

Aus der Medizinischen Klinik und Poliklinik  
der Universität Ulm - Klinikum

Abteilung Transfusionsmedizin  
(Direktor: Prof. Dr. med. Bernhard Kubanek)

**HÄUFIGKEIT SPORADISCHER NICHT-FUNKTIONALER ALLELE  
UND IHRE BEDEUTUNG FÜR DIE GENOTYPISIERUNG AM BEISPIEL DES  
POLYMORHISMUS IM *FUT1*-BLUTGRUPPENGEN**

Habilitationsschrift  
im Fachgebiet Transfusionsmedizin und Immunologie  
vorgelegt der  
Medizinischen Fakultät

Dr. med. Willy A. Flegel

1997



**In die lenge  
wirt man  
gewar der  
warheit als: lange  
gelernet, etwas ge-  
kundet.**

Johannes von Tepl, 1401

# Bombay-Allelpolymorphismus

Inhaltsverzeichnis		Seite
<b>1.</b>	<b>EINLEITUNG</b>	
1.1	Allelie bei Genen und in der Population	5
1.1.1	Nicht-funktionale Allele	5
1.1.2	Verteilung der Allele eines Gens	6
1.1.3	Häufigkeit nicht-funktionaler Allele in Populationen	7
1.1.4	Bedeutung nicht-funktionaler Allele für Strategien der Genotypisierung	8
1.2	H-Blutgruppensystem	8
1.2.1	Phänotyp und Serologie der Bombay ( $O_h$ )-Blutgruppe	9
1.2.2	Molekularbiologie der Fukosyltransferase-Gene <i>FUT1</i> und <i>FUT2</i>	10
1.2.3	Interaktion zwischen verschiedenen Glykosyltransferasen	11
1.3	Biologische Funktion von Blutgruppenantigenen	13
1.3.1	Proteinantigene	13
1.3.2	Zuckerantigene	15
1.4	Fragestellung: <i>FUT1</i> -Gen als Modellsystem für nicht-funktionale Allele	17
<b>2.</b>	<b>METHODEN</b>	
2.1	Identifikation von Individuen mit Bombay ( $O_h$ )-Blutgruppe in der Bevölkerung	19
2.1.1	Blutgruppenbestimmung	19
2.1.2	Blutspender-Datenbasis	19
2.1.3	Statistik	20
2.2	Probanden	20
2.3	Molekularbiologische Untersuchungen	21
2.4	Populationsgenetische Analyse	23

	Seite
<b>3. ERGEBNISSE</b>	
3.1	Frequenzen von seltenen Blutgruppen-Phänotypen 27
3.2	Analyse nicht-funktionaler Allele ( <i>h</i> -Allele) im <i>FUT1</i> -Gen 28
	3.2.1 Mutationen in <i>h</i> -Allelen 29
	3.2.2 Homozygotie für private <i>h</i> -Allele 36
3.3	Statistische Betrachtung zum <i>h</i> -Polymorphismus 36
	3.3.1 Hardy-Weinberg Equilibrium bei seltenen Allelen 36
	3.3.2 Unabhängige Mutationen als Ursache von <i>h</i> -Allelen 39
3.4	Populationsgenetische Analyse 39
	3.4.1 Mittlerer Konsanguinitätskoeffizient $\alpha$ 39
	3.4.2 Frequenz sporadischer <i>h</i> -Allele in der Population 40
<b>4. DISKUSSION</b>	
4.1	Häufigkeit sporadischer nicht-funktionaler Allele im <i>FUT1</i> -Gen 41
4.2	Polymorphismus nicht-funktionaler Allele im <i>FUT1</i> -Gen 41
4.3	Mittlerer Konsanguinitätskoeffizient $\alpha$ in der deutschen Bevölkerung 44
4.4	Verlässlichkeit der Frequenzabschätzung von Null-Allelen in der Population 45
4.5	Bedeutung sporadischer nicht-funktionaler Allele für die Geno- typisierung am Beispiel des AB0-Blutgruppensystems 46
4.6	Bedeutung bei Onkogenen 48
4.7	Bedeutung bei erbgenetischen Fragestellungen 49
4.8	Genotypisierungsstrategien in Gegenwart von sporadischen nicht-funktionalen Allelen 49
4.9	Qualitätsgesicherte Genotypisierung 50
4.10	Ausblick 52
<b>5. ABKÜRZUNGEN</b>	53
<b>6. ZUSAMMENFASSUNG UND ÜBERSICHT EIGENER ARBEITEN</b>	
6.1	Zusammenfassung 55
6.2	Summary 56
6.3	Immunantwort auf bakterielle Toxine 57
	6.3.1. Endotoxin-Inhibition durch humane Lipoproteine 57
	6.3.2. Aktivierung humaner Monozyten durch <i>Clostridium difficile</i> Toxin B 58
6.4	Homolog eines humanen Leukämie-Onkogens in <i>Drosophila melanogaster</i> 58
6.5	Molekularbiologie und Populationsgenetik von Blutgruppengenen 59
	6.5.1 <i>Rhesus</i> -Gene 59
	6.5.2 <i>FUT1</i> -Gen 60
<b>7. LITERATUR</b>	61
<b>8. DANKSAGUNG</b>	71

## 1. EINLEITUNG

### 1.1 Allelie bei Genen und in der Population

#### 1.1.1 Nicht-funktionale Allele

Als Gen wurde ursprünglich eine diskrete Einheit der Vererbung definiert, die als physisches Element die Grundlage für die Ausprägung einer Erbeigenschaft bildet. Dieses Verständnis setzt eine bestimmte Funktion eines Gens voraus. Allele bezeichnen die unterschiedlichen Formen eines Gens, die an seinem Genlocus auftreten können. Zumindest ein Allel - oft das am häufigsten auftretende Allel - muß eine Funktion aufweisen. Meistens ist die Funktionalität eines Allels durch die Expression seines Genprodukts bedingt. In diesem Fall setzt die Funktionalität bestimmte Eigenschaften der chromosomalen DNA-Sequenz voraus (Tabelle 1.1). Als Pseudogene oder Genfragmente werden chromosomale DNA-Sequenzen bezeichnet, die viele dieser Charakteristika aufweisen, bei denen aber ein funktionales Allel ausgeschlossen wurde oder aufgrund bestimmter Sequenzmerkmale sehr unwahrscheinlich ist.

