleles at the candidate gene locus conveying the opposite phenotypes are linked to different alleles at the marker locus investigated. Hence, the lack of association between any D12S889 allele with severe CF disease among the twins and sibs typed and the lack of association of any βENaCGT allele with an extreme disease phenotype fail to exclude a risk allele at SCNN1A or the modulating role of SCNN1B and SCNN1G for CF disease severity.

In conclusion, the observed effect of an association of mild CF disease with the allele D12S889-10 can be explained by assuming a benign SCNN1A allele. However, other allelic variants at SCNN1A, conveying a risk for lung disease, might be existent. Likewise, a modulating effect of the SCNN1B and SCNN1G loci cannot be excluded.

3.2.3.8.2 CF disease severity modulated by CFTR linked loci

As described in detail within chapter 3.2.3.7.2, three effects on the clinical phenotype could be described as being associated with loci close to the *CFTR* gene (table 28):

(I) Located further towards the centromer than *CFTR*, modulation of wfh% was observed at MetH and D7S525. The effect was seen in an allelic association of allele MetH-1 and, less pronounced, allele D7S525-9 with the phenotype "concordant / mild disease". Pairs of sibs whose genotype was composed of the allele MetH-1 had a significantly better wfh% than sibpairs with other genotypes ($p < 0.01$).

(II) Located further towards the telomer than *CFTR*, an influence on the pulmonary status was noticed at J3.11. The effect was detected by an association of the allele J3.11-1 with the phenotype "discordant". Pairs of sibs homozygous for J3.11-1 were significantly more discordant in FEVPerc ($p < 0.01$), but not in wfh%, compared to other sibpairs.

(III) An influence on the nutritional status was detected at D7S495 by an association of allele D7S495-10 with the phenotype "concordant / mild disease", of the allele D7S495-14 with the phenotype "concordant / severe disease" and of the allele D7S495-15 with the phenotype "discordant". Pairs where each sib carried at least one allele D7S495-10 had a significantly better wfh% ($p < 0.01$) and pairs of sibs carrying at least one allele D7S495-14 had a significantly worse wfh% ($p < 0.01$) than pairs of sibs with other genotypes. When excluding pairs composed of sibs carrying one or two alleles D7S495-15, those pairs whose sibs were sharing two D7S495 alleles were less discordant than sibpairs carrying one or no allele IBS at D7S495.

All effects observed are not directly mediated by *CFTR*: both patients of a pair share the *CFTR* alleles as CF is inherited in the recessive condition and all pairs have been typed ΔF508 homozygous. Consequently, unless recombination events have taken place, both siblings will have identical sequences at loci near *CFTR* as well. Hence, an observed allelic association can only be mediated by a sequence close to the marker typed – i.e. by an element encoded in *cis* – if the association is observed for pairs ranked ***concordant*** (allele MetH-1 associated with "concordant / mild disease" (I), allele D7S495-10 associated with "concordant / mild disease" (III)and allele D7S495-14 associated with "concordant / severe disease" (III)).

However, at two loci located further towards telomer than *CFTR*, an association of an allele with the phenotype ***discordant*** has been detected (allele J3.11-1 associated with "discordant" (II) and allele D7S495-15 associated with "discordant" (III)). At J3.11, no recombination event was observed among the pairs typed (table 23[23]). At D7S495, with an observed recombination frequency 1/3, discordant sibs shared the allele D7S495-15 which was observed predominantly for pairs ranked discordant in absence of evidence for a recombination event between *CFTR* and D7S495 (table 23). As this allelic association cannot be related to either a recombination event or an element encoded in

23 Table 23 is shown in chapter 3.2.3.5.2.4 on page 111.

cis, the modulating effect has to be located outside the *CFTR* linkage group, probably on another chromosome. In conclusion, an allelic association with the phenotype ***discordant*** observed at loci linked to *CFTR* indicates a modulating factor encoded in *trans* acting on a sequence close to the locus studied.

TABLE 28: MODULATION OF WFH% AND FEVPerc AT LOCI NEAR *CFTR*

ALLELE	ALLELIC IMBALANCE OBSERVED FOR PAIRS RANKED	+: strong allelic association (+): faint allelic association	EFFECT ON PHENOTYPE OBSERVED FOR	PAIRS ANALYSED	⋆
D7S525-9	CON+	(+)	ΣRANKwfh%	CON+, CON-	$p > 0.05$ ⋆
MetH-1	CON+	+	ΣRANKwfh%	CON+, CON-	$p < 0.01$ ⋆
J3.11-1	DIS	+	ΔRANKFEVPerc	DIS	$p < 0.01$ ⋆
D7S514-11	CON+	(+)	ΣRANKDfO	CON+, CON-	$p > 0.05$ ⋆
D7S514-12	DIS	(+)	—	DIS	—
D7S495-10	CON+	+	ΣRANKwfh%	CON+, CON-	$p < 0.01$ ⋆
D7S495-14	CON-	+	ΣRANKwfh%	CON+, CON-	$p < 0.01$ ⋆
D7S495-15	DIS	+	—	DIS	—
D7S495	Analysis of IBS at D7S495 shows decreasing intrapair rank number difference in wfh% with increasing number of shared alleles.			all pairs	$p > 0.05$ ⋆

⋆*: quantitative influence of allele on clinical parameter judged by Mann-Whitney rank test of values shown in figures 37 (a), 37 (b), 38 (b), 38 (c), 39 (a), 39 (d)*

— : no quantitative influence observed
$p > 0.05$: quantitative influence observed, but not statistically significant
$p < 0.01$: statistically significant quantitative influence detected

3.2.3.8.2.1 Selected genes mapped near *CFTR*

Figure 43 shows a map of the investigated 40 cM region encompassing the *CFTR* gene, displays the effects observed and gives the position of some selected candidate genes. Based on the known functions of the gene products, the candidate genes were classified with respect to their roles:

A) Involved in energy metabolism and/or bioenergetics

PRKAR2B *(protein kinase, cAMP-dependent, type II, beta)*

PRKAR2B knock-out mice have been described as "healthy, lean and protected against developing dietary induced obesity and fatty livers" with reduced leptin mRNA and plasma levels (Cummings *et al.* 1996). PRKAR2B was mapped to 7q22 near MET (Solberg *et al.* 1992) and located on a physical map 600 Mb further towards the centromer than DRA (Höglung *et al.* 1995).

***LEP/OB** (leptin; human homologue of murine obese)*

The human homologue of the murine obese gene LEP was physically mapped at 7q31.3 (Green *et al.* 1995). The involvement of leptin in the long term regulation of energy balance has been characterised (reviewed by Woodsa *et al.* 1998): the gene product leptin is secreted in proportion to the adipose tissue. Low plasma leptin levels result in increased food intake and weight gain.

***PEG1/MEST** (paternally expressed gene 1; mesoderm specific transcript)*

PEG1/MEST was identified as an imprinted gene near D7S649 (Kobayashi *et al.* 1997) with homology to a gene known to be imprinted in mice. Imprinting was verified by demonstrating that the paternal allele of PEG1 was expressed. A role of PEG1 in energy metabolism is inferred based on the observation that maternal isodisomy 7 leads to growth retardation (Spence *et al.* 1988, Voss *et al.* 1989, Spotila *et al.* 1992, Eggerding *et al.* 1994) while paternal isodisomy 7 does not (Höglund *et al.* 1994).

***NRF1** (nuclar respiratory factor 1)*

NRF1 has been identified as a transcription factor for mitochondrial proteins (Gopalakrishnan and Scarpulla 1995). The genes responsive to NRF1 have been identified as components of the oxidative chain. The gene was mapped to 7q32.

B) Tissue differentiation, integrity and development

***MET/HGFR** (met proto oncogene; hepatocyte growth factor receptor)*

The proto oncogene MET encodes for a receptor tyrosine kinase (Dean *et al.* 1985). The ligand of MET is the hepatocyte growth factor HGF (Bottaro *et al.* 1991). MET and HGF are involved in the mediation of motility, invasiveness and proliferation of target cells (reviewed by Weidner 1993).

***PAX4** (paired box homeotic gene 4)*

The PAX4 gene product was described as a regulator for the differentiation of pancreatic cells (Sosa-Pineda *et al.* 1994). It has been localised between *CFTR* and LEP (Green *et al.* 1995).

***CASP2/NEDD2** (caspase 2; apoptosis related cysteine protease)*

Caspase-2 deficient mice have been demonstrated as exhibiting a dual apoptosis defect: while some CASP2 deficient cells – such as motor neurons – showed an accelerated cell death, other CASP2 deficient cell types – such as B lymphoblasts – demonstrated resistance against apoptosis signals (Bergeron *et al.* 1998). CASP2 was mapped to 7q35 (Tiso *et al.* 1996).

C) Involved in host defence, modulation of infection or inflammation

***NM** (neutrophil migration)*

NM, presumed to encode for the surface glycoprotein GP130 was mapped to 7q22-7qter (de la Chapelle *et al.* 1979). Granulocytes that are defective in locomotion, chemotaxis and chemokinesis were shown to express reduced amounts of GP130 (de la Chapelle *et al.* 1982). Thus, NM is most likely to be involved in general host defence.

FIGURE 43: INTEGRATED MAP OF SELECTED GENES NEAR CFTR

The relative positions of the map elements are consistent with Tsui 1985, Rommens 1989, Green 1995, Höglund 1995, Bouffard 1997, Fisher 1998, Li 1998. Positions in cM are given as published by Bouffard 1997.

MARKER: *of the displayed markers, all but* D7S496 *and* D7S649 *have been typed.*

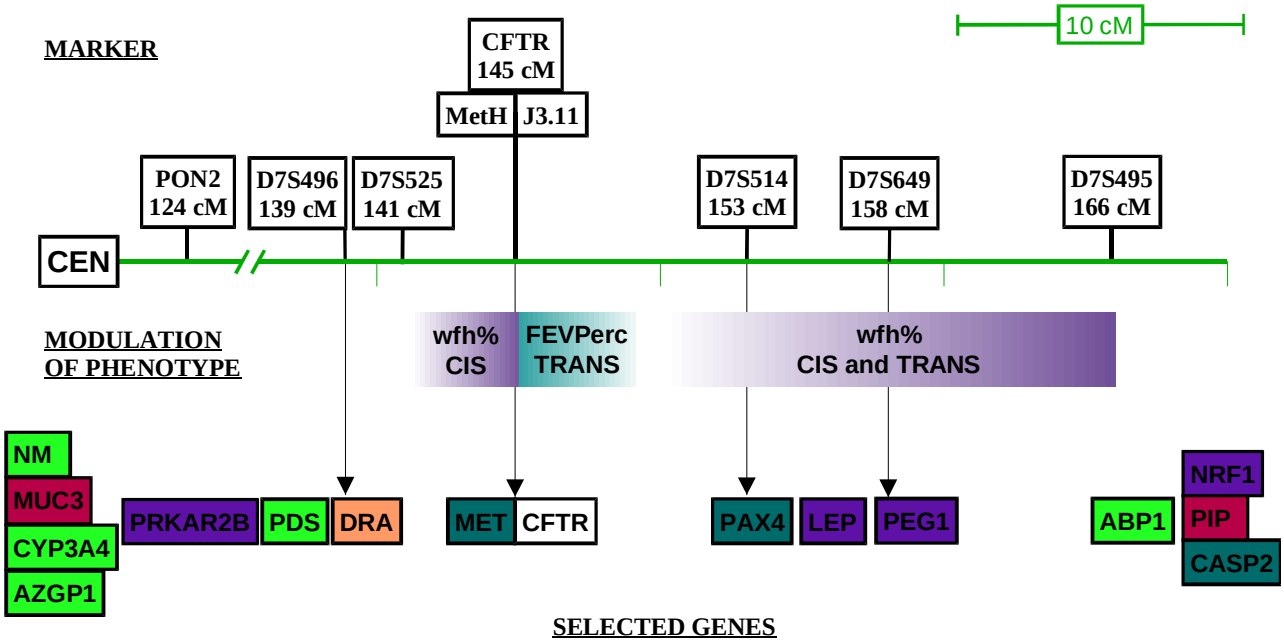
MODULATION OF PHENOTYPE: *representation of observed effects as summarised on page 126: association of alleles at* METH, J3.11, D7S514 *and* D7S495 *with disease severity.* CIS *designates an effect disclosed for pairs ranked concordant while* TRANS *refers to an effect observed for pairs ranked discordant.*

SELECTED GENES: *the genes displayed are abbreviated according to the HUGO/GDB approved nomenclature. Map positions are inferred from cytogenetic localisation using* PON2 *at 7q21.3,* DRA/CLD *at 7q22-7q31.1 (mapped near* D7S496*),* LEP *at 7q31.3 (mapped near* D7S514*) and* D7S495 *at 7q32-7q33 as reference points.*

Role of gene product ‡	Symbol		*Localisation*
C	**NM**	neutrophil migration	*7q22-qter*
D	**MUC3**	mucin 3	*7q22*
C	**AZGP1**	alpha-2-glycoprotein 1, zinc	*7q22.1*
C	**CYP3A4**	cytochromeP450, subfamily IIIA, polypeptide 4	*7q22.1*
A	**PRKAR2B**	protein-kinase, cAMP-dependent, type II, beta	*7q22*
C	**PDS**	pendred syndrome (prostaglandin-h2 d-isomerase precursor)	*7q31*
E	**CLD/DRA**	congenital chloride diarrhoea /down-regulated in adenoma	*7q22-7q31.1*
B	**MET**	met proto oncogene (hepatocyte growth factor receptor)	*7q31*
B	**PAX4**	paired box homeotic gene 4	*7q22-qter*
A	**LEP**	leptin (murine obesity homolog)	*7q31.3*
A	**PEG1**	paternally expressed gene 1 (mesoderm specific transcript)	*7q32*
C	**ABP1**	amiloride-binding protein 1	*7q31-7q32*
A	**NRF1**	nuclear respiratory factor 1	*7q32*
D	**PIP**	prolactin-induced protein	*7q32-q36*
B	**CASP2**	caspase 2 (apoptosis related cysteine protease)	*7q35*

‡ *Functional classification of gene products (details see chapter 3.2.3.8.2.1):*

■	A	*role in energy metabolism and/or bioenergetics*
■	B	*role in tissue differentiation, integrity and development*
■	C	*involved in host defence, modulation of infection or inflammation*
■	D	*CFTR-like expression pattern or relation to CF pathophysiology*
■	E	*basic defect: ion transport*



***AZGP1** (alpha-2-glycorotein-1, zinc)*

AZGP1 was localised to 7q22.1 (Ueyama *et al.* 1993) and shown to encode for a soluble glycoprotein. AZGP1 is partly homologous to class I MHC genes and an involvement in intracellular recognition processes has been proposed for the AZGP1 gene product (Freije *et al.* 1993).

***PDS/PGHD** (pendred syndrome; prostaglandin-h2 d-isomerase precursor)*

PDS was identified as a gene causing pendred syndrome and mapped in a location 300kb further towards the centromer than DRA (Coyle *et al.* 1998). PDS was identified as prostaglandin isomerase.

***CYP3A4** (cytochrome P450, subfamily IIIA, polypeptide 4; nifedipine oxidase)*

CYP3A4 encodes for a member of the cytochrome P450 family (reviewed by Wrighton and Stevens 1992) and was mapped to 7q22.1 (Inoue *et al.* 1992). CF patients and their parents were described as poor metabolisers for the CYP3A4 substrate nifedipine compared to non-CF adults (Daly *et al.* 1992).

***ABP1** (amiloride binding protein)*

ABP1 encodes for a diamine oxidase (Novotny *et al.* 1994) and was mapped to 7q34-36 (Barbry *et al.* 1990).

D) CFTR like expression pattern or relation to CF pathophysiology

***MUC3** (mucin 3)*

MUC3 was described by Gum *et al.* (1990) as an intestinal mucin. The gene has been localised to 7q22 (Fox *et al.* 1992). The pathophysiology of CF makes a modulation of intestinal disease through MUC3 plausible.

***PIP** (prolactin induced protein)*

The PIP gene product is expressed in several exocrine organs and was mapped to 7q32-7q36 (Myal *et al.* 1989).

E) Basic defect: ion transport

***CLD/DRA** (congenital chloride diarrhoea; down-regulated in adenoma)*

The gene transmitting congenital chloride diarrhoea (CLD) was mapped near D7S496 (Höglund *et al.* 1995). DRA (down regulated in adenoma), located at a distance of 230 kb to D7S496 (Höglund *et al.* 1996a) was identified as the gene responsible for CLD (Höglund *et al.* 1996b). CLD was described as a disease with impaired electrolyte transport and hence, DRA qualifies as a candidate gene for the modulation of CF disease.

3.2.3.8.2.2 Hypothesis: modulation of pulmonary status by *CFTR* linked loci

Regulation of CFTR transcription near J3.11: a trans effect

The existence of a modifier gene near J3.11 for the severity of CF lung disease was noticed before in a group of unrelated patients (Santis *et al.* 1990b) where allele J3.11-2 was found to be associated with mild lung disease. Within this thesis, an allelic association of J3.11-1 with the phenotype "discordant" has been identified. ΔF508 homozygous patient pairs homozygous for J3.11-1 were significantly more discordant in FEVPerc than pairs with other J3.11 genotypes. As both siblings, though discordant, must share the same sequence at J3.11, an effect mediated by a gene located elsewhere – i.e. in *trans* – has to be assumed.

A variety of genes located within the *CFTR* linkage group have been mapped (chapter 3.2.3.8.2.1). However, no clear evidence for a modulation of lung disease by a gene located near J3.11 was found. Generally, all gene products participating in host defence, modulation of inflammation or immunology – as NM, CYP3A4, AZGP1, PDS and ABP1 – can be considered capable of modulating respiratory illness. Nevertheless, the only gene located in the region of interest with a known role in respiratory disease is the *CFTR* gene itself (figure 43). The role of *CFTR* mutations as a risk factor for lung disease conditions are well established (Dumur *et al.* 1990, Miller *et al.* 1996, Pignatti *et al.* 1996, Friedman *et al.* 1997, Dahl *et al.* 1998) and hence, the most plausible explanation for the effect observed at J3.11 is a modulation of CFTR expression by a transcription factor that is acting on a regulatory sequence linked to J3.11. Intrapair-discordance could then arise if both siblings have inherited different alleles for the transcription factor, which consequently act differently on a responsive element near the *CFTR* gene. It has to be postulated that this responsive element occurs in different allelic variations whereby the allele associated with J3.11-1 might depend most on the allelic variation of the transcription factor, hence resulting in a discordant phenotype when siblings are homozygous for J3.11-1.

3.2.3.8.2.3 Hypothesis: modulation of nutritional status by *CFTR* linked loci

Involvement of LEP in the inheritance of obesity

Obesity is presumed to be 40-70% inheritable and is considered to be a multigenic trait. 17 genes with homologies to murine obesity loci have been identified so far (reviewed by Comuzzie 1998). Linkage analysis and some association studies have provided evidence for an involvement of the LEP gene region in obesity related traits (Duggirala *et al.* 1996, Clement *et al.* 1996, Reed *et al.* 1996, Oksanen *et al.* 1997), but two research groups found no evidence for a role of LEP in obesity (Norman *et al.* 1996, Bray *et al.* 1996). Nevertheless, a compilation of 5 studies has provided evidence for strong association of the body mass index to markers within the LEP gene region (Allison *et al.* 1998). Mutations within the LEP gene have been identified in patients suffering from morbid obesity (Montague *et al.* 1997, Strobel *et al.* 1998), thus sustaining the involvement of leptin

in obesity. Other genes suspected of being involved in the inheritance of obesity have been investigated in obese patients and mutations have been described for the leptin receptor (Clement *et al.* 1998).

Growth retardation caused by an imprinted gene on chromosome 7

Growth retardation has been described in patients with maternal isodisomy 7 (Spence *et al.* 1988, Voss *et al.* 1989, Spotila *et al.* 1992). In contrast, paternal isodisomy 7 was reported to be compatible with normal growth and development (Höglund *et al.* 1994). The imprinted gene was presumed to be localised on 7q as a patient with partial paternal isodisomy 7p and partial maternal isodisomy 7q was growth retarded (Eggerding *et al.* 1994). In addition, maternal isodisomy has been reported for 3 out of 25 patients with Silver-Russel-syndrome and for 1 out of 10 patients with primordial growth retardation (Kotzot *et al.* 1995). PEG1, a paternally imprinted gene, was mapped to 7q32 (Kobayashi *et al.* 1997). PEG1 is regarded as a candidate gene for the growth retardation caused by paternal isodisomy 7.

Obesity related traits and chromosome 7: evidence for more than one locus involved

Several approaches towards unravelling the role of the LEP gene region have investigated a subportion of chromosome 7. Interestingly, some results point to the existence of more than one locus on chromosome 7q determining obesity or obesity related traits (see table 29):

> Stirling *et al.* (1995) investigated patient pairs with non-insulin dependent diabetes mellitus (NIDDM) with respect to allele sharing at marker loci spanning a 4 cM region. While not statistically significant, a surplus of allele sharing was detected at S7S480 and D7S635, but not at the two tested markers localised in between.
>
> Duggirala *et al.* (1996) has found linkage of three different obesity-related traits to three different regions on chromosome 7 upon analysis of 15 markers spanning the whole chromosome: variation in extremity skinfolds were linked to D7S514, fasting 32,33-split proinsulin levels were linked to D7S530 and D7S479, waist circumference was linked to D7S495.
>
> Clement *et al.* (1996) has investigated sib pairs with respect to the sharing of alleles near LEP. In obese/obese pairs linkage to D7S651 and neighbouring markers D7S680, D7S514 and D7S530 was reported. In obese/lean pairs, an excess of sharing was found at D7S530 and D7S640.
>
> Oksanen *et al.* (1997) has studied association of serum leptin levels and response to a very-low-calory dietary program (VLCD) to markers flanking LEP in a group of patients suffering from morbid obesity. Weak association of an allele at D7S649 with response to VLCD and significant association of an allele at D7S530 with serum leptin levels was detected.

In conclusion, markers associated with obesity related traits have been identified at 7q21 and at 7q31-33. Among the candidate genes discussed within this thesis, LEP and PEG1 – both within the region 7q31-7q33 – can be considered to be involved in weight regulation and growth development. Analysing ΔF508 homozygous CF twins and sibling pairs, a significant association of MetH and

D7S495 genotypes with wfh% was noticed (figures 37[24] and 38[25]). While D7S495 is localised near the obesity-linked region containing LEP and PEG1, MET does not match the position of the obesity related locus at 7q21. Hence, the modulating effect on wfh% observed with markers located further towards 7qtel than *CFTR* can be related to LEP and/or PEG1 while, for the effect observed at MetH, a modifier which has not been identified in previous previous studies of the LEP region has to be considered.