Bereits die Änderung eines einzelnen Nukleotids in einem der aufgeführten Merkmale kann zu einem Verlust der Funktionalität des Gens führen. Die auf der genomischen DNA-Ebene relevanten Merkmale können je nach Größe des

**Tabelle 1.1. Einige für Gene charakteristische DNA- und mRNA-Merkmale**

Merkmal	Bedeutung
<i>genomische DNA</i>	
Promotor	Bindungsstelle der „basal transcription machinery“; beeinflusst Transkriptionseffizienz
Introns	beeinflussen Transkriptions- und Spleißeffizienz
Spleißstellen	beeinflussen Spleißeffizienz und mRNA-Sequenz
Enhancer	beeinflusst Transkriptionseffizienz
<i>mRNA oder cDNA</i>	
cDNA	zur mRNA komplementäre DNA
Start- und Stopkodon	bestimmen Beginn und Ende des „open reading frame (ORF)“
Translationsstart	Bindungsstelle für Ribosomen
kodierende Sequenz (CDS)	bestimmt Aminosäuresequenz des Proteinprodukts
Exons	bestimmen wesentlich die mRNA-Sequenz
poly(A)-Schwanz	beeinflusst mRNA-Stabilität
nicht-translatierte Sequenzen	umfaßt unter anderem Translationsstart und poly(A)-Schwanz

Genlocus üblicherweise einige 10.000 bis zu mehreren Millionen Nukleotide umfassen. Die systematische Bestimmung aller relevanten Merkmale auf der genomischen DNA-Ebene ist somit zur Zeit nicht praktikabel. Demgegenüber genügt für den Nachweis der relevanten Merkmale auf der mRNA- beziehungsweise cDNA-Ebene meist die Untersuchung von einigen tausend Nukleotiden, die deshalb technisch prinzipiell einfacher darstellbar ist.

Auch die chromosomale Organisation oder Proteine in trans-Position, zum Beispiel Enhancer-bindende Proteine, haben einen Einfluß auf die Transkription und somit auf die Präsenz einer mRNA. Die Funktionalität eines Proteins wird weiterhin von anderen Genloci beeinflusst, wenn die Funktion von Komplexen mit Proteinen dieser Genloci abhängt. So setzt die Integration der Kell- beziehungsweise Rhesus-Proteine in die Erythrozytenmembran die Präsenz anderer Proteine voraus. Einige  $K_0$ - beziehungsweise  $Rh_{null}$ -Phänotypen treten auf, wenn begleitende Proteine fehlen. Wird kein funktionales  $K_x$ -Protein exprimiert, das nicht mit dem Kell-Protein identisch ist, so tritt der  $K_0$ -Phänotyp auf. Deshalb fehlt bei dem  $K_0$ -Phänotyp das Kell-Protein in der Erythrozyten-Membran, obwohl funktionale Allele im *Kell*-Genlocus vorhanden sind [49]. Ähnlich verhält es sich beim  $Rh_{null}$ -Phänotyp mit einer  $Rh_{null}$ -Regulator Variante. In diesem Fall verhindert ein rezessiver Inhibitor eines autosomalen Genlocus, der nicht dem *Rhesus*-Genlocus entspricht, eine effiziente Expression der funktionalen *Rhesus*-Genprodukte in der Erythrozytenmembran [14].

Erbkrankheiten können durch verschiedene Arten von Mutationen in krankheitsassoziierten Genloci bedingt sein. Im allgemeinen überwiegen Punktmutationen. Seltener sind kleine Deletionen mit oder ohne Veränderung des Leserahmens. Im übrigen kommen Insertionen, Inversionen, große Deletionen und komplexe Rearrangements vor [16]. In verschiedenen Genloci können die Arten von Mutationen mit unterschiedlicher Häufigkeit auftreten. Diese Variabilität kann durch Charakteristika der Genloci bedingt sein wie Genlänge, Anzahl der Spleißstellen, DNA-Methylierung und Besonderheiten der Nukleotidsequenzen [16].

### 1.1.2 Verteilung der Allele eines Gens

Typischerweise sind in einem Genlocus ein oder wenige Allele prävalent. Im Vergleich hierzu kommen die meisten Allele insgesamt recht selten vor. Eine solche Verteilung von Allelen wird oft ungenau als „asymptotisch“ bezeichnet; mathe-

matische Betrachtungen ergaben [57], daß seltene Allele annähernd einer Poisson-Verteilung folgen. Interessanterweise läßt sich ableiten, daß ihre Verteilung im wesentlichen bestimmt wird von der effektiven Mutationsrate und weniger von einem etwaigen Selektionsdruck [57].

Viele Mutationen führen zu keinen Aminosäuresubstitutionen („silent mutations“). Etliche Aminosäureaustausche („missense mutations“) beeinflussen die Funktion eines Proteins nicht oder nur geringgradig. Für das Gen des Gerinnungsfaktors IX (Hämophilie B) wurde geschätzt [9], daß lediglich 40% aller möglichen „missense“-Mutationen einen Funktionsverlust bedingen und Ursache einer Hämophilie B sein können. Mutationen, die durch den Funktionsverlust eines Proteins und den dadurch bedingten klinischen Phänotyp nachgewiesen werden, stellen deshalb nur einen kleinen Teil aller vorkommenden Mutationen dar [16]. Die Mehrzahl aller Allele, die sich vom prävalenten funktionalen Allel (Wildtyp) unterscheiden, haben wahrscheinlich keine klinische Bedeutung.

### 1.1.3 Häufigkeit nicht-funktionaler Allele in Populationen

Abschätzungen zur Frequenz von nicht-funktionalen Allelen wurden bisher nur in der Fruchtfliege (*Drosophila melanogaster*) durchgeführt. Bei der Untersuchung von 25 Enzymen, deren Genloci auf Autosomen lokalisiert waren, wurde eine mittlere Frequenz von Null-Allelen von 0,0023 gefunden [65], wobei diesen Genloci ein Selektionsdruck unterstellt wurde. Bei X-chromosomalen Genloci wurden keine Null-Allele nachgewiesen, was mit dem bei X-chromosomalen Genen zu erwartenden hohen Selektionsdruck begründet wurde. Die mittlere Mutationsrate ist mit  $3,86 \times 10^{-6}$  angegeben [113]. In den beiden zitierten Studien um 1980 konnte die molekulare Ursache der beobachteten Null-Allele nicht geklärt werden. Eine vergleichbare Abschätzung bei vertebralen insbesondere menschlichen Genloci oder Genloci, die keinem Selektionsdruck unterliegen, ist unseres Wissens bisher nicht veröffentlicht worden.

Null-Allele für Blutgruppengene können durch serologische Untersuchungen in großen Populationen nachgewiesen werden. Für fast alle Blutgruppenantigene sind Null-Phänotypen belegt [55], und oft wurden ihre Frequenzen bestimmt.

#### 1.1.4 Bedeutung nicht-funktionaler Allele für Strategien der Genotypisierung

Viele der gegenwärtigen Genotypisierungsstrategien identifizieren die Allele durch den Polymorphismus bestimmter kurzer Nukleotidsequenzen. Bei dieser Strategie hängt die Vorhersage des Phänotyps von der funktionalen Integrität des Allels ab. Die korrekte Vorhersage des Phänotyps ist oft nicht möglich, wenn sporadische nicht-funktionale Allele (Null-Allele) angetroffen werden. In diesen sporadischen Allelen können für die Funktionalität wichtige Merkmale (Tabelle 1.1) verändert sein, die jedoch nicht die von der Genotypisierungsstrategie getesteten kurzen Nukleotidsequenzen betreffen. Solche Allele können durch bisher nicht bekannte Polymorphismen verursacht werden und kommen grundsätzlich in allen Genen vor. Sie wurden bereits in vielen verschiedenen Genen, einschließlich Blutgruppengenen [46,76,88,89,124] und Onkogenen [25,66], beschrieben. Ihre Häufigkeit in der Population und ihre Ursache sind von herausragender Bedeutung, um verlässliche Strategien der Genotypisierung zu entwickeln. Bisher sind ihre Häufigkeiten in der Population für kein Gen unter niedrigem Selektionsdruck oder für Vertebraten bestimmt worden.

#### 1.2 H-Blutgruppensystem

Von außerordentlich seltenen Ausnahmen abgesehen, tragen die Erythrozyten aller Menschen das H-Antigen. Das H-Antigen ist ein Zwischenprodukt von Membran-ständigen Zuckerketten, die in Gegenwart von funktionalen Allelen des *ABO*-Gens mit den entsprechenden A- beziehungsweise B-Antigenen enden. Die Erstbeschreibung von drei H-defizienten Personen geht auf Bende et al. (1952) in Bombay, Indien, zurück [8]. Im Speichel der drei Personen fand sich keine A-, B-, H- oder Le(b)-Substanz, wohl aber Le(a)-Substanz. Phänotypisch waren ihre Erythrozyten Le(a+ b-).

Die klinische Versorgung von Patienten mit der Bombay ( $O_h$ )-Blutgruppe ist äußerst schwierig. In vivo-Studien mit radioaktiven Chrom-markierten Erythrozyten belegen [19], daß  $O$ -Erythrozyten von dem anti-H innerhalb von Minuten hämolysiert werden. H-defiziente Patienten, bei denen klinisch relevantes anti-H nachweisbar ist (Wärmetechniken: 37°C, LISS, Antiglobulin), müssen deswegen mit Bombay ( $O_h$ )-Erythrozyten versorgt werden, die nur in sehr begrenztem Umfang in spezialisierten Blutspendediensten als tiefgefrorene Erythrozytenpräparate zur Verfügung stehen.

### 1.2.1 Phänotyp und Serologie der Bombay ( $O_h$ )-Blutgruppe

Levine et al. [74] führte die Bezeichnung  $O_h$  für den Bombay ( $O_h$ )-Phänotyp ein und belegte, daß dieser Phänotyp durch den zum Teil epistatischen Effekt eines Gens bedingt ist, wenn ein seltenes rezessives Allel homozygot vorliegt. Diese aufgrund der Suppression des Antigens B abgeleitete Hypothese wurde später bestätigt, als auch die analoge Suppression von Antigen  $A_1$  beziehungsweise Antigen  $A_2$  nachgewiesen wurde [1,51,75]. Yunis et al. [130] schlugen zwei unterschiedliche Genotypen als Ursache der serologischen Polymorphismen bei H-defizienten Probanden vor. Auch die Arbeitsgruppe von Issitt fand serologische Hinweise auf mindestens zwei verschiedene Bombay ( $O_h$ )-Phänotypen [81]. Biochemisch konnte die deutlich verminderte enzymatische Aktivität der  $\alpha$ -2-L-Fukosyltransferase bei H-defizienten Probanden gesichert werden [71].

Bombay ( $O_h$ )-Phänotypen treten am häufigsten in der Bevölkerung des Bundesstaates Maharashtra in Indien auf [7,99]. Weitere Häufungen finden sich auf Reunion im Indischen Ozean und in Durban, Südafrika [80].

Eine wesentliche Klärung wurde schließlich durch Oriol et al. [91] erreicht. Mit einer Reihe von systematischen Untersuchungen in den Familien von Probanden mit Bombay ( $O_h$ )-Phänotyp, die hauptsächlich unter der Bevölkerung von Reunion durchgeführt wurde, konnte diese Arbeitsgruppe ein genetisches Modell zur Entstehung des Phänotyps entwickeln [44,68-70]. In dem genetischen Modell wurden die geringen phänotypischen Unterschiede mit der Interaktion von zwei chromosomal eng benachbarten Genen ( $H$  und  $SE$  beziehungsweise  $FUT1$  und  $FUT2$ ) erklärt, die beide für Fukosyltransferasen kodieren. Schließlich wurde die epistatische Interaktion verschiedener Glykosyltransferasen für die Fukosylierung der Grundsubstanz beschrieben [90]. Das Modell ist im wesentlichen durch die molekularbiologischen Befunde bestätigt worden (siehe Abschnitte 1.2.2 und 1.2.3).

Personen, die keine funktionalen  $H$ - und  $Se$ -Allele tragen (Genotyp  $hh\ sese$ ), weisen serologisch die Bombay ( $O_h$ )-Blutgruppe auf. Der Phänotyp H-defizienter Sekretoren (Genotyp  $hh\ Sese$  oder  $hh\ SeSe$ ) unterscheidet sich von der Bombay ( $O_h$ )-Blutgruppe und wird häufig als serologischer „para Bombay“-Phänotyp bezeichnet [118,129]. Der Begriff „para Bombay“-Phänotyp ist allerdings nicht genau definiert und wird bisweilen auch auf Individuen mit schwach ausgeprägtem Antigen H angewandt.

### 1.2.2 Molekularbiologie der Fukosyltransferase-Gene *FUT1* und *FUT2*

An die wegen ihrer grundsätzlichen biologischen Bedeutung bei allen Menschen vorhandene Grundsubstanz (siehe Abschnitt 1.3.2) können verschiedene Zucker konjugiert werden. Hierbei interagieren verschiedene Glykosyltransferasen (siehe Tabelle 1.2 und Abschnitt 1.2.3).