TABLE 29: HYPOTHESIS: TWO LOCI DETERMINING OBESITY ON 7q

Abbreviations used in the table (details see text): ex shar = surplus of shared alleles in NIDDM sib pairs (Stirling et al. 1995), FSPL = fasting 32,33-split proinsulin levels, extr skflds = extremity skinfolds, wst circ = waist circumference (Duggirala et al. 1996), ob/ob = obese/obese sib pair, ob/l = obese/lean sib pair (Clement et al. 1996), VLCD resp = response to very-low-calorie-diet, lep lev = serum leptin levels (Oksanen et al. 1997), CON = concordant CF sib pair, DIS = discordant CF sib pair (this work).
All marker positions are given based on the chromosome 7 physical map (Bouffard et al. 1997). On this map, PON2 is localised at 124 cM, CFTR is localised at 145 cM, LEP is localised at 153 cM and PEG1 is localised at 158 cM.

POSITION:	near *PON2*	7q22-31	near *CFTR*	near *LEP*	near *PEG1*	7q32-33
Stirling *et al.* 1995 markers investigated:	—	—	D7S480 (147 cM)	D7S635 (153 cM)	D7S530 (157 cM)	—
obesity associated:			ex shar	—	ex shar	
Duggirala *et al.* 1996 markers investigated:	D7S479 (125 cM)	D7S471 (135 cM)	D7S466 (144 cM)	D7S514 (153 cM)	D7S530 (157 cM)	D7S495 (166 cM)
obesity associated:	FSPL	—	—	extr skflds	FSPL	wst circ
Clement *et al.* 1996 markers investigated:	D7S651 (128 cM)	D7S692 (140 cM)	D7S677 (144 cM)	D7S514 (153 cM)	D7S640 (159 cM)	D7S509 (165 cM)
obesity associated:	ob/ob	—	—	ob/ob and ob/l	ob/l	—
Oksanen *et al.* 1997 markers investigated:	—	—	—	D7S514 (153 cM)	D7S649 (158 cM)	—
obesity associated:				VLCD res	lep lev	
This work markers investigated:	—	D7S525 (141 cM)	MetH (145 cM)	D7S514 (153 cM)	—	D7S495 (166 cM)
wfh% modulation:		—	CON	—		CON and DIS

Modulation of wfh% near MetH: a cis effect

The putative modifier of CF disease detected through an association of MetH-1 alleles with the phenotype "concordant/mild disease" (table 19[26]) can be considered to be distinct from the candidate genes PRKAR2B, PDS and DRA as the significant association of wfh% with MetH genotypes was not seen at D7S525 (figure 37[27]), located within a 4 cM distance to MetH between MET and PRKAR2B, PDS and DRA (figure 43). Thus, the modulation of the CF phenotype might be related

24 Figure 37 is shown in chapter 3.2.3.7.2.1 on page 116.
25 Figure 38 is shown in chapter 3.2.3.7.2.1 on page 117.
26 Table 19 is shown in chapter 3.2.3.5.2.1 on page 107.
27 Figure 37 is shown in chapter 3.2.3.7.2.1 on page 116.

to MET itself. A possible mechanism for the influence of the nutritional status by MET relies on the signal transduction pathways controlled by this receptor tyrosine kinase: the ligand of MET, the hepatocyte growth factor, was shown to be involved in the mediation of motility and proliferation of various cell types (reviewed by Weidner 1993). As coordinated cell growth and cell differentiation are obligatory for tissues with rapid proliferation – as, for example, for the intestinal epithelium – a modulation of the respective signal transduction pathway might influence the integrity of the organ and hence modulate the nutritional status of the patient.

Modulation of wfh% by POMC: a trans effect mediated through leptin expression

Within this thesis, an association of alleles D7S514-12 and D7S495-15 with the phenotype "discordant" was observed (figures 33[28]and 34[29]). As both siblings share not only their *CFTR* genotype, but also flanking sequences unless recombination as occurred, the discordance has to be attributed to sequences which are not located at D7S514 or D7S495 but elsewhere in the genome, thereby facilitating the inheritance of different alleles to both siblings. A modulation of the phenotype by a modifier acting in *trans* will then be detectable through D7S514 and D7S495 if the modifying factor is interacting with a sequence near the marker loci , e.g. by modulating the expression of a gene through binding to a regulatory element. Discordance might then occur even if both siblings share their sequences at the regulatory element when they have inherited different alleles for the *trans* acting factor, thereby resulting in different expression of the target gene. An association of alleles D7S514-12 and D7S495-15 with the phenotypic trait "discordance" will then be noticed if the respective allele at the marker locus is linked to an allele of the regulatory element that is sensitive to the allelic heterogeneity of the *trans* acting factor. The two candidate genes LEP and PEG1 are located near D7S514 and D7S495. However, only for LEP has the regulation of gene expression been studied in detail.

It has been demonstrated that plasma leptin levels and the body mass index correlate to a certain degree (Oksanen *et al.* 1997). The expression of the murine OB gene was shown to be inducible by corticosteroids (de Vos *et al.* 1995). For humans, a genome wide linkage scan demonstrated linkage of D2S1788 to serum leptin levels (Comuzzie *et al.* 1997). D2S1788 is localised near POMC. The gene product of POMC is a precursor for the adrenocorticotropic hormone which induces glucocorticoid production of the adrenal glands. Hence, a modulation of LEP expression by POMC can be presumed. Recently, a mutation in the POMC gene was described for two patients suffering from early-onset obesity (Krude *et al.* 1998).

D2S1788 was typed in order to assess the role of leptin expression in CF disease severity. At D2S1788, a significant decrease of intrapair discordance in wfh% with the number of shared alleles at D2S1788 was found (figure 42[30]). Thus, modulation of leptin expression by POMC might play a role in the modulation of CF disease severity.

28 Figure 33 is shown in chapter 3.2.3.5.2.1 on page 108.
29 Figure 34 is shown in chapter 3.2.3.5.2.3 on page 110.
30 Figure 42 is shown in chapter 3.2.3.7.3 on page 124.

4 SYNOPSIS

4.1 *CFTR* genotype – the basis of CF disease phenotype?

4.1.1 Determination of CF disease features by the CFTR genotype

Genetic determination was inferred for the pancreatic status based on high familial concordance (Corey *et al.* 1989). Mutations associated with pancreatic sufficiency were classified as "mild" while those conveying pancreatic insufficiency were termed "severe" (Kristidis *et al.* 1992). A direct correlation of *CFTR* genotype and phenotype might be expected for the sweat gland as this is the only organ where the condition "CF carrier" can be distinguished from the condition "two wild type *CFTR* alleles" (Behm *et al.* 1987). Consistent with the theory, elevated sweat chloride concentrations were observed for patients carrying mutations surmised to yield no or only non-functional CFTR protein while lower sweat chloride concentrations were observed for patients carrying a *CFTR* allele associated with reduced chloride channel function (Wilschanski *et al.* 1995). Inconsistent with the theory, mutations shown to produce low levels of wild type protein were not associated with a lower sweat chloride concentration (Wilschanski *et al.* 1995). No tight correlation between *CFTR* genotype and CF disease phenotype was noticed for the pulmonary status as pulmonary disease was described as being variable among patients carrying the same *CFTR* mutation genotype (Kerem *et al.* 1990, Johannsen *et al.* 1991, Gasparini *et al.* 1992, Shoshani *et al.* 1992, Highsmith *et al.* 1994). However, the risk for colonisation of the airways with *Pseudomonas aeruginosa* could be related to the *CFTR* genotype (Kubesch *et al.* 1993) whereby nonsense mutations were associated with the highest risk. For adult CF patient pairs, concordance among siblings in pulmonary parameters has been reported (Santis *et al.* 1990a). For a group of adult patients, the severity of lung disease was associated with the *CFTR* mutation genotype (Hubert *et al.* 1996).

In summary, the pancreatic status was correlated with the *CFTR* genotype. The sweat chloride concentration was mostly, but not always, deducible from the functional defect expected for the *CFTR* mutation. A high degree of variability was described for the pulmonary phenotype whereby correlation of genotype and phenotype seemed to be closer for adult patients.

4.1.2 Disease severity of patients homozygous for non-ΔF508 genotypes

Atypical CF has been described as "*Pseudomonas* bronchitis, associated with low sweat chloride concentrations and pancreatic sufficiency" before the *CFTR* gene was identified (Stern *et al.* 1978). Expression of symptoms restricted to only one of the many organs typically affected in CF has been disclosed for patients homozygous for G551S (normal sweat chloride concentrations, Strong *et al.* 1991), R117H (only male infertility, Bienvenue *et al.* 1993) and T338I (isolated hypotonic

dehydration, Leoni *et al.* 1995). Homozygosity for 3849+10 kb C⇨T has been associated with normal sweat chloride concentrations and pancreatic sufficiency (Highsmith *et al.* 1994, Gilbert *et al.* 1995) and a case of male fertility has been reported (Dreyfuß *et al.* 1996). Pancreatic sufficiency has been described for patients homozygous for G85E, both cases associated with mild lung disease and highly elevated sweat chloride concentrations (Chalkey and Harris 1991, Vasques *et al.* 1996). Pancreatic insufficiency and elevated sweat chloride concentrations have been shown for patients homozygous for 621+1 G⇨T – associated with normal pulmonary function – and 2183 AA⇨G – associated with moderate lung disease. All patients carrying two nonsense mutations have been reported as pancreatic insufficient (Cuppens *et al.* 1990, Cutting *et al.* 1990, Bal *et al.* 1991, Cheadle *et al.* 1992, Gasparini *et al.* 1992, Shoshani *et al.* 1992, Castaldo *et al.* 1996, Castaldo *et al.* 1997). Pancreatic insufficiency was described for patients homozygous for R1162X – associated with wild type mRNA levels (Rolfini *et al.* 1993) – and likewise for patients homozygous for R553X or G542X – associated with reduced mRNA levels (Hull *et al.* 1994b, Will *et al.* 1995). In all cases, elevated sweat chloride concentrations were associated with homozygosity for two nonsense mutations. Pulmonary disease varied from mildly affected to severely impaired among these patients.

Within this thesis, patients homozygous for R553X, CFTRdelEx2,3 , E92K, CFTRdelEx2, 3849+10kb C⇨T and 1898+3 A⇨G have been investigated during the study "*rare genotypes and atypical CF*"[31]. The clinical picture of these patients was consistent with similar cases described in the literature: the patient homozygous for R553X and two patients compound heterozygous for two stop mutations were all pancreatic insufficient and had an elevated sweat chloride concentration as reported for all other cases with two nonsense mutations. The moderate to severe pulmonary impairment among these three patients likewise reflects the results of variable lung impairment published by other authors. The E92K homozygous patient was comparable to the cases of G85E homozygosity: patients homozygous for these genotypes had a highly elevated sweat test, pancreatic sufficiency and little pulmonary disease in common (Chalkey and Harris 1991, Vasques *et al.* 1996). The patients homozygous for 3849+10 kb C⇨T were, as described for other cases with this mutation, presenting atypical normal sweat chloride levels and were pancreatic sufficient (Highsmith *et al.* 1994, Gilbert *et al.* 1995).

Three cases investigated carried genotypes in the homozygous condition not described in the literature: mutations 1898+3 A⇨G, CFTRdelEx2 and CFTRdelEx2,3 .The patient homozygous for 1898+3 A⇨G was pancreatic sufficient as were patients with 3849+10kbC⇨T or 2789+5G⇨A (Highsmith *et al.* 1994, Highsmith *et al.* 1997). The index case homozygous for 1898+3 A⇨G was investigated at 3 years of age and hence, prognosis for pulmonary disease and pancreatic status cannot be given at this point. CFTRdelEx2,3 was detected in a pancreatic insufficient patient with an impaired lung function and highly elevated sweat electrolytes. Based on the genotype, a condition as observed for patients carrying two nonsense mutations was expected for the out-of-frame deletion CFTRdelEx2,3 and the clinical findings are consistent with this hypothesis. CFTRdelEx2 was found

31 See chapter 3.1.2 for results and chapters 3.1.3.1, 3.1.3.2 and 3.1.2.3 for discussion.

in a pancreatic insufficient patient with a normal lung function. This in-frame deletion might theoretically give rise to functional CFTR. However, homozygotes for E92K and G85E, associated with CFTR proteins of an altered function, were described as pancreatic sufficient (Nunes *et al.* 1993, Chalkey and Harris 1991, Vasques *et al.* 1996). Thus, CFTRdelEx2 should be considered as a severe mutation.

4.1.3 Variability of CF lung disease– modulating factors or functional diversity of CFTR*?*

The manifestation of pulmonary disease is variable among (a) patients without CFTR, e.g. homozygotes for two nonsense mutations associated with mRNA reduction (Cuppens *et al.* 1990, Cutting *et al.* 1990, Bal *et al.* 1991, Cheadle *et al.* 1992, Shoshani *et al.* 1992, Castaldo *et al.* 1996, Castaldo *et al.* 1997) (b) patients with reduced wild type levels of CFTR, e.g. homozygotes for 3849+10 kb C⇨T (Highsmith *et al.* 1994, Gilbert *et al.* 1995, Dreyfuß *et al.* 1996, this work) and (c) patients carrying only one defective *CFTR* allele while suffering from pulmonary disease such as chronic bronchial hypersecretion, atypical sinopulmonary disease or disseminated bronchiectasis (Dumur *et al.* 1990a, Poller *et al.* 1991, Pignatti *et al.* 1995, Pignatti *et al.* 1996, Friedman *et al.* 1997). Thus, while *CFTR* mutations predispose to disease conditions in the lung, the severity of pulmonary disease in an individual cannot be predicted since for all groups of patients carrying the same *CFTR* mutation genotype both exceptionally mild and exceptionally severe cases of lung disease have been described. The pulmonary phenotypes displayed by the groups (a) "CFTR absent", (b) "CFTR reduced" or (c) "CF carrier status" overlap. In conclusion, the variability of lung disease is independent of the *CFTR* mediated risk. While this might indicate an involvement of factors distinct from CFTR modulating lung disease in general, a functional diversity of CFTR itself is an alternative to be considered.

Lung disease in CF was related to the CFTR transmitted defect of impaired chloride conductance by inferring an elevated NaCl concentration of the airway surface fluid, which in turn was thought to alter the defence against bacteria mediated by epithelial β-defensin-1 (Smith *et al.* 1996, Goldmann *et al.* 1997). However, the ionic composition of the airway surface fluid was also reported to be normal (Knowles *et al.* 1997) and no β-defensin-1, but neutrophil defensins were detected in bronchioalveolar lavage fluid (Schnapp and Harris 1998). Apart from chloride conductance, other roles of CFTR in lung disease have been suggested: Lallemand *et al* (1997) inferred a role of CFTR in the export of glutathione conjugated substrates, including leukotrienes which act as inflammatory agents. Lindsdell and Hanrahan (1998) demonstrated that CFTR mediates glutathione transport. Thus, apart from conducting chloride, CFTR seems to transport other substances as is known for other members of the ABC transporter family (reviewed by Higgins 1995).

In conclusion, CFTR has been shown to act as a chloride channel, to regulate the function of other ion channels (Egan *et al.* 1992, Gabriel *et al.* 1993, Ismailov *et al.* 1996, Loussouarn *et al.* 1996, Stutts *et al.* 1997, McNicholas *et al.* 1997) and might act as a transporter of glutathione and derivatives thereof. Thus, CFTR can be considered to perform roles typical for other ABC transporters.

4.1.4 Influence of intragenetic and extragenetic background on the basic defect

The lack of correlation between *CFTR* genotype and CF disease phenotype has been attributed to frequent polymorphisms within the *CFTR* gene that modulate the expression of the *CFTR* disease-causing lesion (Kiesewetter *et al.* 1993, Cuppens *et al.* 1998) or to rare polymorphisms directly compensating the disease-causing lesion (Dörk *et al.* 1991). Both mechanisms can be summarised as "intragenetic background". Modulation of the effect of a disease-causing *CFTR* sequence variation can only be expected if the sequence alteration occurs in combination with more than one intragenetic background, i.e. more than one haplotype has to be associated with a disease-causing lesion. In contrast, all *CFTR* alleles will occur in combinations with more than one allele at loci distinct from *CFTR* and thus, this "extragenetic background" might modulate the CF phenotype (as will be discussed in detail in chapter 4.2).

Within the study "*rare genotypes and atypical cystic fibrosis*", three ΔF508 homozygous patients with a second sequence alteration on one ΔF508 allele have been characterised with respect to the basic defect of impaired chloride conductance[32]. In these three cases, residual chloride channel activity was detected. However, residual chloride conductance was also measured in ΔF508 homozygous patients without a second sequence alteration on any ΔF508 allele[33]. Additionally, in two siblings carrying the genotype ΔF508/ΔF508-V1212I, the residual conductance was DIDS-sensitive indicating a non-CFTR origin. Hence, the second sequence alteration V1212I probably does not have a functional effect on the ΔF508 allele. In contrast, the polymorphism M470V was shown to influence the chloride conductance of mutant CFTR channels (Cuppens *et al.* 1998). Finally, the diversity of ion conductances displayed by ΔF508 homozygous patient pairs[34] – having an identical intragenetic background but not an identical extragenetic background – indicates that the expression of the basic defect among ΔF508 homozygous patients is not modulated by the *CFTR* gene itself.

Modulation of residual chloride conductance among CF patients can either be mediated by an influence on the expression and/or function of the mutant CFTR protein or by the presence of alternative channels. Within this thesis, three chloride conductances differing in their sensitivity to

32 See chapter 3.1.2 for results.
33 See chapter 3.2.2 for results.
34 See figure 25, shown in chapter 3.2.2 on page 95.

DIDS and cAMP have been described[35]. Thus, more than one alternative pathway for chloride conductance exists among CF patients. However, when analysing the relation of residual chloride channel activity in relation to the pancreatic status, any residual chloride conductance displayed by pancreatic insufficient patients was sensitive to DIDS (i.e. indicative of non-CFTR mediated chloride conductance) while the residual chloride conductance displayed by pancreatic sufficient patients was sensitive to cAMP (i.e. indicative of CFTR mediated chloride conductance)[36]. Thus, concordant with the literature published on a close correlation of *CFTR* genotype and pancreatic status (Corey *et al.* 1989, Kristidis *et al.* 1992), the pancreatic status appears to be determined by CFTR mediated residual chloride conductance. Recently it was demonstrated that the DIDS sensitive chloride channel is not expressed in pancreatic tissue (Gruber *et al.* 1998b) and consequently cannot compensate for defective CFTR in the pancreas. In agreement with these findings, an elevated frequency of mutant *CFTR* alleles was found among patients with chronic pancreatitis (Sharer *et al.* 1998), thus emphasising the importance of CFTR for the pancreatic status.

4.2 CF disease variability – the influence of non-*CFTR* genetics?

4.2.1 Genotype-phenotype correlation for ΔF508

ΔF508 was classified as a "severe" mutation associated with pancreatic insufficiency in most, but not all, cases (Santis *et al.* 1990a, Kristidis *et al.* 1992). However, a broad variability of lung disease severity was noted among ΔF508 homozygotes (Kerem *et al.* 1990, Johannsen *et al.* 1991). When ΔF508 homozygotes were compared to ΔF508 compound heterozygotes, pulmonary disease was reported to be similar for ΔF508 homozygous patients and patients carrying G551D/ΔF508 (Hamosh *et al.* 1992a) or ΔF508 compound heterozygous patients with various mutations (Al Jader *et al.* 1992). A higher risk of meconium ileus of ΔF508 homozygous patients compared to patients carrying G551D/ΔF508 was observed (Hamosh *et al.* 1992a). The disease severity of ΔF508 homozygotes and ΔF508 compound heterozygotes was reported to be identical in an ethnically diverse population (Lester *et al.* 1994). When Caucasian and Afro-American ΔF508 homozygous CF patients were compared, the patient groups were found identical except for lower height and weight percentiles of the Afro-American patients (Hamosh *et al.* 1998). As the Afro-American CF patients were diagnosed at a younger age than the Caucasian patients and the poorer nutritional status was already detected at the age of diagnosis, it was concluded by the authors that the dissimilarities observed "may extend beyond socioeconomic differences".

In summary, as described for patients homozygous for non-ΔF508 genotypes (chapter 4.1), variability of disease severity was noted among patients homozygous for the most frequent *CFTR* disease-causing lesion.

35 See figure 18, shown in chapter 3.1.2.5 on page 68.
36 See chapter 3.1.3.2.3.2 for detailed discussion.

4.2.2 The molecular phenotype of ΔF508 CFTR might be variable

Does the variability of CF disease severity among ΔF508 homozygotes reflect a variable molecular phenotype of ΔF508-CFTR? Common pathways, consisting of several cell compartments and multiple proteins, have been shown to contribute to CFTR processing (Jensen *et al.* 1995, Ward *et al.* 1995). Degradation of CFTR in intracellular compartments was shown for wild type CFTR in heterologous expression systems (Ward *et al.* 1994) as well as immortalised airway cells (Wei *et al.* 1996). However, the ΔF508-CFTR protein was characterised by a reduced stability within all systems investigated (Lukacs *et al.* 1993, Ward *et al.* 1994, Jensen *et al.* 1995). The equilibrium between processed and degraded CFTR was reported to be influenced by a variety of factors, ranging from general biophysical properties such as the temperature (Denning *et al.* 1992b) to chaperones interacting specifically with CFTR (Yang *et al.* 1993, Pind *et al.* 1994, Loo *et al.* 1998). In conclusion, the maturation of CFTR – wild type as well as ΔF508 protein – can be considered as a dynamic system and thus provides opportunities for variation in the expression of the maturation defect between different cell types or individual patients.

4.2.3 The basis of clinical concordance

In general, the variability of CF disease can be attributed to three different factors:

The *CFTR* mutation genotype, as the *CFTR* gene is characterised by allelic heterogeneity and correlation of *CFTR* genotype and some aspects of the disease phenotype – i.e. pancreatic status – are known;

Epigenetic factors, such as the socioeconomic status, therapeutic measures and the patient's compliance, but also the exposure to chemicals and/or bacterial and viral lung pathogens;

The genetic background, i.e. inherited factors other than the *CFTR* gene itself.