Oriol et al. [91] postulierte zwei genetisch eng benachbarte Strukturgene für  $\alpha$ -2-L-Fukosyltransferasen. Die eine Fukosyltransferase (*FUT1*- oder *H*-Genprodukt) wird in mesodermalen Geweben exprimiert, synthetisiert das H-Antigen an der Grundsubstanz Typ 2 und ist beim Bombay ( $O_h$ )-Phänotyp nicht funktional. Eine zweite Fukosyltransferase (*FUT2*- oder *SE*-Genprodukt) tritt in epithelialen Geweben auf und kann das H-Antigen an beiden Grundsubstanzen („Precursor substance“ Typ 1 und Typ 2) synthetisieren. Sie bestimmt den Sekretorstatus. Wenn die häufig vorkommenden nicht-funktionalen *se*-Allele homozygot vorhanden sind, wird kein A-, B- oder H-Antigen im Speichel ausgeschieden.

Larson et al. [67] klonierten das *FUT1*-Gen. Sein Produkt ist die GDP-L-Fukose: $\beta$ -D-Galaktosid 2- $\alpha$ -L-Fukosyltransferase (H-Transferase;  $\alpha(1,2)$ Fukosyltransferase; E.C. 2.4.1.69). Die cDNA kodiert für ein Protein mit 365 Aminosäuren. Die genomische Struktur des *FUT1*-Gens besteht aus zwei Exons [58,59,67,97]. Das einzige Intron ist 5' zum Startkodon lokalisiert [58]. Die gesamte CDS findet sich deshalb im Exon 2. Als Spleißvariante wurde in Malignomen auch eine CDS beschrieben, die Teile des Exons 1 umfaßt [60]. Die chromosomale Position ist 19q13.3 [97].

Die Funktion des *FUT1*-Gens wurde durch Expression in COS-1 Zellen bewiesen [67]. Die transfizierten Zellen exprimierten H-Antigen an ihrer Oberfläche. Eine  $\alpha$ -2-L-Fukosyltransferase-Aktivität analog der humanen H-Transferase konnte ebenfalls nachgewiesen werden. In zwei Individuen mit „para Bombay“-Phänotyp wurden nicht-funktionale Allele des *FUT1*-Gens zusammen mit funktionalen *Se*-Allelen des *FUT2*-Gens gesichert [59].

Die Klonierung des *FUT2*-Gens belegte die hohe Homologie zum *FUT1*-Gen [59] und die benachbarten chromosomalen Positionen mit lediglich zirka 80 kb physischer Distanz [97], die genetisch vorher auf 1 cM geschätzt wurde [91]. Beide *FUT*-Gene werden deshalb praktisch immer gemeinsam vererbt („linkage disequilibrium“) und sind wegen ihrer hohen Homologie sicher durch eine

Genduplikation entstanden. Diese Ergebnisse werden als formaler Beweis des genetischen Modells von Oriol et al. [91] betrachtet.

Die genauere Charakterisierung der nicht-funktionalen Allele im *FUT1*-Gen ist Gegenstand dieser Arbeit und wird im Abschnitt 4.2 diskutiert. Die *h*-Allele gelten in der weißen Bevölkerung als sehr selten, während die Frequenz der *se*-Allele in der deutschen Bevölkerung 0,4516 (zirka 45% aller Allele des *FUT2*-Gens) beträgt [83]. Beim *FUT2*-Gen ist ein unter Europäern häufiges Null-Allel mit einer Stopkodon-Mutation nachgewiesen worden [59]. Bei der beginnenden Charakterisierung der Null-Allele in anderen Bevölkerungen zeigt sich bereits, daß dort mit anderen Null-Allelen und einem erheblichen Polymorphismus zu rechnen ist [61,127].

Null-Allele (nicht-funktionale Allele) der *FUT1*- beziehungsweise *FUT2*-Gene werden als neutral betrachtet. Individuen, die für Null-Allele an einem oder beiden Genloci homozygot sind, haben keine gesundheitlichen Nachteile, sind fertil und weisen (im englischen Sprachgebrauch) eine normale „fitness“ auf.

Grundsätzlich könnten für den Bombay ( $O_h$ )-Phänotyp auch andere Ursachen als nicht-funktionale Allele der *FUT1*- und *FUT2*-Gene in Betracht kommen. Bisher wurde jedoch nur ein einziger Proband mit Bombay ( $O_h$ )-Phänotyp beobachtet, der eine funktionale  $\alpha$ -2-L-Fukosyltransferase im Blutplasma aufwies [102]. Die Autoren vermuteten, daß primär der Fukose-Metabolismus gestört war und deswegen die funktionalen Fukosyltransferasen das H-Antigen nicht bilden konnten. Formal wurde bei diesem Probanden jedoch ein Defekt der Grundsubstanz nicht ausgeschlossen. Die weitere Analyse des neu beschriebenen Rambam-Hasharon-Syndroms [40,41], auch als „Leukocyte adhesion deficiency type 2“ bekannt, wird hier Klärung bringen.

### 1.2.3. Interaktion zwischen verschiedenen Glykosyltransferasen

Endständige Zucker werden an die Grundsubstanz konjugiert und bilden - unter anderem - die Antigene der Blutgruppen H, ABO und Lewis. Hierbei wirken die Glykosyltransferasen der verschiedenen Genloci zusammen (Tabelle 1.2). Es wurde geschätzt, daß mindestens 100 Glykosyltransferasen für die Synthese aller menschlichen Kohlenhydratketten erforderlich sind. Glykosyltransferasen sind Membranständige Proteine des endoplasmatischen Retikulums und des Golgi-Apparats. Sie können auch gelöst in Blutplasma und in Sekreten nachgewiesen werden. Die Synthese der Zuckerketten wird wesentlich bestimmt durch die Interaktion mit den Substraten beziehungsweise die Konkurrenz um die Substrate der verschiedenen

Glykosyltransferasen. Der Genotyp (Allele für Enzyme unterschiedlicher Aktivität, nicht-funktionale Allele) hat ebenso Einfluß auf die Ausbildung der Zuckerketten wie das Expressionsmuster der Genprodukte in den verschiedenen Gewebetypen.

**Tabelle 1.2. Gene einiger Glykosyltransferasen für Blutgruppenantigene**

Gen	Synonym	chromosomale Position	Allele	Transferase
<i>FUT1</i>	<i>H</i>	19q13.3	<i>H</i> <i>h</i>	$\alpha$ -2-L-Fukosyltransferase nicht-funktionale Allele
<i>FUT2</i>	<i>SE</i>	19q13.3	<i>Se</i> <i>se</i>	$\alpha$ -2-L-Fukosyltransferase nicht-funktionale Allele
<i>FUT3</i>	<i>LE</i>	19p13.3	<i>Le</i> <i>le</i>	$\alpha$ -3/4-L-Fukosyltransferase nicht-funktionale Allele
<i>ABO</i>	-	9q34	<i>A</i> <i>B</i> <i>O</i>	$\alpha$ -3-N-Acetyl-D-Galaktosaminyltransferase a-3-D-Galaktosyltransferase nicht-funktionale Allele

Weitere Fukosyltransferase-Gene (*FUT5* und *FUT6*) kodieren für  $\alpha$ -3-L-Fukosyltransferasen. Ihre chromosomalen Positionen befinden sich zusammen mit dem *LE*-Gen in einem „Gencluster“ auf dem Chromosom 19p13.3 [94].