The analysis of the clinical data collected on patients participating in the "*European CF twin and sibling study*" has provided evidence of the influence of the genetic background on the disease severity [37]. CF twins and siblings have the same *CFTR* mutation genotype, they share several environmental factors and their genetic background. Surprisingly, the disease variability between two siblings of a pair (intrapair variance) was independent of the *CFTR* mutation genotype: ΔF508 homozygous patient pairs were comparable to all other patient pairs – displaying all other *CFTR* mutation genotypes – with respect to their intrapair variance in wfh%[38], representing the nutritional status, and FEVPerc[39], representing the pulmonary status. Likewise, the intrapair variance was identical for pairs composed of two pancreatic insufficient patients (associated with severe *CFTR* mutation genotypes) compared to pairs composed of two pancreatic sufficient patients (associated with mild *CFTR* mutation genotypes). In contrast, the influence of an identical extragenetic

37 For detailed results see chapter 3.2.1.
38 Defined in chapter 2.2.1.2 on page 47.
39 Defined in chapter 2.2.1.2 on page 47.

background had a significant influence on the intrapair variation of disease severity: monozygous twins were significantly more concordant than dizygous patient pairs, even though a group of monozygous twin pairs is heterogeneous with respect to the *CFTR* mutation genotype. In conclusion, the extragenetic background determined the concordance of CF patient pairs and thus the disease severity must be modulated by inherited factors other than the *CFTR* gene itself.

4.2.4 Non-CFTR genetics (I): analysis of ENaC as modifier of CF disease

Within this thesis, the influence of the epithelial amiloride sensitive Na^+ channel ENaC on the disease severity of selected ΔF508 homozygous patient pairs with extreme phenotypes was assessed using polymorphic markers near genes encoding for the three subunits of the ENaC. An association of an allele at D12S889, localised near SCNN1A encoding for the α subunit of the ENaC (Baens *et al.* 1995), with a mild disease phenotype was noticed[40]. Patients carrying one or two D12S889-10 alleles had a significantly better nutritional and pulmonary status than patients with other D12S889 genotypes[41]. Thus, disease severity among ΔF508 homozygotes is modulated by a gene linked to D12S889, presumably the α subunit of the ENaC itself.

The ENaC is related to CF on the molecular level and in terms of the disease phenotype. A direct modulation of the ENaC by CFTR has been demonstrated (Ismailov *et al.* 1996, Stutts *et al.* 1997). In addition, mutations within the genes encoding for ENaC subunits have been shown to cause pseudohypoaldosteronism type I (Chang *et al.* 1996, Strautnieks *et al.* 1996a) which in turn was reported to display symptoms known for CF disease (Hanakoglu *et al.* 1994, Marthinsen *et al.* 1998). Thus, as discussed detail in chapter 3.2.3.8.1, the ENaC most probably modulates the CF disease severity. Furthermore, the ENaC appears to be one of the factors determining the severity of lung disease among CF patients: the association of D12S889-10 alleles with a mild phenotype was more pronounced for the parameter FEVPerc than for wfh% and a patient suffering from pseudohypoaldosteronism type I was reported to be colonised with *Pseudomonas aeruginosa* (Marthinsen *et al.* 1998), a condition typical for CF lung disease.

4.2.5 Non-CFTR genetics (II): other modifiers

Within the last few years, modulation of CF disease by genes other than *CFTR* has been reported for the antiprotease α1-antitrypsin (Döring *et al.* 1994), immunglobulin G allotypes (Ciofu *et al.* 1997), the glutathione-S-transferase (Becker *et al.* 1997) and the low molecular weight protein 2 (Becker *et al.* 1997, Mekus *et al.* 1998). A modifier of CF disease severity in CF mice was mapped to mouse chromosome 7 (Rozmahel *et al.* 1996) and the homologous gene CFM1, reported to convey a risk for meconium ileus, was localised to human chromosome 19q13 (Zielenski *et al.* 1998).

40 See figure 30 shown in chapter 3.2.3.4 on page 103.
41 See figure 36 shown in chapter 3.2.3.7.1 on page 115.

Apart from these genetic approaches, an influence of residual chloride secretion on CF disease severity has been shown for intestinal tissue (Veeze *et al.* 1994). Among the genes encoding for non-CFTR chloride channels, CACLN1 – coding for the Ca^{2+} activated chloride channel – has recently been cloned. The presence of alternative chloride channels in patients investigated within the study "*Rare genotypes and atypical cystic fibrosis*" and the "*European CF twin and sibling study*" has demonstrated the existence of DIDS-sensitive residual chloride currents among CF patients[42]. Furthermore, ΔF508 homozygous patients have shown various combinations of DIDS-sensitive, alternative residual chloride secretion and DIDS-insensitive residual chloride secretion indicating a variable expression of the basic defect among patients with the same *CFTR* mutation genotype.

4.3 Advantage of CF carriers – none, *CFTR* transmitted or hitchhiking genes?

4.3.1 Allele frequency, allelic heterogeneity and heterozygote advantage

A high frequency of mutant alleles, conveying lethality in the homozygous state, has been attributed to a selective advantage of the heterozygous condition. It has been proposed that the carrier advantage is one of several reasons for the dominance of the diploid phase in higher plants and animals (Goldstein 1992). Examples for a heterozygote advantage have been suggested for thalassaemia traits and glucose-6-phosphate-deficiency, both resulting in an increased resistance to the malaria parasite due to the altered erythrocyte metabolism (reviewed by Friedman and Träger 1981). The high frequency of mutant *CFTR* alleles has been related to an advantage of the carrier state. However, the question remains unresolved as to whether an advantage is transmitted by the *CFTR* gene itself or is due to the action of other genes inherited in linkage to *CFTR*. Apart from the allele frequency, the diversity of mutant *CFTR* alleles is noteworthy. Clark (1998) has speculated on the mechanisms causing and leading to maintenance of allelic heterogeneity. The author concluded that systems in which complementation of alleles occurs favour a higher number of alleles. Strictly speaking, this mechanism is referring again to a group of linked genes and not to an isolated sequence.

4.3.2 A balance: carrier risk and carrier advantage

The carrier status of CF was recognised by a reduced β-adrenergically stimulated response of sweat glands (Behm *et al.* 1987). An elevated frequency of mutant *CFTR* alleles has been found in several disease conditions, predominantly but not exclusively for pulmonary diseases (Warner 1976, Dumur *et al.* 1990a, Dumur *et al.* 1990b, Poller *et al.* 1991, Chillon *et al.* 1995, Pignatti *et al.* 1995, Pignatti

42 See chapter 3.1.2 and 3.2.2 for results.

et al. 1996, Dahl *et al.* 1998b, Sharer *et al.* 1998; table 30). In parallel, hypotheses regarding the advantage of the heterozygous state have been made, mostly suggesting an enhanced resistance against viral and bacterial pathogens (Stuart and Burdon 1974, Quinton 1982, Hollander 1982, Cassano 1985, Rodman and Zamundio. 1991, Gabriel *et al.* 1994, Meindl 1987, Pier *et al.* 1998; table 30).

TABLE 30: RISK AND ADVANTAGE OF CF CARRIER STATUS

CARRIER RISK	REFERENCE
symptoms of allergy more frequent among CF carriers	Warner (1976)
elevated frequency of ΔF508 heterozygotes among adults with chronic bronchial hypersecretion	Dumur *et al.* (1990a)
elevated frequency of ΔF508 heterozygotes among azoospermic men with congenital aplasia of epididymis and vas deferens	Dumur *et al.* (1990b)
mutant *CFTR* alleles among patients with disseminated bronchiectatic lung disease	Poller *et al.* (1991)
elevated frequency of mutant *CFTR* alleles among CBAVD patients	Chillon *et al.* (1995)
elevated frequency of mutant *CFTR* alleles among patients with disseminated bronchiectasis	Pignatti *et al.* (1995), Pignatti *et al.* (1996)
CFTR is *Pseudomonas* receptor and thus non-functional CFTR enhances risk of infection	Pier *et al.* (1997)
elevated frequency of IVS8-5T alleles among patients with atypical sinopulmonary disease	Friedman *et al.* (1997)
reproductive disadvantage for smoking ΔF508 heterozygotes	Dahl *et al.* (1998a)
elevated frequency of ΔF508 heterozygotes among patients suffering from asthma	Dahl *et al.* (1998b)
elevated frequency of mutant *CFTR* alleles among patients with chronic pancreatitis	Sharer *et al.* (1998)
CARRIER ADVANTAGE	REFERENCE
resistance to typhus	Stuart and Burton (1974)
resistance to diarrhoea	Quinton (1982), Baxter *et al.* (1988)
resistance to venereal syphillis	Hollander (1982)
resistance to bubonic plague	Cassano (1985)
resistance to tuberculosis	Meindl *et al.* (1987)
resistance to cholera	Rodman and Zamundio (1991)
CF mice heterozygotes are resistant to cholera toxin	Gabriel *et al.* (1994)
family size related to carrier status	de Vries *et al.* (1997)
reproductive advantage for non-smoking ΔF508 heterozygotes	Dahl *et al.* (1998a)
resistance to *Salmonella typhi*	Pier *et al.* (1998)

In order to result in an elevated reproductive fitness, the carrier advantage has to outweigh the disadvantages resulting from the elevated vulnerability of CF carriers to other diseases. A reproductive advantage of CF carriership was detected by de Vries *et al* (1997). However, the authors concluded that the observed effect was too complex to be attributed to a heterozygote advantage alone. Dahl *et al* (1998) have reported an interrelationship of ΔF508 heterozygosity and smoking on the family size: while the family size of smoking ΔF508 heterozygotes was lower, an increase in family size among non-smoking ΔF508 heterozygotes was noticed. As smoking primarily affects the lung, it is tempting to speculate whether this finding reflects the balance of *CFTR* mediated risk – e.g. for lung disease – and *CFTR* mediated advantage – i.e. enhanced longevity due to defence against viral and bacterial pathogens.

4.3.3 Hitchhiking genes near CFTR

The discussion of risk and advantage of the CF carrier status is focussed on pulmonary disease or resistance to viral and bacterial pathogens in general. However, in terms of hitchhiking genes an influence on the nutritional status can be inferred as LEP, the human homologue to the murine obesity gene, is localised close to *CFTR* (Green *et al.* 1995). An influence of the haplotype at MetH and J3.11, both markers flanking the *CFTR* gene (see figure 31[43]), on growth parameters has been described (Tümmler *et al.* 1990). Furthermore, an interrelation of the pulmonary status and the J3.11 genotype was noticed (Santis *et al.* 1990b) and an association of postnatal female survival with the CS.7 genotype was detected (Macek *et al.* 1997).

The polymorphisms near the *CFTR* locus are in a pronounced linkage disequilibrium with the gene (Kerem *et al.* 1989a, Dörk *et al.* 1992, Cuppens *et al.* 1994), i.e. the "normal" *CFTR* allele is associated with other genotypes at the flanking polymorphisms compared to disease-causing *CFTR* alleles. The following two observations provide evidence for a selective advantage of the chromosomal background associated with most mutant *CFTR* alleles:

(a) The number of mutations associated with an intragenic haplotype was shown to be proportional to the frequency of the haplotype among non-CF chromosomes, indicating that mutations occur on each chromosomal background with the same probability (Morral *et al.* 1996). However, the most frequent *CFTR* mutations which have been caused by unrelated mutation events are found on the same intragenic dimorphic marker haplotype (Dörk *et al.* 1992, Cuppens *et al.* 1994, Morral *et al.* 1996). Consequently, mutations which must have occurred on other chromosomal backgrounds than the haplotype found on most CF chromosomes did not persist within the population.

43 Figure 31 is shown in chapter 3.2.3.5 on page 105.

(b) The most common mutation has been analysed at three intragenic microsatellite markers and several intra- and extragenic dimorphic markers (Dörk *et al.* 1992, Morral *et al.* 1996). Recombination between the two dimorphic markers KM.19 and TUB20, located in a 1 Mb distance to each other, is considered to be more frequent than a mutation event at a microsatellite locus (Morral *et al.* 1996). Thus, it is expected that each microsatellite haplotype is associated with several dimorphic marker haplotypes. For ΔF508 chromosomes, the opposite has been observed: ΔF508 is associated with several haplotypes at three intragenic dinucleotide polymorphism (Morral *et al.* 1994). In contrast, ΔF508 occurs nearly exclusively on one dimorphic marker haplotype spanning markers between KM.19 and TUB20. Consequently, other dimorphic haplotypes which must have been generated by recombination events did not persist within the population.

Hence, the chromosomal background associated with most *CFTR* mutations is suspected to transmit a selective advantage compared to the other haplotypes which must have existed on CF chromosomes but did not remain within the gene pool.

In this thesis, the hypothesis of hitchhiking genes within the *CFTR* linkage group was tested by analysing polymorphisms in a 40 cM region encompassing the *CFTR* gene. Allele frequencies were than compared among ΔF508 homozygous CF twins and siblings with extreme disease phenotypes. Patient pairs were classified as "*concordant/mild disease*" (CON+), "*concordant/severe disease*" (CON-) and "*discordant*" (DIS) based on their pulmonary function, their growth parameters and the intrapair variation thereof (for details see chapters 2.2.1.3 and 2.2.1.4 and figures 27, 28 and 29[44]). An association of an allele at a marker locus with the phenotype CON+, CON- or DIS was considered as indicative for a modulation of the disease phenotype by a gene linked to the marker locus studied. As all patient pairs investigated were ΔF508 homozygous and hence the *CFTR* disease-causing lesion was identical for all patients studied, the phenotype – CON+, CON- or DIS – cannot be caused by the *CFTR* gene itself.

Published results and results obtained within this work on flanking markers localised within 1 Mb distance to the *CFTR* gene (see figure 31[45]) are summarised in table 31. A benign phenotype – better growth (Tümmler *et al.* 1990), survival (Macek *et al.* 1997) and an overall mild disease (this work)– was found to be associated with alleles composing the flanking marker haplotype. Thus, the compiled data support the hypothesis of a selective advantage transmitted by the chromosomes carrying the CF haplotype.

44 Figure 27 is shown in chapter 3.2.3.1 on page 97, figure 28 is displayed in chapter 3.2.3.2 on page 99 and figure 29 is shown in chapter 3.2.3.2 on page 100.

45 Figure 31 is shown in chapter 3.2.3.5 on page 105.

TABLE 31: ASSOCIATION OF ALLELES AT *CFTR* FLANKING LOCI WITH PHENOTYPE

Allele frequencies are given for European populations (Estivill et al. 1987, Schmidke et al. 1987).

Published results: Tümmler et al. 1990: haplotype 1-1-2-1 associated with better growth, Santis et al. 1990a: allele J3.11-2 associated with mild lung disease, Macek et al. 1997: allele KM.19-2 associated with postnatal female survival.

ALLELE FREQUENCIES

		MetH	XV-2c	KM.19	J3.11
non-CF chromosomes	Allele 1	54	45	73	43
	Allele 2	45	55	27	57
CF- chromosomes	Allele 1	70	89	12	60
	Allele 2	30	11	88	40
"CF-haplotype"		**1**	**1**	**2**	**1**

ASSOCIATION OF ALLELE AND PHENOTYPE

	MetH	XV-2c	KM.19	J3.11
Tümmler (1990): better growth	1	1	2	1
Santis (1990a): mild lung disease				2
Macek (1997): female survival			2	
This work: associated with CON+	1	—	—	—
This work: associated with DIS	—	—	—	1

In addition to markers located within 1 Mb distance to *CFTR*, polymorphic loci within a 40 cM region encompassing the *CFTR* gene have been investigated within this thesis. As discussed in detail before[46], more than one modulating element had to be assumed to account for all effects observed. Figure 44 summarises the hypothesis, based on the known genes within the region investigated (see figure 43[47] for an overview of genes mapped near *CFTR*).

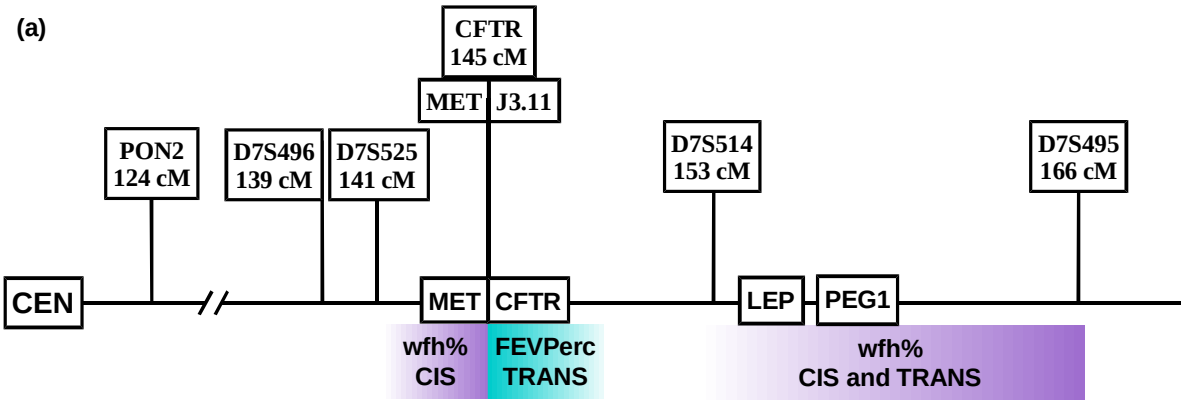
46 RESULTS: allelic association: For results on the association of alleles at marker loci near CFTR with the phenotypes CON+, CON- and DIS see chapter 3.2.3.5. Results on allelic association of D2S1788 alleles with the phenotype of the investigated patient pairs are shown in chapter 3.2.3.6.
RESULTS: correlation of genotype and clinical data: For loci near CFTR, results are displayed in chapter 3.2.3.7.2. For D2S1788, results are shown in chapter 3.2.3.7.3.
The effect are summarised in chapter 3.2.3.8.2.
DISCUSSION: Modulation of the pulmonary status is discussed in chapter 3.2.3.8.2.2. Influence of genes near CFTR and at D2S1788 on the nutritional status is dissected in chapter 3.2.3.8.2.3.

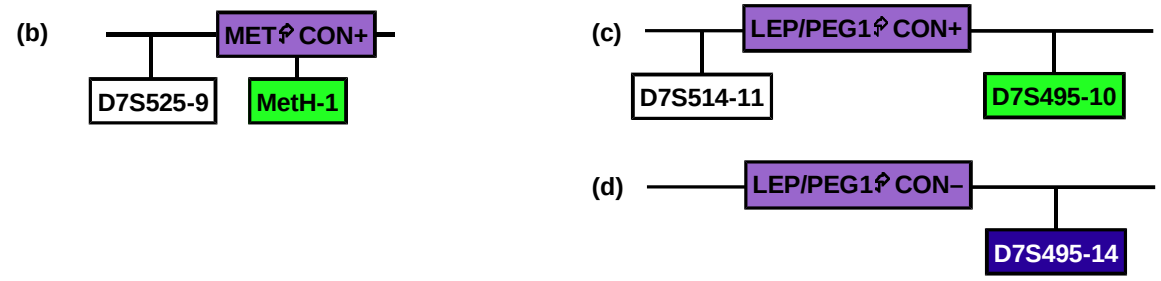
47 Figure 43 is shown in chapter 3.2.3.8.2.1 on page 129.

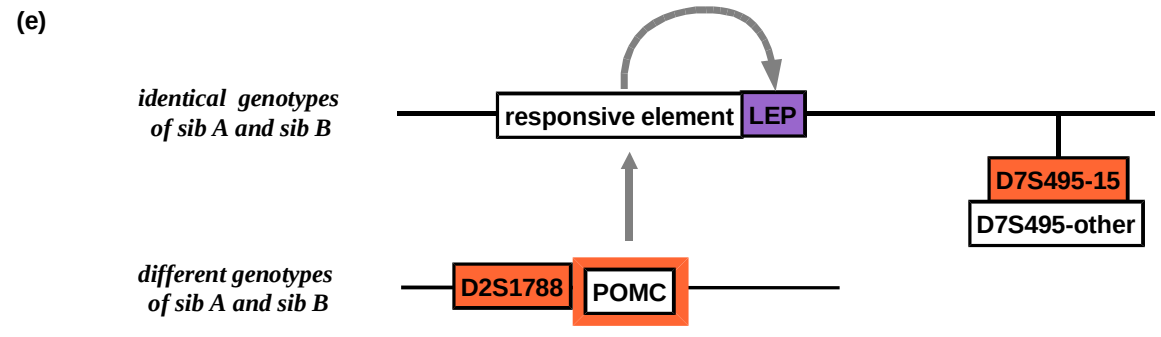
FIGURE 44: SUMMARY OF HYPOTHESES: HITCHHIKING GENES NEAR *CFTR*

(a) Map of investigated region, all typed markers, candidate genes on 7q31 and summary of observed effects.
(b) - (f) Hypotheses to explain observed allelic associations. For detailed discussion see text (page 148).
(b) - (d) Allelic association observed for concordant pairs (modulation in CIS).
(e) - (f) Allelic association observed for discordant pairs (modulation in TRANS).

Alleles at polymorphic loci: locus name - allele name, *observed associated with phenotype* CON+, CON – *or* DIS
Alleles at gene loci: gene name ↑ CON+, *this allele causes phenotype CON+*
gene name ↑ CON– , *this allele causes phenotype CON-*
gene name , *printed in orange box: allelic diversity at this locus causes phenotype DIS*







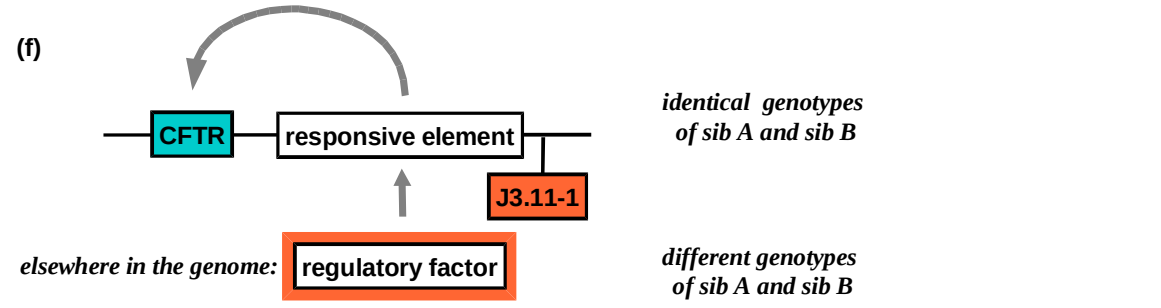


Figure 44 (b) MetH-1 was observed associated with the phenotype "*concordant/mild disease*" (figure 32). Patient pairs homozygous for MetH-1 had a significant better wfh% than patient pairs heterozygous at MetH (figure 37(b), $p < 0.01$). The most plausible mediator for this effect is the gene MET (discussed detailed in chapter 3.2.3.8.2.3).