Die Synthese des H-Antigens durch die H-Transferase (*FUT1*-Gen) ist Voraussetzung dafür, daß die funktionalen A- und B-Transferasen des *ABO*-Gens ihr Substrat (das H-Antigen) zu Antigen A beziehungsweise Antigen B umwandeln können. Weiterhin beeinflusst die H-Transferase des *FUT2*-Gens die Expression der Lewis-Antigene Le(a) und Le(b). Liegt eine funktionale Le-Transferase (*FUT3*-Gen) vor, bestimmt der „Se-Status“ die Spezifität des Lewis-Antigens. Bei Genotyp *Le/- Se/-* wird das Antigen Le(b) gebildet, beim Genotyp *Le/- sese* das Antigen Le(a). Ist der Proband homozygot für nicht-funktionale Allele im *FUT3*-Gen, Genotyp *lele*, tritt der erythrozytäre Phänotyp Le(a- b-) unabhängig davon auf, welches Allel des *FUT2*-Gens vorliegt.

Auch das „Gencluster“ auf 19p13.3 könnte das Ergebnis von Genduplikationen sein [87,94]. Die  $\alpha$ -3-L-Fukosyltransferasen unterscheiden sich - ähnlich den H- und Se-Transferasen - weniger durch ihre Zuckerspezifitäten als durch die geringfügig variablen Expressionsmuster in den Geweben (zum Beispiel *FUT4* myeloisch und *FUT6* plasmatisch) [80]. Auch hier interagieren die verschiedenen Glykosyltransferasen. Zum Beispiel kann die Expression des CA19-9 „Tumor“-Antigens nur erfolgen, wenn ein funktionales *FUT3*-Allel vorliegt [48,78]. Bei Patienten mit dem

Phänotyp Le(a- b-) hat die Untersuchung auf CA19-9 Antigen deshalb keine Aussagekraft, weil das „Tumor“-Antigen nicht effektiv gebildet werden kann [62].

### 1.3 Biologische Funktion von Blutgruppenantigenen

Erst in den letzten Jahren wurden zur Charakterisierung der Blutgruppenantigene der Erythrozytenmembran molekulare Methoden eingesetzt, mit denen bis auf die P- und die Dombrock-Blutgruppe inzwischen alle Gene für Blutgruppensysteme im Sinne der aktuellen Definition [18] molekular kloniert wurden. Hierbei ergaben sich unerwartete Einblicke in die Biologie der Erythrozytenmembran. Einige dieser überraschenden Querverbindungen zu bereits bekannten Proteinfamilien eröffneten neue Ansätze zum Verständnis der Mechanik, Membranphysiologie und immunologischen Kompetenz von Erythrozyten. Eine ganze Reihe von Übersichtsarbeiten wurde kürzlich veröffentlicht [2,42,82,84,107,108], die auch aktuelle Aufstellungen zu Proteinfamilien (siehe Tabelle 1.3) und Rezeptoren (siehe Tabellen 1.4 und 1.5) enthalten [84,108].

#### 1.3.1 Proteinantigene

Aufgrund der Möglichkeit zur direkten Charakterisierung der molekularen Struktur über die cDNA sind die Proteinantigene bisher am besten untersucht (Tabelle 1.3). Einige Blutgruppenantigene werden von Proteinen getragen, die wesentliche strukturelle Funktionen haben (Bande 3-Protein, Glykophorine). Andere antigentragende Proteine stellen Transportkanäle in der Erythrozytenmembran dar (ebenfalls Bande 3-Protein, Aquaporin, Harnstofftransporter).

Die Beschreibung von Null-Phänotypen, die krankheitsassoziiert auftreten, kann richtungsweisend sein, um die Proteinfunktion zu charakterisieren. So konnten die  $K_0$ -Phänotypen zur molekularen Klonierung des  $K_x$ -Gens und Charakterisierung des McLeod-Syndroms herangezogen werden. Bei dem Syndrom wurde sowohl eine 50 kb umfassende Deletion als auch eine Punktmutation in der CDS des  $K_x$ -Gens als Ursache nachgewiesen. Ein aberranter Phänotyp der Colton-Blutgruppe wurde assoziiert mit chronischer Niereninsuffizienz beobachtet [73]. Der Umstand, daß das Colton-Antigen auf dem Wasserkanal-Protein der Nierenepithelien vorliegt, suggeriert einen ursächlichen Zusammenhang und eröffnet Möglichkeiten zur molekularen Analyse in betroffenen Familien.

**Tabelle 1.3. Blutgruppenantigene, Multiproteinfamilien und Funktionen**

Antigen(e)	Protein(e)	Proteinfamilie	Funktion(en)
Bg	HLA	HLA Klasse I	-
Colton	CHIP (Aquaporin-1)	Wasserkanal-Protein	Transporter
Cromer	Decay accelerating factor	Komplement-Regulationsprotein, GPI-verknüpft	Schutz vor Komplement-vermittelter Hämolyse
Diego, Wright	Bande 3-Protein	Anionen-Transporter	Transporter, Membranskelett
Duffy	Duffy	Chemokin-Rezeptor	Interleukin 8-Rezeptor
Indian	CD44	Cartilage link proteins	Membranskelett
K	Kell	cALL-Antigen-Familie	neutrale Endopeptidase
Kidd	Kidd	-	Harnstofftransporter
Knops/McCoy	Complement receptor type 1 (CR1, CD35)	Komplement, Regulationsprotein	Schutz vor Komplement-vermittelter Hämolyse
Lu	Lutheran	Immunglobulin-Superfamilie	-
LW	LW protein	Immunglobulin-Superfamilie	-
MN	Glykophorin A	Glykophorine	-
Rhesus	RhD, RhCE	-	Transporter?
Ss	Glykophorin B	Glykophorine	-
Xg(a)	Xg	CD99-Homolog	Adhäsion?
Yt	Acetylcholinesterase	-	Enzym

Verschiedene Blutgruppenantigene werden von Proteinen determiniert, die als Rezeptoren für Pathogene beschrieben sind (Tabelle 1.4). Ein bekanntes Beispiel ist das Duffy-Blutgruppenantigen, dessen Gen kürzlich kloniert wurde [50] und erythrozytärer Rezeptor für einen Malariaerreger ist. Null-Phänotypen dieser Blutgruppe [Fy(a-b-)] sind unter Europäern eine Rarität, treten aber - wegen dem vorhandenen Selektionsdruck - sehr häufig in Äquatorial-Afrika auf. Interessanterweise unterscheiden sich die molekularen Ursachen der Null-Phänotypen dieser beiden Populationen. Europäer tragen Null-Allele [76], die vermutlich sporadisch sind und eine polymorphe Ursachen haben. Afrikaner besitzen funktionale Allele, die in den verschiedensten Geweben exprimiert werden. Lediglich in hämatopoetischen Geweben wird kein Antigen exprimiert, da dort der Promotor nicht aktiv ist und deshalb kein Transkript gebildet wird [109].