Figure 44 (c) At D7S495, an association of the allele D7S495-10 with the phenotype "*concordant/mild disease*" was seen (figure 34). Pairs with siblings carrying each one or two alleles D7S495-10 had a significantly better wfh% than pairs with other D7S495 genotypes (figure 38(b), $p < 0.01$). This effect might be mediated by LEP, by PEG1 or both (discussed in detail in chapter 3.2.3.8.2.3).

Figure 44 (d) The allele D7S495-14 was found predominantly on pairs with a "*concordant/severe disease*" phenotype (figure 34) Pairs where each sibling carried one or more alleles D7S495-14 had a significantly worse wfh% than pairs with other D7S495 genotypes (figure 38(c), $p < 0.01$). Again, this effect might be mediated by LEP, by PEG1 or both (discussed in detail in chapter 3.2.3.8.2.3).

Figure 44 (e) Located further towards the telomer than *CFTR*, a modulation of the nutritional status by a factor acting in *trans* was inferred: alleles at J3.11, D7S514 and D7S495 were found associated with the phenotype "*discordant*" (figures 32, 33 and 34). As described before, an association of the phenotype "*discordant*" with an allele linked to *CFTR* indicates a modulation mediated by a gene encoded in *trans*, i.e. not within the *CFTR* linkage group: two siblings with CF have the identical *CFTR* gene and identical flanking sequences unless a recombination has occurred. Hence, discordance within such a sib pair has to be caused by a sequence located elsewhere in the genome. LEP, considered as a candidate for the modulation of the nutritional status, is localised between D7S514 and D7S495. Serum leptin levels were shown to be linked to D2S1788 (Commuzzie *et al.* 1997), and thus, an influence in *trans* on LEP can be assumed. D2S1788 was investigated in the selected panel of CF twins and siblings. Patient pairs sharing one or two alleles at D2S1788 were significantly more concordant in wfh% than pairs with no identical allele at D2S1788 (figure 42, $p < 0.01$). The most likely candidate gene near D2S1788 to mediate this effect is POMC (discussed in detail in chapter 3.2.3.8.2.3).

Figure 44 (f) At J3.11, an association of the phenotype "*discordant*" with the allele J3.11-1 was detected (figure 32). Pairs homozygous for J3.11-1 were significantly more concordant in FEVPerc than pairs heterozygous at J3.11 (figure 39(a), $p < 0.01$). As the *CFTR* genes and flanking sequences of two CF affected siblings are identical by descent, the discordance cannot be caused by a sequence linked to J3.11 itself: a gene encoded elsewhere, i.e. in *trans*, expressing allelic diversity, has to mediate the effect. Most likely, a regulatory factor acts on the CFTR expression via a responsive element liked to J3.11. The observed association of allele J3.11-1 with the phenotype "*discordant*" will occur if the allele of the responsive element linked to J3.11-1 is most susceptible to the allelic diversity of the regulatory factor (discussed in detail in chapter 3.2.3.8.2.2).

Figure 44 (c), (d) and (f): three pairs sharing no allele at D7S495 were ranked discordant. This condition was not seen on any pair with a concordant phenotype (table 22). Discordance in siblings sharing no allele at D7S495 could be caused by different alleles at the candidate gene loci LEP and/or PEG1 and/or the regulatory element controlling the expression of LEP, with a probable enhancing effect of an allelic diversity at POMC near D2S1788. Two out of the three pairs ranked discordant and sharing no allele at D7S495 likewise did not share an allele at D2S1788.

In conclusion, hitchhiking genes within the *CFTR* linkage group have a modulating effect on the CF disease phenotype. Whether the genes identified so far within the region investigated are responsible for the effect observed or whether other genes, not identified or mapped till today, are causing the modulation of the CF disease phenotype, will have to be clarified by further investigation of polymorphisms surmised to exist within the candidate genes.

4.4 Cystic fibrosis – monogenic disease or complex trait?

4.4.1 CF disease condition without two mutant CFTR alleles

Within this thesis, for three patients, presenting symptoms compatible with CF, evidence for the absence of *CFTR* disease-causing lesions on at least one allele was provided (see chapter 3.1.2.5 for results and chapter 3.1.3.2.2.4 for discussion). In these three cases – patient Q, R and S – only the lack of two *CFTR* disease-causing lesions has hindered he unequivocal diagnosis of CF. For patients Q, R and S, all *CFTR* exons and flanking intron sequences have been investigated for sequence alterations by SSCP and subsequent sequencing. The promoter region – extended to a 4 kb segment – has been screened for polymorphisms by SSCP in patients Q and R. Large genomic rearrangements have been searched for by pulsed field gel electrophoresis and southern blotting in patients Q and R. The only disease-causing lesion detected in patients Q, R and S was a splice site consensus transition preceding exon 12 in patient Q. The carrier status of patient Q was sustained by the familial history: a sibling was reported to have died from meconium ileus and thus, CF alleles within the family of patient Q are most likely associated with pancreatic insufficiency while patient Q is pancreatic sufficient. Patient S has a healthy sister, carrying the same alleles within a 30 cM region encompassing the *CFTR* gene locus (figure 19) and putative modifiers loci suspected within the linkage group (figure 44). Hence, unless at least one *de novo* mutation has occurred in patient S, he carries at least one wild type *CFTR* allele.

In conclusion, for all three patients the expression of the CF typical basic defect has been detected in at least one tissue. Due to the inheritance of *CFTR* alleles within the family, patient S cannot carry two *CFTR* disease-causing lesions. The carrier status of patient Q is sustained by he family history. For patient R, the absence of two disease-causing lesions can only be inferred based on the genetic analysis performed.

The basic defect of impaired chloride conductance has been investigated in the sweat gland, the nasal tissue and the rectal tissue of these patients. For patient R, evidence for dysfunctional CFTR was found in all three tissues investigated. Patient Q showed an impaired chloride conductance with the respiratory epithelium, but not within the sweat gland or the rectal tissue. In patient S, highly elevated sweat chloride concentrations were detected repetitively, but the nasal and rectal tissue displayed the non-CF condition. These results are summarised in table 32.

TABLE 32: CF WITHOUT TWO DISEASE-CAUSING LESIONS IN THE *CFTR* GENE

All patients present symptoms compatible with CF and have been screened for disease-causing lesions within all CFTR exons and flanking intron sites. For patient Q, 1898+3A⇒G has been identified on one chromosome. Patient S has a healthy sister carrying the same CFTR alleles and consequently cannot have two disease-causing lesions within the CFTR gene. The carrier status of patient Q is sustained by the familial history (see text for details).

	number of CFTR disease-causing lesions	pathological sweat test?	pathological nasal potential difference?	pathological ion currents in rectal tissue?
Patient Q	one	no	yes	no
Patient R	(none)	yes	yes	yes
Patient S	none	yes	no	no

Given the diagnosis of CF in patients Q, R and S – sustained by the pathological condition, due to the impaired chloride conductance typical for CF, which was detected in at least one tissue in each of these patients – other genes than *CFTR* alone must be considered in order to explain the disease condition in these patients. In patient R, a defective factor controlling the *CFTR* expression or the CFTR function might be responsible as an altered CFTR function was demonstrated in all tissues examined. For patient Q, the CF typical picture of defective chloride conductance was revealed exclusively in the nasal tissue and hence, elements of the genetic background involved in the tissue specific expression are likely to be involved. For patient S, pseudohypoaldosteronism type 1 – reported to be associated with CF-like symptoms and elevated sweat electrolytes (Hanokoglu *et al.* 1994, Marthinsen *et al.* 1998) – was excluded by differential diagnosis. It was demonstrated that pseudohypoaldosteronism type 1 can be caused by mutations within the genes encoding for the subunits of the epithelial sodium channel ENaC (Chang *et al.* 1996, Strautnieks *et al.* 1996a). In addition, an involvement of the SCNN1A gene in modulation of the CF phenotype was indicated (see chapter 3.2.3.8.1 for discussion). Thus, the ENaC subunits remain a plausible candidate gene to explain the disease symptoms in patient S as this index case and his healthy sister share only one allele at two polymorphic loci located near genes encoding for ENaC subunits and consequently, the discordance in this sib pair might be caused by different ENaC alleles.

Summarising these findings, the extensive analysis of the *CFTR* alleles from 350 patients (Dörk *et al.* 1994c, this work) has disclosed three index cases of CF-like disease without two defective *CFTR* alleles.

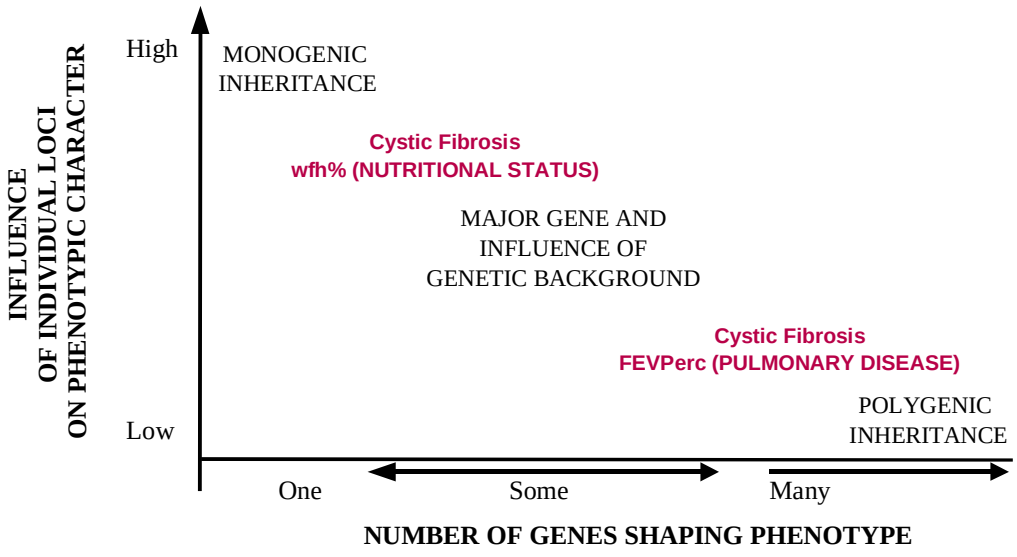
4.4.2 *On the role of the CFTR gene, hitchhiking sequences and the extragenetic background in nutritional status and pulmonary disease*

As discussed in chapter 4.3.3, modulating genes within the *CFTR* linkage group have been detected due to their influence on the nutritional status of CF twins and siblings. Thus, the selective advantage associated with CF carriership is most likely mediated by genes linked to *CFTR* which act on the nutritional status – and not by *CFTR* itself. In contrast, the risk associated with CF carriership appears to be directed to pulmonary disorders being most likely caused by the *CFTR* gene itself – not by flanking sequences. In conclusion, the balance between CF carrier risk and CF carrier advantage appears to be based on the genetic linkage between *CFTR* as a risk factor for lung disease and hitchhiking sequences providing an advantage with respect to the nutritional status.

The prominent role of the *CFTR* gene in determining the course of CF lung disease is not contradicted by the variability of pulmonary disease severity among patients carrying the same *CFTR* mutation genotype: statistical analysis of CF twins and siblings has provided evidence for multiple factors modulating CF lung disease while the nutritional status was shown to be determined by few inherited genes besides *CFTR* (see chapter 3.2.1.2 for details). As introduced in chapter 1.4, the number of loci involved in shaping a phenotypic trait and the impact of each gene determines whether the inheritance is perceived as monogenic or polygenic. Thus, the appearance of the trait "CF-nutritional status" resembles more the condition of a monogenic disease than "CF-pulmonary status" which has the appearance of a genetically complex trait.

FIGURE 45: CF: MONOGENIC DISEASE OR COMPLEX TRAIT ?

The number of loci involved and the impact of each locus on the phenotype determines whether a trait is perceived as monogenic, influenced by genetic background or polygenic (see chapter 1.4). For CF, wfh% (representing the nutritional status) was shown to be modified by few inherited factors besides the CFTR gene itself. The pulmonary status (expressed as FEVPerc) was demonstrated to be modulated by many inherited and environmental factors.



5 SUMMARY

Cystic fibrosis (CF) is the most frequent disease inherited in an autosomal recessive fashion within the Caucasian population. Though CF is considered to be a monogenic disorder transmitted by lesions in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, the disease severity is characterised by a high degree of variability. This thesis reports on two studies addressing the relative impact of the *CFTR* mutation genotype and other inherited factors on the course of CF disease.

Patients participating in the study "*rare genotypes and atypical CF*" were recruited from the Hannover clinic. A total of 19 patients were investigated by intestinal current measurements (ICM), nasal potential difference measurement (NPD) and a pilocarpine iontophoresis sweat test to assess the basic defect, being defined as an impaired CFTR transmitted Cl^- conductance. Mutation analysis of the CFTR gene was done by single strand conformation polymorphism analysis, direct sequencing and Southern blotting. The two deletions CFTRdelEx2 and CFTRdelEx2,3 were identified by analysis of genomic DNA. Different residual Cl^- conductances, distinguishable by their sensitivity to DIDS and cAMP, were seen by ICM in patients participating in the study "*rare genotypes and atypical CF*" but pancreatic sufficiency was found to be associated with a residual current indicative of CFTR mediated Cl^- conductance. The diagnosis of CF was confirmed by investigation of the basic defect for three patients presenting CF symptoms but lacking evidence for two *CFTR* disease causing lesions as judged by extensive mutation screening and, in two of these cases, sustained by the family history or genetic analysis within the family.

In the "*European CF twin and sibling study*", disease severity and intrapair discordance for more than 300 patient pairs was quantified by combining nutritional (weight for height %, wfh%) and pulmonary (FEV1%pred percentiles, FEVPerc) status. A significantly higher concordance of monozygous twins in wfh% and FEVPerc compared to dizygous patient pairs indicated the impact of non-*CFTR* genetics on the CF disease phenotype. 38 patient pairs, selected due to an extreme phenotype while being homozygous for the most frequent CF causing allele ΔF508, were analysed at candidate gene loci. The typing of polymorphic markers employing multiwell PCR reactions and continuous direct blotting electrophoresis was optimised. Significant association of an allele at the marker locus D12S889 with a milder clinical phenotype was detected. Hence, SCNN1A, encoding for the α-subunit of the amiloride-sensitive sodium channel (ENaC) and located near D12S889, might modulate CF disease. The analysis of markers within the *CFTR* linkage group pointed to the existence of hitchhiking genes which modulate the course of CF disease and thus might be responsible for a heterozygote advantage among CF carriers. A significant influence on wfh% was detected at MetH and at D7S495 due to an association of alleles at these loci with a milder phenotype. Candidate genes for the modulation of the nutritional status located within the investigated region are the receptor tyrosine kinase MET, the paternally imprinted gene PEG1 and the human homologue of the murine obesity gene LEP. An association with the phenotype "discordance" in FEVPerc, indicative of a modulation mediated by a gene encoded in trans, was seen at J3.11. This effect was attributed to a trans encoded factor acting on a responsive element linked to J3.11, thereby regulating *CFTR* expression. At the loci D7S514 and D7S495, an association of alleles with the phenotype "discordant" was seen in wfh%. As serum levels of leptin, encoded for by the gene LEP located near D7S514 and D7S495, were shown to be in linkage to D2S1788, this marker was analysed among CF twins and sibs. Sharing of alleles at D2S1788 was significantly associated with higher concordance in wfh%. Summarising these findings gained by genetic analysis of ΔF508 homozygous CF twins and siblings, evidence for a role of the ENaC and hitchhiking genes near *CFTR* in modulating CF disease severity has been provided.

6 ZUSAMMENFASSUNG

Cystische Fibrose (CF) wird durch Mutationen im "Cystic Fibrosis Transmembrane Conductance Regulator" (*CFTR*) Gen ausgelöst und ist die häufigste autosomal rezessiv vererbte Erkrankung in der kaukasischen Bevölkerung. CF wird als monogene Erkrankung angesehen, obwohl der Schweregrad der CF durch eine hohe Variabilität gekennzeichnet ist. Diese Dissertation beschreibt Ergebnisse zweier Studien, deren Ziel es ist, den relativen Einfluß des *CFTR* Mutationsgenotypes neben weiteren vererbten Faktoren auf den Verlauf der Erkrankung zu erfassen.

19 Patienten der CF Klinik in Hannover wurden im Rahmen der Studie "*Seltene Genotypen und atypische CF*" mithilfe der Messung der intestinalen Leitfähigkeit (ICM), der Nasalpotentialdifferenzmessung (NPD) und dem Pilocarpin-Iontophorese Schweißtest hinsichtlich ihres Basisdefektes untersucht, der sich als Störung der *CFTR* vermittelten Chloridleitfähigkeit manifestiert. Mutationsanalyse des *CFTR* Gens wurde mithilfe der Analyse von Einzelstrangkonformationspolymorphismen, Direktsequenzierung und Southern-Blot Verfahren durchgeführt. Die Deletionen CFTRDelEx2 und CFTRDelEx2,3 wurden auf genomischer Ebene nachgewiesen. Verschiedene Restleitfähigkeiten, die aufgrund ihrer Sensitivität gegenüber DIDS und cAMP unterschieden werden konnten, wurden mit ICM im Rahmen der Studie "*Seltene Genotypen und atypische CF*" beobachtet. Pankreassuffizienz war dabei mit CFTR vermittelter Restleitfähigkeit assoziiert. Durch die Analyse des Basisdefektes wurde die Diagnose der CF bei drei Patienten bestätigt, für die eine Anwesenheit von krankheitsauslösenden Läsionen auf beiden *CFTR* Allelen nach ausführlicher Mutationsanalyse unwahrscheinlich schien. Für zwei dieser Patienten wurde der genetische Befund durch klinische oder genetische Analyse der Familie unterstützt.

Im Rahmen der "*Europäischen CF Zwillings-und Geschwisterstudie*" wurde für mehr als 300 Patientenpaare der Schweregrad der Erkankung und die Intrapaardiskordanz quantitativ durch einen aus dem Ernährungsstatus (wfh%) und dem pulmonalen Status (FEVPerc) zusammengesetzten Parameter beschrieben. Monozygote Zwillinge waren signifikant konkordanter in wfh% und FEVPerc als dizygote Patientenpaare, so daß der Einfluß des genetischen Hintergrundes auf den Schweregrad der Erkrankung nachgewiesen werden konnte. Aufgrund eines extremen klinischen Phänotyps wurden 38 Patientenpaare, die für das häufigste CF verursachende Allel ΔF508 homozygot waren, zur Analyse von Kandidatengenorten ausgewählt. Die Typisierung polymorpher Marker wurde mithilfe von Multiwell-PCR Reaktionen und kontinuierlicher "Direct Blotting" Elektrophorese optimiert. Ein Allel des Marker Locus D12S889 war mit einem milden klinischen Verlauf signifikant assoziiert. SCNN1A, das Gen welches für die α–Untereinheit des Amilorid-sensitiven Natriumkanals codiert, liegt in der Nähe von D12S889 und könnte daher als Modulator der Erkrankung CF eine Rolle spielen. Eine Analyse von Markern innerhalb der *CFTR* Kopplungsgruppe ergab Hinweise auf die Existenz benachbarter Gene, die den Verlauf der CF beeinflussen und daher auch für den Heterozygotenvorteil von CF Genträgern verantwortlich sein können. Eine Assoziation von Allelen der Loci MetH und D7S495 mit einem milden Krankheitsbild konnte für wfh% beschrieben werden. Kandidatengene dieser Region mit einer möglichen Funktion bei der Modulation des Ernährungsstatus sind MET, PEG1 und LEP. Mit dem Phänotyp "diskordant" konnte eine Assoziation eines J3.11 Allels in FEVPerc gezeigt werden. Dieser Effekt stellt einen Hinweis auf den Einfluß eines in trans kodierten Gens dar und wurde auf einen regulatorischen Faktor zurückgeführt, der durch ein an J3.11 gekoppeltes Element die Expression des *CFTR* Gens beeinflußt. Eine Assoziation von Allelen der Loci D7S514 und D7S495 mit dem Phänotyp "diskordant" wurde in wfh% beobachtet. Da der Serumspiegel von Leptin – einem durch das zwischen D7S514 und D7S495 liegende Gen LEP kodierten Peptid – mit dem Marker D2S1788 in genetischer Kopplung steht, wurde D2S1788 bei den ausgewählten Patientenpaaren typisiert. Ein signifikanter Zusammenhang zwischen der Zahl gemeinsamer D2S1788 Allele und der Konkordanz in wfh% wurde beobachtet. Zusammenfassend läßt sich aus den Ergebnissen der Typisierung von ΔF508 homozygoten Zwillingen und Geschwistern ein modulierender Einfluß des ENaC sowie mehrerer Gene aus der *CFTR* Kopplungsgruppe auf den Schweregrad der Erkrankung ableiten.

7 BIBLIOGRAPHY

Adkins S, Gan KN, Mody M, La Du BN (1993) Molecular basis for the polymorphic forms of human serum paraoxonase / arylesterase: glutamine or arginine at position 191, for the respective A or B allozymes. Am J Hum Genet 52: 598-608

Al-Jader LN, Meredith AL, Ryley HC, Cheadle JP, Maguire S, Owen G, Goodchild MC, Harper PS (1992) Severity of chest disease in cystic fibrosis patients in relation to their genotypes. J Med Genet 29: 883-887

Allison DB, Heo M (1998) Meta-analysis of linkage data under worst-case conditions: a demonstration using the human *OB* region. Genetics 148: 859-866

Anderson MP, Sheppard DN, Berger HA, Welsh MJ (1992) Chloride channels in the apical membrane of normal and cystic fibrosis airway and intestinal epithelial cells. Am J Physiol 263: L1-L14

Antonorakis SE (1994) Genome linkage scanning: systematic or intelligent? Nat Genet 8: 211-212

Anwar R, Murray K, Hedge PJ, Smith JC, Markham AF (1990) DNA sequence analysis of the KM19 locus linked to cystic fibrosis. Design of new oligonucleotides to remove non-specific PCR products. Hum Genet 85: 319-323

Baens M, Aerssens J, van Zand K, van den Berghe H, Marynen P (1995) Isolation and regional assignment of human chromosome 12p cDNAs. Genomics 29: 44-52

Bajnath RB, Dekker K, Vaandrager AB, de Jonge HR, Groot JA (1992) Biphasic increase of apical Cl^- conductance by muscarinic stimulation of HT-29cl.19A human colon carcinoma cell line: evidence for activation of different Cl^- conductances by carbachol and forskolin. J Membrane Biol 127: 81-84

Bal J, Stuhrmann M, Schloesser M, Schmidtke J, Reiss J (1991) A cystic fibrosis patient homozygous for the nonsense mutation R553X. J Med Genet 28: 715-717

Barbry P, Simon-Bouy B, Mattei MG, Le Guern E, Jaume-Roig B, Chassande O, Ullrich A, Lazdunski M (1990) Localisation of the gene for amiloride-binding protein on chromosome 7 and RFLP analysis in cystic fibrosis families. Hum Genet 85: 587-589

Barreto C, Pinto LM, Duarte A, Lavinha J, Ramsay M (1991) A fertile male with cystic fibrosis: molecular genetic analysis. J Med Genet 1991: 420-421

Bateson W, Saunders ER (1902) Experimental studies in the physiology of heredity. Rep Evolut Comm Roy Soc. Rep I, pp 166

Baxter PS, Goldhill J, Hardcastle PT, Taylor CJ (1988) Accounting for cystic fibrosis. Nature 355: 211

Beck S, Pohl FM (1984) DNA sequencing with direct blotting electrophoresis. EMBO J 3: 2905-2909

Becker C, Mekus F, Epplen JT, Tümmler B (1997) Disease modifying loci in CF: study of candidate genes. EWGCF conference 1997, p.129

Behm JK, Hagiwara G, Lewiston NJ, Quinton PM, Wine JJ (1987) Hyposecretion of beta-adrenergically induced sweating in cystic fibrosis heterozygotes. Pediatr Res 22: 271-276

Bensch KW, Raida M, Mägert HJ, Schulz-Knappe P, Forssmann WG (1995) hBD-1: a novel β-defensin from human plasma. FEBS Lett 368: 331-335

Bergeron L, Perez GI, Macdonald G, Shi L, Sun Y, Jurisicova A, Varmuza S, Lathan KF, Flaws JA, Salter JC, Hara H, Moskowitz MA, Li E, Greenberg A, Tilly JL, Yuan J (1998) Defects in regulation of caspase-2-deficient mice. Genes Dev 12: 1304-1314

Bienvenue T, Beldjord C, Adjiman M, Kaplan JC (1993) Male infertility as the only presenting sign of cystic fibrosis when homozygous for the mild mutation R117H. J Med Genet 30: 797

Bienvenue T, Hubert D, Fonknechten N, Dusser D, Kaplan JC, Beldjord C (1994) Unexpected inactivation of acceptor consensus splice sequence by a –3 C to T transition in intron 2 of the CFTR gene. Hum Genet 94: 65-68

Bienvenue T, Lacronique V, Raymondjean M, Cazeneuve C, Hubert D, Kaplan JC, Beldjord C (1995) Three novel sequence variations in the 5' upstream region of the cystic fibrosis transmembrane conductance regulator (CFTR) gene: two polymorphisms and one putative molecular defect. Hum Genet 95: 698-702

Bottaro DP, Rubin JS, Faletto DL, Chan AM, Kmiecik TE, Vande Woude GF, Aaronson SA (1991) Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product. Science 251: 802-804

Boucher RC (1999) Status of gene therapy for cystic fibrosis lung disease. J Clin Invest 103: 441-445

Bouffard GG, Idol JR, Braden VV, Iyer LM, Cunningham AF, Weintraub LA, Touchman JW, Mohr-Tidwell RM, Peluso DC, Fulton RS, Ueltzen MS, Weissenbach J, Magness CL, Green ED (1997) A physical map of human chromosome 7: An integrated YAC contig with average STS spacing of 79 kb. Genome Res 7: 673-692

de Braekeleer M, Allard C, Leblanc JP, Simard F, Aubin G (1997) Genotype-phenotype correlation in five cystic fibrosis patients homozygous for the 621+1 G→T mutation. J Med Genet 34: 788-792

Bradbury NA, Jiling T, Berta G, Sorscher EJ, Bridges RJ, Kirk KL (1992) Regulation of plasma membrane recycling by CFTR. Science 256: 530-532

Bray MS, Boerwinkle E, Hanis CL (1996) *OB* gene not linked to human obesity in Mexican American affected sib pairs from Starr County, Texas. Hum Genet 98: 590-595

Brezillon S, Dupuit F, Hinnrasky J, Marchand V, Kälin N, Tümmler B, Puchelle E (1995) Decreased expression of the CFTR protein in remodeled human nasal epithelium from non-cystic fibrosis patients. Lab Invest 72: 191-200

Brezillon S, Hamm H, Heilmann M, Schäfers HJ, Hinnrasky J, Wagner TO, Puchelle E, Tümmler B (1997) Decreased expression of the cystic fibrosis transmembrane conductance regulator protein in remodeled airway epithelium from lung transplanted patients. Hum Pathol 28: 944-952

Bridges CB, Morgan TH (1923) The third chromosome group of mutant characters of *Drosohila melanogaster*. Publ Carneg Instn Nr 327

Bronsveld I, Bijman J, de Jonge HR, Sinaasappel M, Veeze HJ (1996) Gluconate response of nasal epithelium to discriminate between CF and non-CF in case of high baseline nasal potential difference. Ped Pulmonol Suppl 13: 125

Carter CO (1952) Familial incidence. In: Bodian, ed. Fibrocystic disease of the pancreas. London, Heyneman: 50-64

Casals T, Ramos MD, Giménez J, Larriba S, Nunes V, Estivill X (1997) High heterogeneity for cystic fibrosis in Spanish families: 75 mutations account for 90% of chromosomes. Hum Genet 101: 365-370

Cassano WF (1985) Cystic fibrosis and the plague. Med Hypothesis 18: 51-52

Castaldo G, Rippa E, Raia V, Salvatore D, Massa C, de Ritis G, Salvotore F (1996) Clinical features of cystic fibrosis patients with rare genotypes. J Med Genet 13: 73-76

Castaldo G, Rippa E, Salvatore D, Rafaella S, Raia V, de Ritis G, Salvatore F (1997) Severe liver impairment in a cystic fibrosis-affected child homozygous for the G542X mutation. Am J Med Genet 69: 155-158

Ceppellini R, Cantoni ES, Mattiuz PL, Miggiano V, Scandler G, Serra A (1967) Genetics of the leukocyte antigens. A family study of segregation and linkage. Histocompatibility testing, Munksgaard, Copenhagen, pp 149-189

Chalkey G, Harris A (1991) A cystic fibrosis patient who is homozygous for the G85E mutation has very mild disease. J Med Genet 28: 875-877

Champigny G, Imler JL, Puchelle E, Dalemans W, Gribkoff V, Hinnrasky J, Dott K, Barbry P, Pavirani A, Ladunski M (1995) A change in gating mode leading to increased intrinsic Cl^- channel activity compensates for defective processing in a cystic fibrosis mutant corresponding to a mild form of the disease. EMBO J 14: 2417-2423

Chang SS, Grunder S, Hanakoglu A, Rösler A, Mathew PM, Hanakoglu I, Schild L, Lu Y, Shimkets RA, Nelson-Williams C, Rossier BC, Lifton RP (1996) Mutations in the subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. Nat Genet 12: 248-253

de la Chapelle A, Ghamberg CG, Andersson LC, Ruutu T, Ruutu P, Kosunen TU, Repo H, Vuopio P (1979) A granulocyte membrane glycoprotein possibly involved in locomotion and coded for by chromosome 7. Cytogenet Cell Genet 25: 147

de la Chapelle A, Ruutu P, Ruutu T, Repo H, Vuopio P, Timonen T, Kosunen TU (1982) The locomotion defect of granulocytes in monosomy 7 involves random locomotion, chemotaxis and chemokinesis. Cytogenet Cell Genet 32: 264-265

Cheadle J, Al-Jader L, Goodchild M, Meredith AL (1992) Mild pulmonary disease in a cystic fibrosis child homozygous for R553X. J Med Genet 29: 597

Cheng SH, Gregory RJ, Marshall J, Paul S, Souza DW, White GA, O'Riordan CR, Smith AE (1990) Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. Cell 63: 827-834

Chevalier-Porst F, Bonardot AM, Gilly R, Chazalette JP, Mathieu M, Bozon D (1994) Mutation analysis in 600 French cystic fibrosis patients. J Med Genet 31: 541-544

Chillon M, Casals T, Mercier B, Bassas L, Lissens W, Silber S, Romey MC, Ruiz-Romero J, Verlingue C, Claustres M, Nunes V, Ferec C, Estivill X (1995) Mutations in the cystic fibrosis gene in patients with congenital absence of the vas deferens. N Engl J Med 332: 1475-1480

Chou JL, Rozmahel R, Tsui LC (1991) Characterization of the promoter region of the cystic fibrosis transmembrane conductance regulator gene. J Biol Chem 266: 24471-24477

Chu CS, Trapnell BC, Murtagh JJ Jr, Moss J, Dalemans W, Jallat S, Mercenier A, Pavirani A, Lecoq JP, Cutting GR (1991) Variable deletion of exon 9 coding sequences in cystic fibrosis transmembrane conductance regulator gene mRNA transcripts in normal bronchial epithelium. EMBO J 10: 1355-1363

Chu CS, Trapnell BC, Curristin S, Cutting GR, Crystal RG (1993) Genetic basis of variable exon 9 skipping in cystic fibrosis transmembrane conductance regulator mRNA. Nature Genet 3: 151-156

Ciofu O, Pressler T, Pandey JP, Hoiby N (1997) The influence of allotypes on the IgG subclass response to chromosomal β-lactamase of *Pseudomonas aeruginosa* in humans. Clin Exp Immunol 108: 88-94

Clark AG (1998) Mutation-selection balance with multiple alleles. Genetica 102/103: 41-47

Clement K, Garner C, HagerJ, Philipi A, LeDue C, Carey A, Haris TJR, Jury C, Cardon LR, Basdevant A, Demenais F, Guy-Grand B, North M, Froguel P (1996) Indication for linkage of the human OB gene region with extreme obesity. Diabetes 45: 687-690

Collin A, Sawyer SM, Mickle JE, Oates RD, Milunski A, Amos JA (1996) Pulmonary function and clinical observations in men with congenital bilateral absence of the vas deferens. Chest 110: 440-445

Comuzzie AG, Hixson JE, Almasy L, Mitchell BD, Mahaney MC, Dyer TD, Stern MP, MacCluer JW, Blangero J (1997) A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. Nat Genet 15: 273-276

Comuzzie AG, Allison DB (1998) The search for human obesity genes. Science 280: 1374-1377

Corey M, McLaughlin FJ, Williams M, Levison H (1988) A comparison of survival, growth, and pulmonary function in patients with cystic fibrosis in Boston and Toronto. J Clin Epidemiol 41: 583-591

Corey M, Durie P, Moore D, Forstner G, Levison H (1989) Familial concordance of pancreatic function in cystic fibrosis. J Pediatr 115: 274-277

Cotten JF, Ostedgaard LS, Carson MR, Welsh MJ (1996) Effect of cystic fibrosis-associated mutations in the fourth intracellular loop of cystic fibrosis transmembrane conductance regulator. J Biol Chem 271: 21279-21284

Coyle B, Reardon W, Herbrick JA, Tsui LC, Gausden E, Lee J, Coffrey R, Grueters A, Grossman A, Phelps PD, Luxon L, Kendall-Taylor P, Scherer SW, Trembath RC (1995) Molecular analysis of the PDS gene in Pendred syndrome (sensorineural hearing loss and goitre). Hum Mol Genet 7: 1105-1112

Cummings DE, Brandon EP, Planas JV, Motamed K, Idzerda RI, McKnight GS (1996) Genetically lean mice result from targeted disruption of the RII beta subunit of protein kinase A. Nature 382: 622-626

Cuppens H, Marynen P, DeBoeck C, DeBaets F, Eggermont E, van den Berghe H, Cassiman JJ (1990) A child, homozygous for a stop codon in exon 11, shows milder cystic fibrosis than her heterozygous nephew. J Med Genet 27: 717-719

Cuppens H, Teng H, Raeymakers P, DeBoeck C, Cassiman JJ (1994) CFTR haplotype backgrounds on normal and mutant CFTR genes. Hum Mol Genet 3: 607-614

Cuppens H, Lin W, Jaspers M, Costes B, Teng H, Vankeerberghen A, Jorissen M, Droogmans G, Reynaert I, Goossens M, Nilius B, Cassiman JJ (1998) Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (TG)m locus explains the partial penetrance of the T5 polymorphism as disease mutation. J Clin Invest 101: 487-496

Cutting GR, Kasch LM, Rosenstein BJ, Tsui LC, Kazazian HH, Antonorakis SE (1990) Two patients with cystic fibrosis, nonsense mutations in each cystic fibrosis gene, and mild pulmonary disease. N Engl J Med 323: 1685-1689

The Cystic Fibrosis Genetic Analysis Consortium. Cystic fibrosis mutation database http://www.genet.sickkids.on.ca./cftr/

Dahl M, Tybjærg-Hansen A, Wittrup HH, Lange P, Nordestgaard BG (1998a) Cystic fibrosis ΔF508 heterozygotes, smoking, and reproduction: studies of 9141 individuals from a general population sample. Genomics 50: 89-96

Dahl M, Tybjærg-Hansen A, Lange P, Nordestgaard BG (1998b) ΔF508 heterozygosity in cystic fibrosis and susceptibility to asthma. Lancet 351: 1911-1913

Daly AK, Salh BS, Bilton D, Allen J, Knight AD, Webb AK, Braganzy JM, Idle JR (1992) Deficient nifedipine oxidation: a rare inherited trait associated with cystic fibrosis kindreds. Pharmacogenetics 2: 19-24

Davis PB, Drumm M, Konstan MW (1996) Cystic fibrosis. Am J Respir Crit Care Med 154: 1229-1256

Dean M, Park M, Le Beau MM, Robins TS, Diaz MO, Rowley JD, Blair DG, Vande Woude GF (1985) The human met oncogene is related to the tyrosine kinase oncogenes. Nature 318: 385-388

Denning GM, Ostedgaard LS, Cheng SH, Smith AE, Welsh MJ (1992a) Localization of cystic fibrosis transmembrane conductance regulator in chloride secretory epithelia. J Clin Invest 89: 339-349

Denning GM, Anderson MP, Amara JF, Marshall J, Smith AE, Welsh MJ (1992b) Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. Nature 358: 731-764

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human chromosome based on 5264 microsatellites. Nature 380: 152-154

Döring G, Krogh-Johansen H, Weidinger S, Hoiby N (1994) Allotypes of alpha-1-antitrypsin in patients with cystic fibrosis, homozygous and heterozygous for ΔF508. Pediatr Pulmonol 18: 3-7

Dörk T, Wulbrand U, Richter T, Neumann T, Wolfes H, Wulf B, Maaß G, Tümmler B (1991) Cystic fibrosis with three mutations in the cystic fibrosis transmembrane conductance regulator gene. Hum Genet 87: 441-446

Dörk T, Neumann T, Wulbrand U, Wulf B, Kälin N, Maaß G, Krawcak M, Guillermit H, Ferec C, Horn G, Klinger K, Kerem BS, Zielenski J, Tsui LC, Tümmler B (1992) Intra- and extragenic marker haplotypes of CFTR mutations in cystic fibrosis families. Hum Genet 88: 417-425

Dörk T, Wulbrand U, Tümmler B (1993) Four novel cystic fibrosis mutations in splice junction sequences affecting the CFTR nucleotide binding folds. Genomics 15: 688-691

Dörk T, Fislage R, Neumann T, Wulf B, Tümmler B (1994a) Exon 9 of the CFTR gene: splice site haplotypes and cystic fibrosis mutations. Hum Genet 93: 67-73

Dörk T, Mekus F, Schmidt K, Boßhammer J, Fislage R, Heuer T, Dziadek V, Neumann T, Kälin N, Wulbrand U, Wulf B, von der Hardt H, Maaß G, Tümmler B (1994b) Detection of more than 50 different mutations in a large group of German cystic fibrosis patients. Hum Genet 94: 533-542

Dörk (1994c) Molekulargenetik der Mukoviszidose: Identifizierung und Charakterisierung von CFTR Mutationen. Dissertation Universität Hannover.

Dreyfus DH, Bethel R, Gelfand EW (1996) Cystic fibrosis 3894 + 10 kb C→T mutation associated with severe pulmonary disease and male fertility. Am J Respir Crit Care Med 153: 858-860

Duggirala R, Stern MP, Mitchell BD, Reinhart LJ, Shipman PA, Uresandi OC, Chung WK, Leibel RL, Hales CN, O'Connel P, Blangero J (1996) Quantitative variation in obesity-related traits and insulin precursors linked to the *OB* gene region on human chromosome 7. Am J Hum Genet 59: 694-703

Dumur V, Lafitte JJ, Gervais R, Debaecker D, Kesteloot M, Lalau G, Roussel P (1990a) Abnormal distribution of cystic fibrosis ΔF508 allele in adults with chronic bronchial hypersecretion. Lancet 335: 1340

Dumur V, Gervais R, Rigot JM, Lafitte JJ, Manouvrier S, Biserte J, Mazeman E, Roussel P (1990b) Abnormal distribution of cystic fibrosis ΔF508 allele in azoospermic men with congenital aplasia of epidymis and vas deferens. Lancet 336: 512

Dupuit F, Kälin N, Brezillon S, Hinnrasky J, Tümmler B, Puchelle E (1995) CFTR and differentiation markers expression in non-CF and ΔF508 homozygous CF nasal epithelium. J Clin Invest 96: 1601-1611

Egan M, Flotte T, Afione S, Solow R, Zeitlin PL, Carter BJ, Guggino WB (1992) Defective regulation of outwardly rectifying Cl^- channels by protein kinase A corrected by insertion of CFTR. Nature 358: 581-584

Eggerding FA, Schonberg SA, Chebab FF, Norton ME, Cox VA, Epstein CJ (1994) Uniparental isodisomy for paternal 7p and maternal 7q in a child with growth retardation. Am J Hum Genet 55: 253-265

Eiberg H, Mohr J, Schmiegelow K, Nielsen LS, Williamson R (1985) Linkage relationships of paraoxonase (PON) with other markers: indication of PON-cystic fibrosis synteny. Clin Genet 28: 265-271

Engelhardt JF, Yankaskas JR, Ernst SA, Yang Y, Marino CR, Boucher RC, Cohn JA, Wilson JM (1992) Submucosal glands are the predominant site of CFTR expression in the human bronchus. Nat Genet 2: 240-248

Estivill X, Scambler PJ, Wainwright BJ, Kawley K, Frederick P, Schwartz M, Baiget M, Kere J, Williamson R, Farrall M (1987) Patterns of polymorphism and linkage disequilibrium for cystic fibrosis. Genomics 1: 257-263

European Working Group on Cystic Fibrosis Genetics (1990): Gradient of distribution in Europe of the major CF mutation and of its associated haplotype. Hum Genet 85: 436-445

Fanen P, Ghanem N, Vidaud M, Besmond C, Martin J, Costes B, Plassa F, Goossens M (1992) Molecular characterization of cystic fibrosis: 16 novel mutations identified by analysis of the whole cystic fibrosis conductance transmembrane regulator (CFTR) coding regions and splice site junctions. Genomics 13: 770-776

Fanen P, Labarthe R, Garnier F, Benharough M, Goossens M, Edelman A (1997) Cystic fibrosis phenotype associated with pancreatic insufficiency does not always reflect the cAMP-dependent chloride conductive pathway defect. J Biol Chem 272: 30563-30566

Fischer H, Kreusel KM, Illek B, Machen TE, Hegel U, Clauss W (1992) The outwardly rectifying Cl^- channel is not involved in cAMP-mediated Cl^- secretion in HT 29 cells: evidence for a very low conductance Cl^- channel. Pflügers Arch 422: 159-167

Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME (1998) Localisation of a gene implicated in a severe speech and language disorder. Nat Genet 18: 168-170

Fox MF, Lahbib F, Pratt W, Attwood J, Gum J, Kim Y, Swallow DM (1992) Regional localisation of the intestinal mucin gene MUC3 to chromosome 7q22. Ann Hum Genet 56: 281-287

Frankel WN, Schork NJ (1996) Who's afraid of epistasis? Nat Genet 14: 371-373

Freije JP, Fueyo A, Uria JA, Velasco G, Sanchez LM, Lopez-Boado YS, Lopez-Otin C (1992) Human Zn-alpha-2-glycoprotein: complete genomic sequence, identification of a related pseudogene and relationship to class I major histocompatibility complex genes. Genomics 18: 575-587

French PJ, van Doorninck JH, Peters RH, Verbeek E, Ameen NA, Marino CR, de Jonge HR, Bijman J, Scholte BJ (1996) A ΔF508 mutation in mouse cystic fibrosis transmembrane conductance regulator results in a temperature-sensitive processing defect in vivo. J Clin Invest 98: 1304-1312

Friedman KJ, Heim RA, Knowles MR, Silverman LM (1997) Rapid characterization of the variable length polythymidine tract in the cystic fibrosis (CFTR) gene: association of the 5T allele with selected CFTR mutations and its incidence in atypical sinopulmonary disease. Hum Mutat 10: 108-115

Friedman MJ, Trager W (1981) The biochemistry of resistance to malaria. Sci Am 244: 158-164

Fulmer SB, Schwiebert EM, Morales MM, Guggino WB, Cutting GR (1995) Two cystic fibrosis transmembrane conductance regulator mutations have different effects on both pulmonary phenotype and regulation of outwardly rectified chloride currents. Proc Natl Acad Sci USA 92: 6832-6836

Gabriel SE, Clarke LL, Boucher RC, Stutts MJ (1993) CFTR and outwardly rectifying chloride channels are distinct proteins with a regulatory relationship. Nature 363: 263-268

Gabriel SE, Brigman KN, Koller BH, Boucher RC, Stutts MJ (1994) Cystic fibrosis heterozygote resistance to cholera toxin in the cystic fibrosis mouse model. Science 266: 107-109