**Tabelle 1.4. Blutgruppenantigene (Proteine) als Rezeptoren für Pathogene**

Antigen	Protein	Mikroorganismus
MN	Glykophorin A	<i>Plasmodium falciparum</i> , <i>Escherichia coli</i>
Dr <sup>a</sup>	Decay accelerating factor, CD55	<i>Escherichia coli</i>
Cromer	Decay accelerating factor, CD55	Echovirus
Indian	CD44	Poliovirus
AnWj	CD44	<i>Haemophilus influenzae</i>
Duffy	Duffy	<i>Plasmodium vivax</i> , <i>Plasmodium knowlesi</i>

### 1.3.2 Zuckerantigene

Auch einige Blutgruppenantigene (Tabelle 1.5), die durch Zuckerantigene determiniert werden, sind als Rezeptoren für virale und bakterielle Pathogene beschrieben worden [95]. Meist liegen in vitro-Untersuchungen zugrunde. Es wäre interessant, den relativen Einfluß dieser Antigene auf Krankheitsentstehung und -verlauf im Rahmen einer multifaktoriellen Prädisposition zu bestimmen. So ist beispielsweise das P-Antigen in der „ISBT 209 GLOBO collection of antigens“ als Rezeptor des Parvovirus B19 bekannt, das an der Entstehung aplastischer Anämien beteiligt ist [39,64].

**Tabelle 1.5. Blutgruppenantigene (Zucker) als Rezeptoren für Pathogene**

Antigen	Mikroorganismus
P1	<i>Escherichia coli</i>
Lewis: Le(b)	<i>Helicobacter pylori</i>
H, A	<i>Candida albicans</i>
P	Parvovirus B19, <i>Escherichia coli</i>

Der Grundsubstanz („Precursor substance“, an deren Ende das H- beziehungsweise die A- und B-Antigene synthetisiert werden) wird seit langem große Bedeutung für die „fitness“ und somit für biologische Funktionen wie Zell- und Gewebedifferenzierung beigemessen. Indirekt wird dies durch den fehlenden Nachweis eines Individuums belegt, dem ein Teil der Grundsubstanz (Null-Phänotyp) fehlt. Analog dem Bombay (O<sub>h</sub>)-Phänotyp müßte dieser Phänotyp über die Bildung eines „Isoagglutinins“ gegen die fehlenden Zuckerepitope auffallen. Trotz millionenfacher Untersuchung von Patienten und Blutspendern ist jedoch weltweit noch nie ein solcher Phänotyp beschrieben worden. Dennoch könnte ein solcher Phänotyp eines Tages beobachtet werden. Es wäre deshalb wichtig, daß bei verdächtigen Phänotypen [102] auch formal die Möglichkeit einer defekten Grundsubstanz ausgeschlossen wird.

Die Bedeutung der frühen Vorstufen der Grundsubstanz für die Zellentwicklung wurde neuerdings experimentell in Pflanzen [45] und durch „knock out“-Mäuse auch bei Vertebraten gesichert [53,79]. In diesen Experimenten wurde die Embryonalentwicklung von Mäusen untersucht, die einen homozygoten Defekt des Gens der N-acetylglukosaminyl-Transferase I (GlcNAc-TI) aufweisen. Die Entwicklung verlief weit über die Einnistung der Eizelle hinaus normal. Gegen Ende der Embryonal-

entwicklung kam es jedoch zu Entwicklungsstörungen im Neuronalgewebe und bei der Ausbildung der Rechts-Links-Asymmetrie, die immer mit einem Abort um Tag 9 der Trächtigkeit endeten.

Antigene, die durch eine bestimmte Fukosyltransferase (Fuc-TVII) gebildet werden, wurden als Rezeptoren für L-, E- und P-Selektine bei Leukozyten gesichert [77]. Dadurch bestimmt das Expressionsmuster dieses *FUT*-Gens die Gewebespezifität des entsprechenden Rezeptors und in der Folge auch die Verteilung („trafficking“) von Leukozyten. Zweifellos werden durch die genauere Charakterisierung von Glykosyltransferasen viele weitere wichtige biologische Funktionen von Zuckerepitopen geklärt, die zumindest zum Teil auch als erythrozytäre Antigene vorhanden sind.

#### 1.4 Fragestellung: *FUT1*-Gen als Modellsystem für nicht-funktionale Allele

Ziel der Untersuchung war es, bei einem menschlichen Genlocus unter niedrigem Selektionsdruck die Frequenz von Null-Allelen mit Mittelwert und Vertrauensintervall statistisch zu bestimmen. Dabei sollte ein möglicher Polymorphismus der vergleichsweise seltenen Null-Allele im H-Blutgruppensystem bestimmt werden. Der populationsgenetische Ansatz erlaubte auch eine Aussage zur Konsanguinität in der deutschen Bevölkerung aus individuellen molekularen Daten ohne Kenntnis von Abstammungsverhältnissen.

In bezug auf die Frequenzbestimmung von sporadischen Null-Allelen ist das H-Blutgruppensystem (ISBT Nummer 018, Gennamen *FUT1*) besonders geeignet. Vor allem liegen für Blutgruppensysteme große Datenbasen für viele Bevölkerungen vor, und die vorgeschriebenen serologischen Routinemethoden zum Beispiel bei Blutspendern weisen den Bombay ( $O_h$ )-Phänotyp sicher nach. Im H-Blutgruppensystem gibt es in Europa keinen prävalenten Null-Phänotyp. Weiterhin ist kein Selektionsdruck an diesem Genlocus bekannt. Methodologisch kam uns zugute, daß die gesamte CDS in einem Exon auftritt und mithin technisch vorteilhaft über eine PCR mit genomischer DNA erfaßt werden kann.

Die Häufigkeit von sporadischen Null-Allelen in Genen unter niedrigem oder fehlendem Selektionsdruck kann daher sehr effizient abgeleitet werden, indem man die seltenen Individuen mit der Bombay ( $O_h$ )-Blutgruppe untersucht. Dieser Ansatz nutzt die umfangreichen blutgruppenserologischen Datenbasen, die für große Populationen verfügbar sind, und erlaubt die effiziente Ableitung von molekular-genetischen Daten, die für neutrale Mutationen in vielen Genen Bedeutung haben.



## 2. METHODEN

### 2.1 Identifikation von Individuen mit Bombay ( $O_h$ )-Blutgruppe in der Bevölkerung

#### 2.1.1 Blutgruppen-Bestimmung

Die AB0-Blutgruppen wurden bestimmt mit einem Olympus PK7100 Autoanalyser (Olympus, Hamburg). Entsprechend den während des Untersuchungszeitraums gültigen Richtlinien zur Blutgruppenbestimmung und Bluttransfusion [120] wurden mindestens zwei Anti-A und Anti-B Testseren sowie ein Anti-AB Testserum bei jeder Blutgruppenbestimmung verwendet. Hierbei wurden ausschließlich staatlich zugelassene und kommerziell erhältliche Antiseren eingesetzt. Alle untersuchten Personen wurden auf antierythrozytäre Antikörper untersucht. Bei Individuen mit der Bombay ( $O_h$ )-Blutgruppe liegt obligat ein Anti-H Antikörper vor, der in den eingesetzten Standardtechniken des Antikörpersuchtestes sicher nachweisbar ist [54,101].

#### 2.1.2 Blutspender-Datenbasis

Der DRK-Blutspendedienst Baden-Württemberg unterhält seit 1985 eine zentrale elektronische Datenbasis für die blutgruppenserologischen Daten aller Blutspender. Die Angaben zur Blutgruppenfrequenz wurden abgeleitet von allen gespeicherten Blutspendern einschließlich der Patienten, bei denen eine autologe Blutbereitstellung durchgeführt wurde. Der Anteil autologer Blutbereitstellungen macht nach wie vor einen kleinen Anteil aus ( $< 0,5\%$ ). Personen, die nach Baden-Württemberg eingewandert sind, wurden nicht von der Analyse ausgeschlossen. Die abgeleiteten Frequenzen sind deswegen repräsentativ für die Bevölkerung, die im Untersuchungszeitraum 1985 bis 1993 in Baden-Württemberg lebte. Blutspender, die seltene Phänotypen aufwiesen - hier insbesondere Blutspender mit der Bombay ( $O_h$ )-Blutgruppe - wurden gezielt gefragt, ob er beziehungsweise sie von der seltenen Blutgruppe bereits vor ihren ersten Blutspenden Kenntnis hatten.

### 2.1.3 Statistik

Die Frequenzen der seltenen Blutgruppen-Phänotypen wurden kalkuliert unter Berücksichtigung solcher Individuen, deren seltener Phänotyp anlässlich der ersten Blutspende zufällig gefunden wurde, ohne daß eine Vorkenntnis bestand. Für die Kalkulation der minimalen Frequenz wurden allerdings alle bekannten Individuen mit den seltenen Phänotypen in Baden-Württemberg berücksichtigt, wenn dadurch die untere Grenze des 95% Vertrauensintervalls überschritten wurde.

Im Februar 1994 betrug die Bevölkerungszahl in Baden-Württemberg 10.234.252 Personen (Statistisches Landesamt Baden-Württemberg, Stuttgart). Wegen Geburt, Tod und Wanderungsbewegungen stellt dies lediglich eine minimale Abschätzung der effektiv getesteten Bevölkerungszahl dar. Unsere Stichprobengröße von 624.164 Blutspendern kann dennoch mehr als 5% der effektiven Bevölkerungsgröße repräsentieren. Deshalb kalkulierten wir die Vertrauensintervalle unter der Annahme einer hypergeometrischen Verteilung mit einer minimalen Populationsgröße von 10.200.000 und zusätzlich unter der Annahme einer Poisson-Verteilung, die eine unbegrenzte Populationsgröße voraussetzt; die konservativere Abschätzung des Vertrauensintervalls wurde berücksichtigt (siehe Tabelle 3.1). Deshalb können die angegebenen Vertrauensintervalle nicht zu klein sein für jede effektive Bevölkerungsgröße von mehr als 10.200.000 Personen.

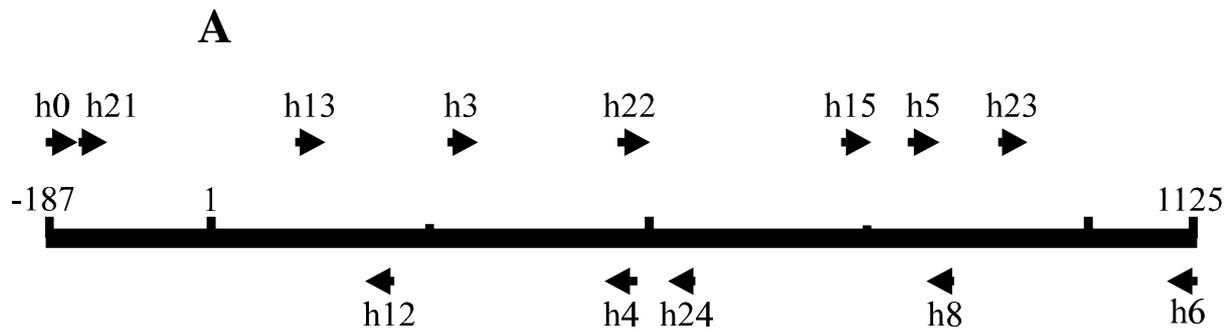
## 2.2 Probanden

Im Rahmen der durchgeführten systematischen Bevölkerungsuntersuchung wurden zwei Individuen mit der Bombay ( $O_h$ )-Blutgruppe (Probanden 1 und 2) unter insgesamt 624.164 Blutspendern gefunden. Drei weitere Probanden (Proband 3, 4 und 5) mit den Phänotypen H-defizienter Sekretoren und Bombay ( $O_h$ ) wurden uns zugewiesen im Rahmen der transfusionsmedizinischen Versorgung mit tiefgefrorenen Erythrozytenpräparate der Blutgruppe Bombay ( $O_h$ ). Die Probanden waren nicht miteinander verwandt. Proband 1, 2, 3 und 5 waren deutscher, Proband 4 war türkischer Abstammung. In den Familien der Probanden deutscher Abstammung waren keine konsanguinen Vorfahren bekannt. Für Proband 4 war diese Frage nicht abschließend zu klären.