Galliard D, Ruocco S, Lallemand A, Dalemans W, Hinnrasky J, Puchelle E (1994) Immunohistochemical localization of cystic fibrosis transmembrane conductance regulator in human fetal airway and digestive mucosa. Pediatr Res 36: 137-143

Garrod AE, Hurley WH (1912) Congenital familial steatorrhea. Q J Med 6: 242-258

Gasparini P, Novelli G, Estivill X, Olivieri D, Savoia A, Ruzzo A, Nunes V, Borgo G, Antonelli M, Williamson R, Pignatti PF, Dallapiccola B (1990) The genotype of a new linked DNA marker, MP6d-9, is related to the clinical course of cystic fibrosis. J Med Genet 27: 17-20

Gasparini P, Borgo G, Mastella G, Bonizzato A, Dognini M, Pignatti PF (1992) Nine cystic fibrosis patients homozygous for the CFTR nonsense mutation R1162X have mild or moderate lung disease. J Med Genet 29: 558-562

Gibson LE, Cooke RE (1959) A test for concentration of electrolytes in cystic fibrosis using pilocarpine by electrophoresis. Pediatrics 23: 549-563

Gilbert F, Li, Z, Arzimanoglou I, Bialer M, Denning C, Gorvoy J, Honorof J, Ores C, Quittel L, Rossoff L, Valverde K (1995) Clinical spectrum in homozygotes and compound heterozygotes inheriting cystic fibrosis mutation 3894 +10kb C→T: significance for geneticists. Am J Med Genet 58: 356-359

Gögelein H (1988) Chloride channels in epithelia. Biochem Biophys Acta 947: 521-547

Goldman MJ, Anderson GM, Stolzenberg ED, Kari UP, Zasloff M, Wilson JM (1997) Human β-defensin-1 is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis. Cell 88: 553-560

Goldstein DB (1992) Heterozygote advantage and the evolution of the dominant diploid phase. Genetics 132: 1195-1198

Gopalakrishnan L, Scarpulla RC (1995) Structure, expression and chromosomal asignment of the human gene encoding nuclear respiratory factor 1. J Biol Chem 270: 18019-18025

Green ED, Maffei M, Braden VV, Procenca R, DeSilva U, Zhang Y, Chua SC Jr, Leibel RL, Weissenbach J, Friedman JM (1995) The human obese (OB) gene: RNA expression pattern and mapping on the physical, cytogenetic, and genetic maps of chromosome 7. Genome Res 5: 5-12

Griffais R, Andre PM, Thibon M (1991) K-tuple frequency in the human genome and polymerase chain reaction. Nucleic Acids Res 19: 3887-3891

Gross-Bellard M, Dudet P, Chambon P (1973) Isolation of high molecular weight DNA from mammalian cells. Eur J Biochem 36: 32-38

Gruber AD, Gandhi R, Pauli BU (1998a) The murine calcium-sensitive chloride channel (mCaCC) is widely expressed in secretory epithelia and in other select tissues. Histochem Cell Biol 110: 43-49

Gruber AD, Elble RC, Ji HL, Schreur KD, Fuller CM, Pauli BU (1998b) Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first member of the family of Ca^{2+} activated Cl^- channel proteins. Genomics 54: 200-214

Guidelines for Human Gene Nomenclature (1997). available from:
URL http://www.gene.ucl.ac.uk/nomenclature/guidelines.html

Gum JR, Hicks JW, Swallow DM, Lagace RL, Byrd JC, Lamport DT, Siddiki B, Kim YS (1990) Molecular cloning of cDNAs derived from a novel human intestinal mucin gene. Biochem Biophys Res Commun 171: 407-415

Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. J Genet 8: pp 299

Hall RA, Ostedgaard LS, Premont RT, Blitzer JT, Rahman N, Welsh MJ, Lefkowitz RJ (1998) A C-terminal motif found in the β_2-adrenergic receptor, P2Y1 receptor and cystic fibrosis transmembrane conductance regulator determines binding to the Na^+/H^+ exchanger regulatory factor family of PDZ proteins. Proc Natl Sci USA 95: 8496-8501

Hamosh A, Trapnell BC, Zeitlin PL, Montrose-Rafizadeh C, Rosenstein BJ, Crystal RG, Cutting GR (1991) Severe deficiency of cystic fibrosis transmembrane conductance regulator mRNA carrying nonsense mutations R553X and W1316X in respiratory epithelial cells of patients with cystic fibrosis. J Clin Invest 88: 1880-1885

Hamosh A, King TM, Rosenstein BJ, Corey M, Levison H, Durie P, Tsui LC, McIntosh I, Keston M, Brock DJH, Macek M Jr, Zemkova D, Krasnocanova H, Vavrova V, Macek M Sr, Golder N, Schwarz MJ, Super M, Watson EK, Williams C, Bush A, O'Mahony SM, Humphries P, DeArce MA, Reis A, Bürger J, Stuhrmann M, Schmidtke J, Wulbrand U, Dörk T, Tümmler B, Cutting GR (1992a) Cystic fibrosis patients bearing both the common missense mutation Gly→Asp at codon 551 and the ΔF508 mutation are clinically indistinguishable from ΔF508 homozygotes, except for a decreased risk of meconium ileus. Am J Hum Genet 51: 245-250

Hamosh A, Rosenstein BJ, Cutting GR (1992b) CFTR nonsense mutations G542X and W1282X asociated with severe reduction of CFTR mRNA in nasal epithelial cells. Hum Mol Genet 1: 542-544

Hamosh A, FitzSimmons SC, Macek M jr, Knowles MR, Rosenstein BJ, Cutting GR (1998) Comparison of the clinical manifestations of cystic fibrosis in black and white patients. J Pediatr 132: 255-259

Hanakoglu A, Bistritzer T, Rakover Y, Mandelberg A (1994) Pseudohypoaldosteronism with increased sweat and saliva electrolyte values and frequent lower respiratory tract infections mimicking cystic fibrosis. J Pediatr 125: 752-755

Hansson GC (1988) Cystic fibrosis and chloride secreting diarrhoea. Nature 333: 711

Hansson JH, Nelson-Williams C, Suzuki H, Schild L, Shimkets R, Lu Y, Canessa C, Iwasaki T, Rossier B, Lifton RP (1995) Hypertension caused by a truncated epithelial sodium channel γ subunit: genetic heterogeneity of Liddle syndrome. Nat Genet 11: 76-82

Hasegawa H, Skach W, Baker O, Calayag MC, Lingappa V, Verkman AS (1992) A multifunctional aqueous channel formed by CFTR. Science 258: 1477-1479

Haws CM, Nepomuceno IB, Krouse ME, Wakelee H, Law T, Xia Y, Nguyen H, Wine JJ (1996) ΔF508-CFTR channels: kinetics, activation by forskolin, and potentiation by xanthines. Am J Physiol 270: C1544-C1555

Higgins CF (1995) The ABC of channel regulation. Cell 82: 693-696

Highsmith WE, Burch LH, Zhou Z, Olsen JC, Boat TE, Spock A, Gorvoy JD, Quittel L, Friedman KJ, Silverman LM, Boucher RC, Knowles MR (1994) A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations. New Engl J Med 331: 974-980

Highsmith WE, Burch LH, Zhou Z, Olsen JC, Strong TV, Smith T, Friedman KJ, Silverman LM, Boucher RC, Collins FS, Knowles MR (1997) Identification of a splice site mutation (2789+5G>A) associated with small amounts of normal CFTR mRNA and mild cystic fibrosis. Hum Mutat 9: 332-338

Höglund P, Holmberg C, de la Chapelle A, Kere J (1994) Paternal isodisomy for chromosome 7 is compatible with normal growth and development in a patient with congenital chloride diarrhoea. Am J Hum Genet 55: 747-752

Höglund P, Sistonen P, Norio R, Holmberg C, Dimberg A, Gustavson KH, de la Chapelle A, Kere J (1995) Fine mapping of the congenital chloride diarrhoea gene by linkage disequilibrium. Am J Hum Genet 57: 95-102

Höglund P, Haila S, Scherer SW, Tsui LC, Green ED, Weissenbach J, Holmberg C, de la Chapelle A, Kere J (1996a) Positional candidate genes for congenital chloride diarrhoea suggested by high-resolution physical mapping in chromosome region 7q31. Genome Res 6: 202-210

Höglund P, Haila S, Socha J, Tomaszewski L, Saarialho-Kere U, Karjalainen-Lindsberg ML, Airola K, Holmberg C, de la Chapelle A, Kere J (1996b) Mutations in the down-regulated in adenoma (DRA) gene cause congenital chloride diarrhoea. Nat Genet 14: 316-319

Hollander DH (1982) Etiogenesis of the European cystic fibrosis polymorphism: heterozygote advantage against veneral syphilis? Med Hypotheses 8: 191-197

Horn GT, Richards B, Wright G, Landes GM, Klinger KW (1989) Analysis and detction of genetic polymorphisms near the cystic fibrosis and Huntington's disease loci. Cytogenet Cell Genet 51: 1015

Hubert D, Bienvenue T, Desmazes-Dufeu N, Fajac I, Lacronique J, Matran R, Kaplan JC, Dusser DJ (1996) Genotype-phenotype relationships in a cohort of adult cystic fibrosis patients. Eur Respir J 9: 2207-2214

Hull J, Shackleton S, Harris A (1993) Abnormal mRNA splicing resulting from three different mutations in the CFTR gene. Hum Mol Genet 2: 689-692

Hull J, Shackleton S, Harris A (1994a) Analysis of mutations and alternative splicing patterns in the CFTR gene using mRNA derived from nasal epithelial cells. Hum Mol Genet 3: 1141-1146

Hull J, Shackleton S, Harris A (1994b) The stop mutation R553X in the CFTR gene results in exon skipping. Genomics 19: 362-364

Illek B, Yankaskas JR, Machen TE (1997) cAMP and genistein stimulate HCO_3^- conductance through CFTR in human airway epithelia. Am J Physiol 272: L752-L761

Inglehearn CF (1997) Intelligent linkage analysis using gene density estimates. Nat Genet 16: 15

Inoue K, Inazawa J, Nakagawa H, Shimada T, Yamazaki H, Guengerich FP, Abe T (1992) Assignment of the human cytochrome P-450 nifedipine oxidase gene (CYP3A4) to chromosome 7 at band 7q22.1 by fluorescence in situ hybridization. Jpn J Hum Genet 37: 133-138

Ismailov II, Awayda MS, Jovov B, Berdiev BK, Fuller CM, Dedman JR, Kaetzel M, Benos DJ (1996) Regulation of epithelial sodium channels by the cystic fibrosis transmembrane conductance regulator. J Biol Chem 271: 4725-4732

Ismailov II, Berdiev BK, Shlyonsky VG, Fuller CM, Prat AG, Jovov B, Cantiello HF, Ausiello DA, Benos DJ (1997) Role of actin in regulation of epithelial sodium channels by CFTR. Am J Physiol 272: C1077-C1086

Jensen TJ, Loo MA, Pind S, Williams DB, Goldberg AL, Riordan JR (1995) Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. Cell 83: 129-135

Jeunemaitre X, Basilana F, Persu A, Dumont C, Champigny G, Lazdunski M, Corvol P, Barby P (1997) Genotype-phenotype analysis of a newly discovered family with Liddle's syndrome. J Hypertens 15: 1091-1100

Johannsen HK, Nir M, Hoiby N, Koch C, Schwartz M (1991) Severity of cystic fibrosis in patients homozygous and heterozygous for ΔF508 mutation. Lancet 337: 631-634

Johannsen W (1909) Elemente der exakten Erblichkeitslehre. Fischer Verlag, Jena.

Jorde LB, Lathrop GM (1988) A test of the heterozygote-advantage hypothesis in cystic fibrosis carriers. Am J Hum Genet 42: 808-815

Kartner N, Augustina O, Jensen TJ, Naismith AL, Riordan JR (1992) Mislocalization of ΔF508 CFTR in cystic fibrosis sweat gland. Nat Genet 1: 321-327

Kerem BS, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC (1989a) Identification of the cystic fibrosis gene: Genetic analysis. Science 245: 1073-1080

Kerem BS, Buchanan JA, Durie P, Corey M, Levison H, Rommens JM, Buchwald M, Tsui LC (1989b) DNA marker haplotype association with pancreatic sufficiency in cystic fibrosis. Am J Hum Genet 44: 827-834

Kerem BS, Zielenski J, Markiewicz D, Bozon D, Gazit E, Yahav J, Kennedy D, Riordan JR, Collins FS, Rommens JM, Tsui LC (1990) Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. Proc Natl Acad Sci USA 87: 8447-8451

Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D, Levison H, Tsui LC, Durie P (1990) The relationship between genotype and phenotype in cystic fibrosis - analysis of the most common mutation (ΔF508). N Engl J Med 323: 1517-1522

Kiesewetter S, Macek M Jr, Davis C, Curristin SM, Chu CS, Graham C, Shrimpton AE, Cashman SM, Tsui LC, Mickle J, Amos J, Highsmith WE, Shuber A, Witt DR, Crystal RG, Cutting GR (1993) A mutation in CFTR produces different phenotypes depending on chromosomal background. Nature Genet 5: 274-277

Kitzis A, Chomel JC, Kaplan JC, Giraud G, Labbe A, Dastugue B, Dumur V, Farriaux JP, Roussel P, Williamson R, Feingold J (1988) Unusual segregation of cystic fibrosis alleles to males. Nature 333: 215

Kimura M (1956) A model of a genetic system which leads to closer linkage by natural selection. Evolution 10: 278

Knowles MR, Clarke LL, Boucher RC (1981) Increased bioelectrical potential difference across respiratory epithelia in cystic fibrosis. N Engl J Med 305: 1489-1495

Knowles MR, Robinson JM, Wood RE, Pue CA, Mentz WM, Wagner GC, Gatzy JT, Boucher RC (1997) Ion composition of airway surface liquid of patients with cystic fibrosis as compared with normal and disease-control subjects. J Clin Invest 100: 2588-2595

Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B (1983) Changes in the normal maximal expiratory flow volume curve with growth and ageing. Am Rev Respir Dis 127: 725-734

Kobayashi S, Kohda T, Miyoshi N, Kuroiwa Y, Aisaka K, Tsutsumi O, Kaneko-Ishino T, Ishino F (1997) Human *PEG*1/*MEST*, an imprinted gene on chromosome 7. Hum Mol Genet 6: 781-786

Koh J, Sferra TJ, Collins FS (1993) Characterization of the cystic fibrosis transmembrane conductance regulator promoter region. Chromatin context and tissue-specificity. J Biol Chem 268: 15912-15921

Kosambi (1944) The estimation of map-distance from recombination values. Ann Eugen 12: 172-175

Kotzot D, Schmitt S, Bernasconi F, Robinson WP, Lurie IW, Ilyina H, Méhes K, Hamel BCJ, Otten BJ, Hergersberg M, Werder E, Schoenie E, Schinzel A (1995) Uniparental disomy 7 in Silver-Russel syndrome and primordial growth retardation. Hum Mol Genet 4: 583-587

Kristidis P, Bozon D, Corey M, Markiewicz D, Rommens J, Tsui LC, Durie P (1992) Genetic determinants of exocrine pancreatic function in cystic fibrosis. Am J Hum Genet 50: 1178-1184

Krude H, Bieberman H, Luck W, Horn R, Brabant G, Grüters A (1998) Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by *POMC* mutations in humans. Nat Genet 19: 155-157

Kubesch P, Dörk T, Wulbrand U, Kälin N, Neumann T, Wulf B, Geerlings H, Weißbrodt H, von der Hardt H, Tümmler B (1993) Genetic determinants of airways' colonization with *Pseudomonas aeruginosa* in cystic fibrosis. Lancet 341: 189-193

van Kuijck MA, van Aubel RA, Busch AE, Lang F, Russel FG, Bindels RJ, van Os CH, Deen PM (1996) Molecular cloning and expression of a cyclic AMP activated chloride conductance regulator: a novel ATP-binding cassette transporter. Proc Natl Acad Sci USA 93: 5401-5406

Lallemand JY, Stoven V, Annereau JP, Boucher J, Blanquet S, Barthe J, Lenoir G (1997) Induction by antitumoral drugs of proteins that functionally complement CFTR: a novel therapy for cystic fibrosis? Lancet 350: 711-712

Lander ES, Botstein D (1986) Consanguinity and heterogeneity: cystic fibrosis need not be homogeneous in Italy. Am J Hum Genet 39: 282-283

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265: 2037-2048

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11: 241-247

Leoni GB, Ptzalis S, Podda R, Zanda M, Silvetti M, Caocci L, Cao A, Rosatelli C (1995) A specific cystic fibrosis mutation (T338I) associated with the phenotype of isolated hypotonic dehydration. J Pediatr 127: 281-283

Lester LA, Kraut J, Lloyd-Still J, Karrison T, Mott C, Billstrand C, Lemke A, Ober C (1994) ΔF508 genotype does not predict disease severity in an ethnically diverse cystic fibrosis population. Pediatrics 93: 114-118

Li XC, Everett LA, Lalwani AK, Desmukh D, Friedman TB, Green ED, Wilcox ER (1998) A mutation in PDS causes non-syndromic recessive deafness. Nat Genet 18: 215-217

Lindsdell P, Hanrahan JW (1996) Disulphonic stilbene block of cystic fibrosis transmembrane conductance regulator Cl- channels expressed in a mammalian cell line and its regulation by a critical pore residue. J Physiol 496: 687-693

Lindsdell P, Hanrahan JW (1998) Glutathione permeability of CFTR. Am J Physiol 275: C323-326

Loo MA, Jensen TJ, Cui L, Hou Y, Chang XB, Riordan JR (1998) Perturbation of Hsp90 interaction with nascent CFTR prevents its maturation and accelerates its degradation by the proteasome. EMBO J 17: 6879-6887

Loussourarn G, Demolombe S, Mohammad-Panah R, Escande D, Baro I (1996) Expression of CFTR controls cAMP-dependent activation of epithelial K^+ currents. Am J Physiol 271: C1565-1573

Lowe CU, May CD, Reed SC (1949) Fibrosis of the pancreas in infants and children: a statistical study of clinical and hereditary features. Am J Dis Child 78: 349-374

Ludwig M, Bolkenius U, Wickert L, Marynen P, Bidlingmaier F (1998) Structural organisation of the gene encoding the α-subunit of the human amiloride-sensitive epithelial sodium channel. Hum Genet 102: 576-581

Lukacs GL, Chang XB, Bear C, Kartner N, Mohamed A, Riordan JR, Grinstein S (1993) The ΔF508 mutation decreases the stability of cystic fibrosis transmembrane conductance regulator in the plasma membrane. Determination of functional half-lives on transfected cells. J Biol Chem 268: 21592-21598

Macek M Jr, Macek M Sr, Krebsova A, Nash E, Hamosh A, Reis A, Varon-Mateeva R, Schmidke J, Maestri NE, Sperling K, Krawcak M, Cutting GR (1997) Possible association of the allele status of the CS.7/*Hha*I polymorphism 5' of the *CFTR* gene with postnatal female survival. Hum Genet 99: 565-572

MacVinish LJ, Pickles RJ, Cuthbert AW (1993) Cyclic AMP and Ca^{2+} interactions affecting epithelial chloride secretion in human cultured colonic epithelia. Br J Pharmacol 108: 462-468

Maiuri L, Raia V, DeMarco G, Coletta S, deRitis G, Londei M, Auricchio S (1997) DNA fragmentation is a feature of cystic fibrosis epithelial cells: a disease with inappropriate apoptosis? FEBS Lett 408: 225-231

Marthinsen L, Kornfält R, Aili M, Andersson D, Westgren U, Schaedel C (1998) Recurrent *Pseudomonas* bronchopneumonia and other symptoms as in cystic fibrosis in a child with type 1 pseudohypoaldosteronism. Acta Paediatrica 87: 472-474

Mayer B, Hemmens B (1997) Biosynthesis and action of nitric oxide in mammalian cells. Trends Biochem Sci 22: 477-481

McNicholas CM, Nason MW jr, Guggino WB, Schwiebert EM, Hebert SC, Giebisch G, Egan ME (1997) A functional CFTR-NBF1 is required for ROMK2-CFTR interaction. Am J Physiol 273: F843-848

Meindl RS (1987) Hypothesis: a selective advantage for cystic fibrosis heterozygotes. Am J Phys Anthropol 74: 39-45

Mekus F (1994) Makrorestriktionsanalyse und Mikrosatelliten im *CFTR*-Gen. Diplomarbeit Universität Hannover

Mekus F, Ballmann M, Becker C, Bijman J, Bronsveld I, Halley D, Laabs U, Otters H, Sandkuyl L, Veeze H, Tümmler B and the European CF twin and sibling study consortium (1998) Genetic modifiers of CF disease. European CF conference PS7-2

Melo CA, Serra C, Stoyanova V, Aguzzoli C, Faraguna D, Tamanini A, Berton G, Carbrini G, Baralle FE (1993) Alternative splicing of a previously unidentified CFTR exon introduces an in-frame stop codon 5' of the R region. FEBS Lett 329: 159-162

Meng QH, Springall DR, Bishop AE, Morgan K, Evans TJ, Habib S, Gruenert DC, Gyi KM, Hodson ME, Yacoub MH, Polak JM (1998) Lack of inducible nitric oxide synthase in bronchial epithelium: a possible mechanism of susceptibility to infection in cystic fibrosis. J Pathol 1998: 323-331

Miller PW, Hamosh A, Macek M Jr, Greenberger PA, MacLean J, Walden SM, Slavin RG, Cutting GR (1996) Cystic fibrosis transmembrane conductance regulator (CFTR) gene mutations in allergic bronchopulmonary aspergillosis. Am J Hum Genet 59: 45-51

Montague CT, Farooqi S, Whitehead JP, Soos MA, Rau H, Wareham NJ, Sewter CP, Digby JE, Mohammed SN, Hurst JA, Cheetham CH, Earley AR, Barnett AH, Prins JB, O'Rahilly S (1997) Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature 387: 903-908

Morgan TH, Cattrell E (1912) Data for the study of sexlinked inheritance in drosophila. J Exp Zool 13: pp79

Morgan TH, Sturtevant AH, Muller HJ, Bridges CB (1915) The mechanisms of Mendelian heredity. H Holt & Company, New York

Morison S, Dodge JA, Cole TJ, Lewis PA, Coles EC, Geddes D, Russel G, Littlewood JM, Scott MT (1997) Height and weight in cystic fibrosis: a cross sectional study. UK cystic fibrosis survey management committee. Arch Dis Child 77: 497-500

Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Giménez J, Reis A, Varon-Mateeva R, Macek M jr, Kaladjieva L, Anglicheva D, Dancheva R, Romeo G, Russo MP, Garnerone S, Restagno G, Ferrari M, Magnani C, Claustres M, Desgorges M, Schwartz M, Schwarz M, Dallapiccola B, Novelli G, Ferec C, de Arce M, Nemeti M, Kere J, Anvret M, Dahl N, Kadasi L (1994) The origin of the major cystic fibrosis mutation (ΔF508) in European populations. Nature Genet 7: 169-175

Morral N, Dörk T, Llevadot R, Dziadek V, Mercier B, Ferec C, Costes B, Girodon E, Zielenski J, Tsui LC, Tümmler B, Estivill X (1996) Haplotype analysis of 94 cystic fibrosis mutations with seven polymorphic CFTR DNA markers. Hum Mutat 8: 149-159

Myal Y, Gregory C, Wang H, Hammerton JL, Shiu RPC (1989) The gene for proclactin-inducible protein (PIP), uniquely expressed in exocrine organs, maps to chromosome 7. Somat Cell Molec Genet 15: 265-270

Naren AP, Quick MW, Collawen JF, Nelson DJ, Kirk KL (1998) Syntaxin 1A inhibits CFTR chloride channels by means of domain-specific protein-protein interactions. Proc Natl Acad Sci USA 95: 10972-10977

National Committee for Clinical Laboratory Standards (1994) Sweat testing: sample collection and quantitative analysis - approved guideline (Document C34-A). Wayne, Pennsylvania

Norman RA, Leibel RL, Chung WK, Power-Kehoe L, Chua SC Jr, Knowler WC, Thompson DB, Bogardus C, Ravussin E (1996) Absence of linkage of obesity and energy metabolism to markers flanking homologues of rodent obesity genes in Pima indians. Diabetes 45: 1229-1232

Novotny WF, Chassande O, Baker M, Lazdunski M, Barbry P (1994) Diamine oxidase is the amiloride-binding protein and is inhibited by amiloride analogues. J Biol Chem 269: 9921-9925

Nunes V, Chillon M, Dörk T, Tümmler B, Casal T, Estivill X (1993) A new missense mutation (E92K) in the first transmembrane domain of the CFTR gene causes a benign cystic fibrosis phenotype. Hum Mol Genet 2: 79-80

Oksanen L, Öhman M, Heiman M, Kainulainen K, Kaprio J, Mustajoki P, Koivisto V, Koskenvuo M, Jänne OA, Peltonen L, Kontula K (1997) Markers for the gene *ob* and serum leptin levels in human morbid obesity. Hum Genet 99: 559-564

Pasyk EA, Foskett JK (1997) Cystic fibrosis transmembrane conductance regulator-associated adenosine-3'phosphate-5'-phosphosulfate channels in endoplasmatic reticulum and plasma membranes. J Biol Chem 272: 7746-7751

Pier GB, Grout M, Zaidi TS (1997) Cystic fibrosis transmembrane conductance regulator is an epithelial cell receptor for clearance of *Pseudomonas aeruginosa* from the lung. Proc Natl Acad Sci USA 94: 12088-12093

Pier GB, Grout M, Zaidi T, Meluleni G, Mueschenborn SS, Banting G, Ratcliff R, Evans MJ, Colledge WH (1998) *Salmonella typhi* uses CFTR to enter intestinal epithelial cells. Nature 393: 79-82

Pignatti PF, Bombieri C, Marigo C, Benetazzo M, Luisetti M (1995) Increased incidence of cystic fibrosis gene mutations in adults with disseminated bronchiectasis. Hum Mol Genet 4: 635-639

Pignatti PF, Bombieri C, Benetazzo M, Casartelli A, Trabetti E, Gilè LS, Martinati LC, Boner AL, Luisetti M (1996) CFTR gene variant IVS8-5T in disseminated bronchiectasis. Am J Hum Genet 58: 889-892

Pilz AJ, Povey S, Gruss P, Abbott CM (1993) Mapping of the human homologues of the murine paired-box homeotic genes. Mamm Genome 4: 78-82

Pind S, Riordan JR, Williams DB (1994) Participation of the endoplasmatic reticulum chaperone calnexin (p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator. J Biol Chem 269: 12784-12788

Poller W, Faber JP, Scholz S, Olek K, Müller KM (1991) Sequence analysis of the cystic fibrosis gene in patients with disseminated bronchiectatic lung disease. Klin Wochenschr 69: 657-663

Prader A, Largo RH, Molinari L, Issler C (1989) Physical growth of Swiss children from birth to 20 years of age. First Zurich longitudinal study of growth and development. Helv Paediatr Acta Suppl 52: 1-125

Pritchard DJ (1987) Why cystic fibrosis is on the increase. Nature 330: 319

Quinton PM (1982) Abnormalities in electrolyte secretion in cystic fibrosis sweat glands due to decreased anion permeability. In: Quinton PM, Martinez RJ, Hopfer U (eds) Fluid and electrolyte abnormalities in exocrine glands in cystic fibrosis. San Francisco Press, San Francisco, Calif, pp 53-76

Ravnik-Glavcak M, Glavcak D, Dean M (1994): Sensitivity of single-strand conformation polymorphism and heteroduplex method for mutation detection in the cystic fibrosis gene. Hum Mol Genet 3: 801-807

Reed DR, Ding Y, Xu W, Cather C, Green ED, Price RA (1996) Extreme obesity may be linked to markers flanking the human *OB* gene. Diabetes 45: 691-694

Riordan JR, Rommens JM, Kerem BS, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, Drumm ML, Iannuzzi MC, Collins FS, Tsui LC (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245: 1066-1071

Risch N, Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 268: 1584-1589

Rodman DM, Zamudio S (1991) The cystic fibrosis heterozygote-advantage in surviving cholera? Med Hypotheses 36: 253-258

Rolfini R, Cabrini G (1993) Nonsense mutation R1162X of the cystic fibrosis transmembrane conductance regulator gene does not reduce messenger RNA expression in nasal epithelial tissue. J Clin Invest 92: 2683-2687

Romeo G, Devoto M, Galietta LJV (1989) Why is the cystic fibrosis gene so frequent ? Hum Genet 94: 1-5

Romeo G, McKusick VA (1994) Phenotypic diversity, allelic series and modifier genes. Nat Genet 7: 451-453

Rommens JM, Iannuzzi MC, Kerem BS, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui LC, Collins FS (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245: 1059-1065

Rosenbloom CL, Kerem BS, Rommens JM, Tsui LC, Wainwright B, Williamson R, O'Brian WE, Beaudet AL (1989) DNA amplification for detection of the XV-2c polymorphism linked to cystic fibrosis. Nucleic Acids Res 17: 7117

Rotoli BM, Bussolati O, Sironi M, Cabrini G, Gazzola GC (1994) CFTR protein is involved in the efflux of neutral amino acids. Biochem Biophys Res Comm 204: 653-658

Rozmahel R, Wilschanski M, Matin A, Plyte S, Oliver M, Auerbach W, Moore A, Forstner J, Durie P, Nadeau J, Bear C, Tsui LC (1996) Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor. Nat Genet 12: 280-287

Di Sant'Agnese PA, Darling RC, Perera GA, Shea E (1953) Abnormal electrolyte composition of sweat in cystic fibrosis of the pancreas, its clinical significance and relationship to the disease. Pediatrics 12: 549-563

Santis G, Osborne L, Knight RA, Hodson ME (1990a) Independent genetic determinants of pancreatic and pulmonary status in cystic fibrosis. Lancet 336: 1081-1084

Santis G, Osborne L, Knight RA, Hodson ME (1990b) Linked marker haplotypes and the ΔF508 mutation in adults with mild pulmonary disease and cystic fibrosis. Lancet 335: 1426-1429

Scambler PJ, Farral M, Stanier P, Bell G, Ramirez F, Wainwright BJ, Bell J, Lench NJ, Kruyer H, Williamson R (1985) Linkage of COLA2 collagene gene to cystic fibrosis, and its clinical implications. Lancet 2: 1241-1242

Schild L, Canessa CM, Shimkets RA, Warnock DO, Lifton RP, Rossier BC (1995) A mutation in the epithelial sodium channel causing Liddle's disease increases channel activity in the *Xenopus laevis* oocyte expression system. Proc Natl Acad Sci USA 92: 5699-5703

Schmidtke J, Krawzcak M, Schwartz M, Alkan M, Bonduelle M, Bühler E, Chemke M, Darnedde T, Domagk J, Wéngel W, Frey D, Fryburg K, Halley D, Hundrieser J, Ladanyi L, Libaers I, Lissens W, Mächler M, Malik NJ, Morreau J, Neubauer V, Oostra B, Pape B, Poncin JE, Schinzel A, Simon P, Trefz FK, Tümmler B, Vassart G, Voss R (1987) Linkage relationships and allelic associations of the cystic fibrosis locus and four marker loci. Hum Genet 76: 337-343

Schnapp D, Harris A (1998): Antibacterial peptides in bronchoalveolar lavage fluid. Am J Respir Cell Mol Biol 19: 352-356

Schultheiss G, Diener M (1997) Regulation of apical and basolateral K+ conductances in rat colon. Br J Pharmacol 122: 87-94

Schulz IJ (1969) Micropuncture studies of the sweat formation in cystic fibrosis patients. J Clin Invest 48: 1470-1477

Seibert FS, Lindsdell P, Loo TW, Hanrahan JW, Clarke DM, Riordan JR (1996a) Disease-associated mutations in the fourth cytoplasmic loop of cystic fibrosis transmembrane conductance regulator comprise biosynthetic processing and chloride channel activity. J Biol Chem 271: 15139-15145

Seibert FS, Lindsdell P, Loo TW, Hanrahan JW, Riordan JR, Clarke DM (1996b) Cytoplasmatic loop three of the cystic fibrosis transmembrane conductance regulator contributes to regulation of chloride channel activity. J Biol Chem 271: 27493-27499

Seibert FS, Jia Y, Mathews CJ, Hanrahan JW, Riordan JR, Loo TW, Clarke DM (1997) Disease-associated mutations in cytoplasmic loops 1 and 2 of cystic fibrosis transmembrane conductance regulator impede processing or opening of the channel. Biochemistry 36: 11966-11974

Sharer N, Schwarz M, Malone G, Howarth A, Painter J, Super M, Braganza J (1998) Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. N Engl J Med 339: 645-652

Sheppard DN, Ostedgaard LS, Winter MC, Welsh MJ (1995) Mechanism of dysfunction of two nucleotide binding domain mutations in cystic fibrosis transmembrane conductance regulator that are associated with pancreatic sufficiency. EMBO J 14: 876-883

Shier WT (1979) Increased resistance to influenza as a possible source of heterozygote advantage in cystic fibrosis. Med Hypothesis 5: 661-667

Shimkets RA, Warnock DG, Bositis CM, Nelson-Williams C, Hansson JH, Schambelan M, Gill JR Jr, Ulick S, Milora RV, Findling JW, Canessa CM, Rossier BC, Lifton RP (1994) Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. Cell 79: 407-414

Short DB, Trotter KW, Reczek D, Kreda SM, Bretscher A, Boucher RC, Stutts MJ, Milgram SL (1998) An apical PDZ protein anchors the cystic fibrosis transmembrane conductance regulator to the cytoskeleton. J Biol Chem 273: 19797-19801

Shoshani T, Augarten A, Gazit E, Bashan N, Yahav Y, Rivlin Y, Tal A, Seret H, Yaar L, Kerem E, Kerem BS (1992) Association of a nonsense mutation (W1282X), the most common mutation in the Ashkenazi Jewish cystic fibrosis patients in Israel, with presentation of severe disease. Am J Hum Genet 50: 222-228

Shoshani T, Kerem E, Szeinberg A, Augarten A, Yahav Y, Cohen D, Rivlin J, Tal A, Kerem BS (1994) Similar levels of mRNA from the W1282X and the ΔF508 cystic fibrosis alleles, in nasal epithelial cells. J Clin Invest 93: 1502-1507

Slomski R, Schloesser M, Berg LP, Wagner M, Kakkar VV, Cooper DN, Reiss J (1992) Omission of exon 12 in cystic fibrosis transmembrane conductance regulator (CFTR) gene transcripts. Hum Genet 89: 615-619

Smit LS, Strong TV, Wilkinson DJ, Macek M Jr, Mansoura MK, Wood DL, Cole JL, Cutting GR, Cohn JA, Dawson DC (1995) Missense mutation (G480C) in the CFTR gene associated with protein mislocalization but normal chloride channel activity. Hum Mol Genet 4: 269-273

Smith AN, Barth ML, McDowell TL, Moulin DS, Nuthall HN, Hollingsworth MA, Harris A (1996) A regulatory element in intron 1 of the cystic fibrosis transmembrane conductance regulator gene. J Biol Chem 271: 9947-9954

Solberg R, Sistonen P, Traskelin AL, Berube D, Simard J, Krajci P, Jahnsen T, de la Chapelle A (1992) Mapping of the regulatory subunits RI beta and RII beta of cAMP-dependent protein kinase genes on human chromosome 7. Genomics 14: 63-69

Sosa-Pineda B, Chowdhury K, Torres M, Oliver G, Gruss P (1997) The Pax4 gene is essential for differentiation of insulin-producing beta-cells in the mammalian pancreas. Nature 386: 399-402

Spence JE, Perciaccante RG, Greig GM, Willard HF, Ledbetter DH, Hejtmancik JF, Pollack MS, O'Brian WE, Beaudet AL (1988) Uniparental disomy as a mechanism for human genetic disease. Am J Hum Genet 42: 217-226

Spilmont C, Hinnrasky J, Zahm JM, Jacquot J, Puchelle E, Kuhry JG (1996) CFTR is involved in membrane endocytosis but not in fluid-phase and receptor mediated endocytosis in human respiratory epithelial cells. Biochem Biophys Res Comm 227: 182-188

Spotila LD, Sereda L, Prockop DJ (1992) Partial isodisomy for maternal chromosome 7 and short stature in an individual with a mutation at the COL1A2 locus. Am J Hum Genet 51: 1396-1405

Stern RC, Boat TF, Abramowsky CR, LeRoy WM, Wood RE, Doershuk CF (1978) Intermediate-range sweat chloride concentration and *Pseudomonas* bronchitis. A cystic fibrosis variant with preservation of exocrine pancreatic function. JAMA 239: 2676-2680

Strautnieks SS, Thompson RJ, Hanukoglu A, Dillon MJ, Hanukoglu I, Kuhnle U, Seckl J, Gardiner RM, Chung E (1996a) Localisation of pseudohypoaldosteronism genes to chromosome 16p12.2-13.11 and 12p13.1-pter by homozygosity mapping. Hum Mol Genet 5: 293-299

Strautnieks SS, Thompson RJ, Gardiner RM, Chung E (1996b) A novel splice mutation in the γ subunit of the epithelial sodium channel gene in three pseudohypoaldosteronism type 1 families. Nat Genet 13: 248-250

Stirling B, Cox NJ, Bell GI, Hanis CL, Spielman RS, Concannon P (1995) Identification of microsatellite markers near the human *OB* gene and linkage studies in NIDDM-affected sib pairs. Diabetes 44: 999-1001

Strobel A, Issad T, Camoin L, Ozata M, Strosberg AD (1998) A leptin missense mutation associated with hypogonadism and morbid obesity. Nat Genet 18: 213-215

Strong TV, Smit LS, Turpin SV, Cole JL, Horn CT, Markiewicz D, Petty TL, Craig MW, Rosenow EC, Tsui LC, Iannuzzi MC, Knowles MR, Collins FS (1991) Cystic fibrosis gene mutation in two sisters with mild disease and normal sweat electrolyte levels. New Engl J Med 325: 1630-1634

Strong TV, Wilkinson DJ, Mansoura MK, Devor DC, Henze K, Yang Y, Wilson JM, Cohn JA, Dawson DC, Frizzell RA, Collins FS (1993) Expression of an abundant alternatively spliced form of the cystic fibrosis transmembrane conductance regulator (CFTR) gene is not associated with a cAMP activated chloride conductance. Hum Mol Genet 2: 225-230

Stuart AB, Burdon MG (1974) Letter: frequency of the cystic-fibrosis gene. Lancet 2: 1521

Sturtevant AH (1913) The linear arrangement of sex-linked factors in Drosophila, as shown by their mode of association. J Exp Zool 14: 43-59

Stutts MJ, Lazarowski ER, Paradiso AM, Boucher RC (1995) Activation of CFTR Cl^- conductance in polarized T84 cells by luminal extracellular ATP. Am J Physiol 268: C425-C433

Stutts MJ, Rossier BC, Boucher RC (1997) Cystic fibrosis transmembrane conductance regulator inverts protein kinase A-mediated regulation of epithelial sodium channel single channel kinetics. J Biol Chem 272: 14037-14040

Sugita M, Yue Y, Foskett JK (1998) CFTR Cl^- channel and CFTR associated ATP channel: distinct pores regulated by common gates. EMBO J 17: 898-908

Sweezy NB, Gauthier C, Gagnon S, Ferretti E, Kopelman H (1996) Progesterone and estradiol inhibit CFTR-mediated ion transport by pancreatic epithelial cells. Am J Physiol 271: G747-G754

Tanksley SD (1993) Mapping polygenes. Ann Rev Genet 27: 205-233

Teem JL, Berger HA, Ostedgaard LS, Rich DP, Tsui LC, Welsh MJ (1993) Identification of revertants for the cystic fibrosis ΔF508 mutation using STE6-CFTR chimeras in yeast. Cell 73: 335-346

Teem JL, Carson MR, Welsh MJ (1996) Mutation R553Q in CFTR-ΔF508 enhances function and partially corrects defective processing. Receptors and channels 4: 62-67

Teng H, Jorissen M, van Poppel H, Legius E, Cassiman JJ, Cuppens H (1997) Increased proportion of exon 9 alternatively spliced CFTR transcripts in vas deferens compared with nasal epithelial cells. Hum Mol Genet 6: 85-90

Thinnes FP, Reymann S (1997) New findings concerning vertebrate porin. Naturwissenschaften 84: 480-498

Tiso N, Pallavicini A, Muraro T, Zimbello R, Apolloni E, Valle G, Lanfranchi G, Danieli GA (1996) Chromosomal localisation of the human genes CPP32, Mch2, Mch3 and Ich-1, involved in cellular apoptosis. Biochem Biophys Res Commun 225: 983-989

Tsui LC, Buchwald M, Barker D, Braman JC, Knowlton R, Schumm JW, Eiberg H, Mohr J, Kennedy D, Plasvic N, Zsiga M, Markiewicz D, Akots G, Brown V, Helms C, Gravius T, Parker C, Rediker K, Donis-Keller H (1985) Cystic fibrosis locus defined by a genetically linked polymorphic marker. Science 230: 1054-1057

Tümmler B, Aschendorff A, Darnedde T, Fryburg K, Maaß G, Hundrieser J (1990) Marker haplotype association with growth in German cystic fibrosis patients. Hum Genet 84: 267-273

Ueyama H, Deng HX, Ohkubo I (1993) Molecular cloning and chromosomal assignment of the gene for human Zn-alpha-2-glycoprotein. Biochemistry 32: 12968-12976

Vajanaphanich M, Schultz C, Rudolf MT, Wasserman M, Enyedi P, Craxton A, Shears SB, Tsien RY, Barret KE, Traynor-Kaplan A (1994) Long-term uncoupling of chloride secretion from intracellular calcium levels by Ins(3,4,5,6)P_4. Nature 371: 711-714

Vajanaphanich M, Schultz C, Tsien RY, Traynor-Kaplan AE, Pandol SJ, Barret KE (1995) Cross-talk between calcium and cAMP-dependent intracellular signalling pathways. J Clin Invest 96: 386-393

Vankeerberghen A, Wei L, Jaspers M, Cassiman JJ, Nilius B, Cuppens H (1998a) Characterization of 19 disease-associated missense mutations in the regulatory domain of the cystic fibrosis transmembrane conductance regulator. Hum Mol Genet 7: 1761-1769

Vankeerberghen A, Wei L, Teng H, Jaspers M, Cassiman JJ, Nilius B, Cuppens H (1998b) Characterization of mutations located in exon 18 of the CFTR gene. FEBS Lett 437: 1-4

Vasquez C, Antiñolo G, Casals T, Dapena J, Elotz J, Seculi JL, Sirvent J, Cabanas R, Soler C, Estivill X (1996) Thirteen cystic fibrosis patients, 12 compound heterozygous and one homozygous for the missense mutation G85E: a pancreatic sufficiency/insufficiency mutation with variable clinical presentation. J Med Genet 33: 820-822

Veeze HJ, Sinaasappel M, Bijman J, Bouquet J, de Jonge HR (1991) Ion transport abnormalities in rectal suction biopsies from children with cystic fibrosis. Gastroenterology 101: 398-403

Veeze HJ, Halley DJJ, Bijman J, de Jongste JC, de Jonge HR, Sinaasappel M (1994) Determinants of mild clinical symptoms in cystic fibrosis patients. Residual chloride secretion measured in rectal biopsies in relation to the genotype. J Clin Invest 93: 461-466

Verlingue C, Vuillaumer S, Mercier B, LeGac G, Elion J, Férec C, Denamur E (1998) Absence of mutations in the interspecies conserved regions of the CFTR promoter region in cystic fibrosis (CF) and CF related patients. J Med Genet 35: 137-140

Voilley N, Linguelia E, Champigny G, Mattéi MG, Waldmann R, Lazdunski M, Barby P (1994) The lung amiloride-sensitive Na^+ channel: Biophysical properties, pharmacology, ontogenesis and molecular cloning. Proc Natl Acad Sci USA 91: 247-251

Voilley N, Bassilana F, Mignon C, Merscher S, Mattei MG, Carle GF, Lazdunski M, Barby P (1995) Cloning, chromosomal localization, and physical linkage of the β and γ subunits of the human epithelial amiloride-sensitive sodium channel. Genomics 28: 560-565

de Vos P, Saladin R, Auwerx J, Staels B (1995) Induction of ob gene expression by corticosteroids is accompanied by body weight loss and reduced food intake. J Biol Chem 270: 15958-15961

Voss R, Ben-Simon E, Avital A, Godfrey S, Zlotogora J, Dagan J, Tikochinski Y, Hillel J (1989) Isodisomy of chromosome 7 in a patient with cystic fibrosis: Could uniparental disomy be common in humans? Am J Hum Genet 45: 373-380

de Vries HG, Collee JM, de Walle HEK, van Veldhuizen MHR, Sibinga CTS, Scheffer H, ten Kate LP (1997) Prevalence of ΔF508 cystic fibrosis carriers in The Netherlands: logistic regression on sex, age, region of residence and number of offspring. Hum Genet 99: 74-79

Walsh PS, Metzger DH, Higuchi R (1991): Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. Biotechniques 10: 506-513

Ward CL, Kopito RR (1994) Intracellular turnover of cystic fibrosis transmembrane conductance regulator. Inefficient processing and rapid degradation of wild-type and mutant proteins. J Biol Chem 269: 25710-25718

Ward CL, Omura S, Kopito RR (1995) Degradation of CFTR by the ubiquitin-proteasome pathway. Cell 83: 121-127

Warner JO, Norman AP, Soothill JF (1976) Cystic fibrosis heterozygosity in the pathogenesis of allergy. Lancet 1: 990-991

Watt WC, Lazarowski ER, Boucher RC (1998) Cystic fibrosis transmembrane regulator-independent release of ATP. Its implications for the regulation of P2Y2 receptors in airway epithelia. J Biol Chem 273: 14053-14058

Weber E. Grundriss der biologischen Statistik. 8th ed. Stuttgart, Fischer Verlag; 1980

Wei X, Eisman R, Xu J, Harsch AD, Mulberg AE, Bevins CL, Glick MC, Scanlin TF (1996) Turnover of the cystic fibrosis transmembrane conductance regulator (CFTR): slow degradation of wild-type and ΔF508 CFTR in surface membrane preparations of immortalized airway epithelial cells. J Cell Physiol 168: 373-384

Weidner KM, Sachs M, Birchmeier W (1993) The *Met* receptor tyrosine kinase transduces motility, proliferation, and morphogenic signals of scatter factor / hepatocyte growth factor in epithelial cells. J Cell Biol 121: 145-154

Welsh MJ, Tsui LC, Boat TF, Beaudet AL (1995) Cystic Fibrosis. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease. McGraw Hill, New York, pp 3799-3876

Wilschanski M, Zielenski J, Markiewicz D, Tsui LC, Corey M, Levinson H, Durie PR (1995) Correlation of sweat chloride concentration with classes of the cystic fibrosis transmembrane conductance regulator gene mutations. J Pediatr 127: 705-710

Will K, Dörk T, Stuhrmann M, von der Hardt H, Ellemunter H, Tümmler B, Schmidtke J (1995) Transcript analysis of CFTR nonsense mutations in lymphocytes and nasal epithelial cells from cystic fibrosis patients. Hum Mut 5: 210-220

Witt M, Erickson RP, Ober C, Howatt WF, Farber R (1991) Correlation of phenotypic and genetic heterogeneity in cystic fibrosis: variability in sweat electrolyte levels contributes to heterogeneity and is increased with the XV-2c/KM.19 haplotype. Am J Med Genet 39: 137-143

Woods SC, Seeley RJ, Porte D Jr, Schwartz MW (1998) Signals that regulate food intake and energy homeostasis. Science 280: 1378-1382

Wrighton SA, Stevens JC (1992) The human hepatic cytochromes P450 involved in drug metabolism. Crit Rev Toxicol 22: 1-21

Xie J, Drumm M, Ma J, Davies PB (1995) Intracellular loop between transmembrane segments IV and V of cystic fibrosis transmembrane conductance regulator is involved in regulation of chloride channel conductance state. J Biol Chem 270: 28084-28091

Yang Y, Janich S, Cohn JA, Wilson JM (1993) The common varant of cystic fibrosis transmembrane conductance regulator is recognized by hsp70 and degraded in a pre-golgi nonlysosomal compartment. Proc Natl Acad Sci USA 90: 9480-9484

Yoshimura K, Chu CS, Crystal RG (1993) Alternative splicing of intron 23 of the human cystic fibrosis transmembrane conductance regulator gene resulting in a novel exon and transcript coding for a shortened intracytoplasmatic C terminus. J Biol Chem 268: 686-690

Zeitlin PL (1999) Novel pharmacologic therapies for cystic fibrosis. J Clin Invest 103: 447-452

Zielenski J, Rozmahel R, Bozon D, Kerem BS, Grzelczak Z, Riordan JR, Rommens J, Tsui LC (1991) Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Genomics 10: 214-228

Zielenski J, Bozon D, Markiewicz D, Aubin G, Simard F, Rommens JM, Tsui LC (1993) Analysis of transcripts in nasal epithelial cells and lymphoblasts of a cystic fibrosis patient with 621+1G→T and 711+1G→T mutations. Hum Mol Genet 2: 683-687

Zielenski J, Tsui LC (1995) Cystic fibrosis: genotypic and phenotypic variations. Annu Rev Genetics 29: 777-807

Zielenski J, Corey M, Markiewicz D, Aznarez I, Rozmahel R, Durie P, Tsui LC and the CF modifier study collaborative group (1998) Fine mapping of the cystic fibrosis modifier 1 (CFM1) gene on human chromosome 19 and assessment of candidate genes. Pediatr Pulmonol 16: P339

8 APPENDIX

8.1 List of figures and tables

FIGURES

TABLES

8.2 List of frequently used abbreviations

ARU	arbitrary repeat units (for definition see figure 8 in chapter 2.1.2.2.4 on page 43)
CON+	concordant/mild disease (for definition see figure 12 in chapter 2.2.1.4 on page 53)
CON-	concordant/severe disease (for definition see figure 12 in chapter 2.2.1.4 on page 53)
DELTA	for definition see figure 10 in chapter 2.2.1.3 on page 48
DfO	distance from origin (for definition see figure 10 in chapter 2.2.1.3 on page 48)
DIDS	4-4'-diisothiocyanostilbene-2,2'-disulfonic acid
DIS	discordant (for definition see figure 12 in chapter 2.2.1.4 on page 53)
ENaC	amiloride-sensitive epithelial sodium channel
FEV1	forced expiratory volume in 1 s
FEV1%	FEV1 as % of predicted FEV1
FEVPerc	percentiles for FEV1%pred (for definition see figure 10 in chapter 2.2.1.3 on page 48)
IBD	identity by descent
IBS	identity by state
ICM	intestinal current measurement
NPD	nasal potential difference
PCR	polymerase chain reaction
RFLP	restriction fragment length polymorphism
SNP	single nucleotide polymorphism
SSCP	single strand conformation polymorphism
wfh%	weight as % of weight predicted for height centile (for definition see figure 10 in chapter 2.2.1.3 on page 48)

8.3 Appendix to chapter 2

8.3.1 References for primer sequences

Chromosome 12 microsatellite locus near SCNN1A:

Marker locus	GDB accession number / reference
D12889	363193

Chromosome 16 microsatellite locus near SCNN1B, SCNN1C:

Marker locus	GDB accession number / reference
βENaCGT	434736

Chromosome 7 marker loci:

Marker locus	GDB accession number / reference	
D7S495	188164	
D7S514	188404	
D7S525	188521	
PON2	—	Adkins *et al* 1993 (modified) 5'-GGCACAAATGATCACTATTTTCTTGAC 5'-GAGTAAATCCACTACATTTCA
MetH		Horn *et al.* 1989
XV-2c		Rosenbloom *et al.* 1989
KM.19		Anwar *et al.* 1990
TUB20		Dörk *et al.* 1992
J3.11		Horn *et al.* 1989

4 kb sequence preceding *CFTR* exon 1:

genomic sequence:	M58478	(EMBL accession number)
and	M58479	(EMBL accession number)
Primers used for amplification:		5'-CAAAGGAAAACATAAGATGCAATTCG 5'-CAAATCCTGAATATCATGTCGCAA 5'-TGTCCAGATGCACTAATTGCGA 5'-GACCCTTGCCTTAGATGTGTCG 5'-TGGCCTGATTTTATTGCCG 5'-TGAAAAAAAGTTTGGAGACAACGC

These primer sequences were chosen based on the algorithm of Griffais *et al* (1991).

***CFTR* exons and flanking intron sequences:**

Unless noted differently in Dörk *et al.* 1994b, the primers described by Zielenski *et al.* 1991 were used for amplification of the CFTR exons and flanking intron sequences. For mutation screening of patients D, I, P, Q, R and S, the CFTR exons 1, 2, 3, 5, 6a, 6b, 8, 12, 14a, 14b, 15, 16, 18, 23, 24 were analysed by SSCP (see tables 5 and 10 for details). The following primer sequences were used to amplify the respective exons and flanking intron sites:

Exon	primer sequences	
1	5'-GGTCTTTGGCATTAGGAGCTTG	5'-ACACGCCCTCCTCTTTCGTG
2	5'-CATAATTTTCCATATCGGAG	5'-TAGCCACCATACTTGGCTCC
3	5'-CTTGGGTTAATCTCCTTGG	5'-ATTCACCAGATTTCGTAGTC
5	5'-ATTTCTGCCTAGATGCTGGG	5'-AACTCCGCCTTTCCAGTTG
6a	5'-GTGCTCAGAACCACGAAGTG	5'-GGGATGACAGATCTAGACTC
6b	5'-GGAATGAGTCTGTACAGCG	5'-GAGGTGGAAGTCTACCATG

8	5'-AAGATGTAGCACAATGAGAG	5'-GTGAGTGATCCTCCTTCCAG
12	5'-GTGAATCGATGTGGTGACCA	5'-CTGGTTTAGCATGAGGCGGT
14a	5'-AAAAGGTATGCCACTGTTAA	5'-GTATACATCCCCAAACTATCT
14b	5'-GAACACCTAGTACAGCTGC	5'-CACCTCACCCAACTAATGGTCATC
15	5'-GTGCATGCTCTTCTAATGC	5'-AAGGCACATGCCTCTGTGCA
16	5'-CAGAGAAATTGGTCGTTAC	5'-CTAAATGTGGGATTGCCTC
18	5'-GTAGATGCTGTGATGAACTG	5'-AGTGGCTACTCATGAGAAGG
23	5'-AGCTGATTGTGCGTAACGC	5'-TAAAGCTGGATGGCTGTATG
24	5'-GGACACAGCAGTTAAATGTG	5'-CCAATTCCATGAGGTGACTGTCCC

8.3.2 Common reagents

Common reagents:

SSC (20×)	3M NaCl, 0.3M NaCitrat
STE	50mM Tris-HCl pH 7.5, 100mM NaCl, 1mM Na_2EDTA
TBE (10×)	0.9M Tris-HCl pH 8.3, 0.9M Boric acid, 0.02M Na_2EDTA
TE	10mM Tris-HCl pH 8.0, 1mM Na_2EDTA
Meyler buffer	126.2 mM Na^+, 114.3 mM Cl^-, 20.2 mM HCO_3^-, 0.3 mM HPO_4^{2-}, 0.4 mM $H_2PO_4^-$, 10 mM Hepes, pH7.4

Reagents provided by manufacturers:

TaqPolymerase reaction buffer (Invitek):
670mM Tris-HCl pH 8.8; 160 mM $(NH_4)_2SO_4$, 0.1%Tween20
DapGoldstar TaqPolymerase reaction buffer (Eurogentec):
750mM Tris-HCl pH 9.0; 200mM $(NH_4)_2SO_4$, 0.1%Tween20

8.3.3 Regression formulas for height, weight and FEV1%pred percentiles

All centiles were described by areas defined by a polynomial equation z = f(x,y). Censoring of data was unavoidable for some patients with unusual combination of extreme values in age and pulmonary function or weight and height, because the regression curves resulted in implausible percentiles of larger than 100 or lower than 0. As negative values or values of more than 100 are not defined for percentiles, these patients could not be described with the regression parameters employed and thus were excluded from further analysis.

In detail, the following calculation parameters were used:

length centiles for females:

x = age [years], y = length [cm], area defines z= length centiles

z=a00+a01*y+a02*y^2+a03*y^3+a04*y^4+a05*y^5+a10*x+a11*x*y+a12*x*y^2+a13*x*y^3+a14*x*y^4+a15*x*y^5+a20*x^2+a21*x^2*y+a22*x^2*y^2+a23*x^2*y^3+a24*x^2*y^4+a25*x^2*y^5+a30*x^3+a31*x^3*y+a32*x^3*y^2+a33*x^3*y^3+a34*x^3*y^4+a35*x^3*y^5+a40*x^4+a41*x^4*y+a42*x^4*y^2+a43*x^4*y^3+a44*x^4*y^4+a45*x^4*y^5+a50*x^5+a51*x^5*y+a52*x^5*y^2+a53*x^5*y^3+a54*x^5*y^4+a55*x^5*y^5

a00 = 48,3151
a01 = 0,34407
a02 = -0,0151914
a03 = 0,000340005
a04 = -0,00000346078
a05 = 0,0000000131611
a10 = 22,8933
a11 = 0,0867497
a12 = -0,00173392
a13 = 0,0000373274
a14 = -0,000000590913
a15 = 0,00000000343674
a20 = -4,20072
a21 = -0,00833215
a22 = -0,000819702
a23 = 0,0000290106
a24 = -0,000000306396
a25 = 0,000000000996832
a30 = 0,432829
a31 = 0,00350085
a32 = 0,00000752944
a33 = -0,00000214429
a34 = 0,0000000288737
a35 = -0,000000000105779
a40 = -0,020147
a41 = -0,000318564
a42 = 0,00000597415
a43 = -0,0000000189258
a44 = -0,000000000292056
a45 = 1,26083E-12
a50 = 0,000338875
a51 = 0,00000783121
a52 = -0,000000206802

a53 = 0,00000000247261
a54 = -1,57831E-11
a55 = 5,64011E-14

length centiles for males:

x = age [years], y = length [cm], area defines z= length centiles

'z=a00+a01*y+a02*y^2+a03*y^3+a04*y^4+a10*x+a11*x*y+a12*x*y^2+a13*x*y^3+a14*x*y^4+a20*x^2+a21*x^2*y+a22*x^2*y^2+a23*x^2*y^3+a24*x^2*y^4+a30*x^3+a31*x^3*y+a32*x^3*y^2+a33*x^3*y^3+a34*x^3*y^4+a40*x^4+a41*x^4*y+a42*x^4*y^2+a43*x^4*y^3+a44*x^4*y^4+a50*x^5+a51*x^5*y+a52*x^5*y^2+a53*x^5*y^3+a54*x^5*y^4

a00 = 50,4847
a01 = 0,185647
a02 = -0,00473329
a03 = 0,0000599715
a04 = -0,000000243954
a10 = 19,3204
a11 = 0,194157
a12 = -0,00197079
a13 = -0,0000144381
a14 = 0,000000229552
a20 = -2,44298
a21 = -0,0725234
a22 = 0,00102231
a23 = -0,0000000974229
a24 = -0,0000000532814
a30 = 0,169349
a31 = 0,00985
a32 = -0,000145367
a33 = 0,000000182424
a34 = 0,0000000062406
a40 = -0,00437449
a41 = -0,000522302
a42 = 0,00000731626
a43 = -0,00000000301641
a44 = -0,000000000371049
a50 = 0,0000160633
a51 = 0,00000933779
a52 = -0,000000114827
a53 = -0,000000000259881
a54 = 8,51463E-12

weight centiles for females:

x = age [years], y = length [cm], area defines z= weight centiles

z=a00 + a01 * y + a02 * y ^ 2 + a03 * y ^ 3 + a04 * y ^ 4 + a05 * y ^ 5 + a10 * x + a11 * x * y + a12 * x * y ^ 2 + a13 * x * y ^ 3 + a14 * x * y ^ 4 + a15 * x * y ^ 5 + a20 * x ^ 2 + a21 * x ^ 2 * y + a22 * x ^ 2 * y ^ 2 + a23 * x ^ 2 * y ^ 3 + a24 * x ^ 2 * y ^ 4 + a25 * x ^ 2 * y ^ 5 + a30 * x ^ 3 + a31 * x ^ 3 * y + a32 * x ^ 3 * y ^ 2 + a33 * x ^ 3 * y ^ 3 + a34 * x ^ 3 * y ^ 4 + a35 * x ^ 3 * y ^ 5 + a40 * x ^ 4 + a41 * x ^ 4 * y + a42 * x ^ 4 * y ^ 2 + a43 * x ^ 4 * y ^ 3 + a44 * x ^ 4 * y ^ 4 + a45 * x ^ 4 * y ^ 5 + a50 * x ^ 5 + a51 * x ^ 5 * y + a52 * x ^ 5 * y ^ 2 + a53 * x ^ 5 * y ^ 3 + a54 * x ^ 5 * y ^ 4 + a55 * x ^ 5 * y ^ 5

a00 = 2,71691
a01 = 0,103324

a02 = -0,00521835
a03 = 0,000125116
a04 = -0,00000137348
a05 = 0,00000000560491
a10 = 5,26919
a11 = -0,0582872
a12 = 0,00646361
a13 = -0,000179684
a14 = 0,00000214365
a15 = -0,00000000909209
a20 = -1,13665
a21 = 0,0652821
a22 = -0,00510643
a23 = 0,000137522
a24 = -0,00000160224
a25 = 0,00000000667259
a30 = 0,132086
a31 = -0,012776
a32 = 0,000966179
a33 = -0,0000258209
a34 = 0,000000297988
a35 = -0,00000000122846
a40 = -0,00611844
a41 = 0,000943505
a42 = -0,0000687829
a43 = 0,0000018189
a44 = -0,0000000208099
a45 = 8,51988E-11
a50 = 0,0000971125
a51 = -0,0000230746
a52 = 0,00000162686
a53 = -0,0000000425435
a54 = 0,000000000482632
a55 = -1,96363E-12

weight centiles for males:

x = age [years], y = length [cm], area defines z= weight centiles

z= a00 + a01 * y + a02 * y ^ 2 + a03 * y ^ 3 + a04 * y ^ 4 + a05 * y ^ 5 + a10 * x + a11 * x * y + a12 * x * y ^ 2 + a13 * x * y ^ 3 + a14 * x * y ^ 4 + a15 * x * y ^ 5 + a20 * x ^ 2 + a21 * x ^ 2 * y + a22 * x ^ 2 * y ^ 2 + a23 * x ^ 2 * y ^ 3 + a24 * x ^ 2 * y ^ 4 + a25 * x ^ 2 * y ^ 5 + a30 * x ^ 3 + a31 * x ^ 3 * y + a32 * x ^ 3 * y ^ 2 + a33 * x ^ 3 * y ^ 3 + a34 * x ^ 3 * y ^ 4 + a35 * x ^ 3 * y ^ 5 + a40 * x ^ 4 + a41 * x ^ 4 * y + a42 * x ^ 4 * y ^ 2 + a43 * x ^ 4 * y ^ 3 + a44 * x ^ 4 * y ^ 4 + a45 * x ^ 4 * y ^ 5 + a50 * x ^ 5 + a51 * x ^ 5 * y + a52 * x ^ 5 * y ^ 2 + a53 * x ^ 5 * y ^ 3 + a54 * x ^ 5 * y ^ 4 + a55 * x ^ 5 * y ^ 5

a00 = 3,0688
a01 = 0,0999505
a02 = -0,00413342
a03 = 0,0000958518
a04 = -0,00000103228
a05 = 0,00000000408829
a10 = 4,4228
a11 = 0,011136
a12 = -0,000399856
a13 = 0,0000155581

a14 = -0,000000254645
a15 = 0,00000000163697
a20 = -0,620168
a21 = 0,0304986
a22 = -0,0016539
a23 = 0,0000391129
a24 = -0,000000397227
a25 = 0,00000000133317
a30 = 0,0567971
a31 = -0,00793663
a32 = 0,000415316
a33 = -0,00000988668
a34 = 0,000000102572
a35 = -0,000000000363985
a40 = -0,00231202
a41 = 0,000684378
a42 = -0,0000334448
a43 = 0,000000776112
a44 = -0,00000000796942
a45 = 0,000000000028344
a50 = 0,0000421339
a51 = -0,0000184095
a52 = 0,000000839923
a53 = -0,0000000187932
a54 = 0,000000000188535
a55 = -6,58876E-13

FEVPerc – centiles for FEV1%pred:

x = FEV1%pred (defined by Knudson et al. 1983), y = age [years], area defines z=FEVPerc

z= 54,2282 - 6,48596 * y + 0,160673 * y ^ 2 - 5,79524 * x + 0,686198 * x * y - 0,0183229 * x * y ^ 2 + 0,229493 * x ^ 2 - 0,0267976 * x ^ 2 * y + 0,000755553 * x ^ 2 * y ^ 2 - 0,00399186 * x ^ 3 + 0,000423935 * x ^ 3 * y - 0,000010645 * x ^ 3 * y ^ 2 + 0,0000303584 * x ^ 4 - 0,00000261721 * x ^ 4 * y + 0,0000000507948 * x ^ 4 * y ^ 2 - 0,0000000785657 * x ^ 5 + 0,00000000480301 * x ^ 5 * y - 3,86546E-11 * x ^ 5 * y ^ 2

LEBENSLAUF

Name: Mekus
Vorname: Frauke

Geboren am 12.03.1970 in Essen als Tochter von
Eckhard und Inge Mekus

Besuch der Grundschule Vinnhorst, Hannover
von August 1976 bis Juli 1980

Besuch der Orientierungsstufe Entenfangweg, Hannover
von August 1980 bis Juni 1982

Einschulung im Gymnasium Lutherschule, Hannover
im Juni 1982
Erwerb des Sekundarabschlusses I im Juni 1986
Schulabschluß mit Erwerb der allgemeinen Hochschulreife im
Mai 1989

Immatrikulation an der Universität Hannover
im Studiengang Biochemie
zum Wintersemester 1989

Abschluß des Grundstudiums mit Vordiplom
im September 1991

Mündliche Diplomprüfung im April 1994,
daran anschließend Diplomarbeit
"*Makrorestriktionsanalyse und Mikrosatelliten im CFTR Gen*"
in der klinischen Forschergruppe
"*Molekulare Pathologie der Mukoviszidose*"
an der medizinischen Hochschule Hannover
Abschluß des Studienganges zum Diplom-Biochemiker
im Oktober 1994

Beginn der Promotion November 1994
in der klinischen Forschergruppe
"*Molekulare Pathologie der Mukoviszidose*"
an der medizinischen Hochschule Hannover