### 2.3 Molekularbiologische Untersuchungen

Die volle Länge der kodierenden Sequenz (CDS) des *FUT1*-Gens wurde amplifiziert mit dem Primerpaar h0 und h6 (Abbildung 2.1). Wir führten die Nukleotidsequenzierung mit Hilfe eines DNA-Sequenzierungsgeräts ABI 373A durch. Hierbei wurden kommerziell erhältliche Sequenziermethoden verwendet (Prism Dye Terminator Sequenase Kit, Prism Dye Terminator Cycle Sequencing Kit mit AmpliTaq DNA-Polymerase und Prism Dye Terminator Sequencing Kit mit AmpliTaq FS DNA-Polymerase; Applied Biosystems, Weiterstadt). Die meisten PCR-Produkte wurden in einen Plasmidvektor subkloniert (pMosBlueT-Vektor Kit, United States Biochemicals, Cleveland OH, USA). Die PCR-Produkte wurden sequenziert (Abbildung 2.1). Alle nachgewiesenen Abweichungen von der veröffentlichten *H*-Sequenz (Nukleotid Sequenz Datenbanken Zugriffsnummer M35531) wurden bestätigt durch die direkte Nukleotidsequenzierung von PCR-Produkten, die in unabhängigen PCR-Reaktionen amplifiziert wurden (Prism Dye Terminator Cycle Sequencing Kit).

Zusätzlich amplifizierten wir die DNA von allen Probanden mit dem Primerpaar h6 und h10 („sense primer“ in dem nicht-translatierten Exon 1: GCTCCGGCC ACGAAAAGC). Die PCR-Produkte dieses Primerpaars überspannen die volle Länge des Exons 2 und des einzigen Introns sowie eines Teils des Exons 1. Wir konnten die PCR-Produkte der erwarteten Größe bei allen Probanden nachweisen, was bei den Probanden 3 und 4 die Möglichkeit einer größeren Deletion im Intron ausschloß (Daten nicht gezeigt). PCR-Produkte, die in den Bereichen der mutierten Nukleotide positioniert waren, wurden sequenziert in fünf unabhängigen Klonen für den Probanden 3 und neun unabhängigen Klonen für den Probanden 4, die jeweils von wenigstens drei unabhängigen PCR-Reaktionen stammten; in allen Klonen konnten die für  $h^3$  und  $h^4$  typischen Nukleotidmutationen nachgewiesen werden. Um die Möglichkeit von alternativen Spleißstellen im Allel  $h^7$  auszuschließen, subklonierten wir PCR-Produkte, die das gesamte Intron umfaßten (Primerpaar h6 und h10). Die Sequenzierung wurde mit Primern in der Vektor-DNA durchgeführt. Wir konnten die Integrität der bekannten Spleißstellen nachweisen (Daten nicht gezeigt).



**B**

Sense primers		
h0	CTCATCCCTGAAACCTAGGACTC	-187 bis -165
h21	TCAGCCCCAGGAGTCTAAC	-150 bis -131
h13	CCACATGGCCTAGGCCTGTCC	97 bis 117
h3	TAATCAGATGGGACAGTATGCC	270 bis 291
h22	CGGACTTGAGAGATCCTTTCTGA	464 bis 487
h15	CCTACCTCCGGCAGGCCATGG	719 bis 739
h5	GGAGTGGTGTAAGAAAACATCG	795 bis 817
h23	TGCAACCACACCATTATGAC	898 bis 917
Antisense primers		
h6	GGCTACTTCAGAAAGTCTCCC	1125 bis 1105
h8	CCAGCAAACGTCACATCG	848 bis 831
h24	ACTCTCTGCGGATCTGTTC	553 bis 536
h4	TCAGGAAAAGGATCTCTCAAGTCCG	487 bis 464
h12	ACAGGAAGAGGAGGCGTTGG	210 bis 191

**Abbildung 2.1. Sequenzierungsstrategie und Primer.** Im Teil A der Abbildung ist das *FUT1*-Gen mit dem 3' Bereich des einzigen Introns und dem Exon 2 dargestellt, das die gesamte kodierende Sequenz (CDS) umfaßt. Die Position 1 stellt das erste Nukleotid des Startkodons dar. Das Stopkodon liegt an den Nukleotidpositionen 1096 - 1098. Die Positionen, Kennzeichnungen und Richtungen der für die Sequenzierung verwendeten Primer sind eingezeichnet. Die maximale erforderliche Leselänge beträgt 250 Nukleotide in der Vorwärtsleserichtung und 300 Nukleotide in der reversen Leserichtung der kodierenden Sequenz. Im Teil B sind die Bezeichnungen der verwendeten synthetischen Oligonukleotidprimer (ODN) angegeben zusammen mit ihren Nukleotidsequenzen und ihren Positionen im Verhältnis zur ersten Nukleotidposition des Startkodons ATG in der veröffentlichten *H*-Sequenz [67].

## 2.4 Populationsgenetische Analyse

Der „inbreeding coefficient  $F$ “ wurde ursprünglich definiert von Wright als relatives Maß für die Inzucht in einer Abstammungslinie [123]. Die durchschnittliche Konsanguinität innerhalb einer Population wird mit dem mittleren Konsanguinitätskoeffizienten  $\alpha$  (englisch: average inbreeding factor  $\alpha$ ) beschrieben und stellt den Durchschnitt aller  $F$  der Mitglieder der Population dar. Das Verhältnis von Probanden mit einem homozygoten  $h$ -Allel zu Probanden mit zwei unterschiedlichen  $h$ -Allelen innerhalb einer Population wird durch den mittleren Konsanguinitätskoeffizienten  $\alpha$  bestimmt. Ein geeignetes Gleichungssystem (Formel 1.1 und 1.2) zur Kalkulation des mittleren Konsanguinitätskoeffizienten  $\alpha$  und der Frequenz von  $h$ -Allelen,  $f_h$ , in der Bevölkerung findet sich bei Cavalli-Sforza und Bodmer [12].

Die Summe der Frequenzen aller verschiedenen Haplotypen  $hse$ ,  $f_{hse}$ , kann mathematisch wie folgt dargestellt werden:

$$f_{hse} = \sum_{i=1}^n q_i ,$$

wobei  $q_i$  die Frequenz eines einzelnen „privaten“ Haplotyps  $hse$  bezeichnet und

$$\sum_{i=1}^n q_i$$

die Summe der Frequenzen aller unterschiedlichen  $hse$ -Haplotypen, die in der Population auftreten. Das folgende Gleichungssystem aus einem Lehrbuch der Populationsgenetik wurde angewandt [12]:

$$\text{(Formel 1.1.) } f_{\text{homo+hetero}} = (1 - a) \left( \sum_{i=1}^n q_i \right)^2 + a \sum_{i=1}^n q_i$$

$$\text{(Formel 1.2.) } f_{\text{homo}} = (1 - a) \sum_{i=1}^n (q_i)^2 + a \sum_{i=1}^n q_i$$

In dem Gleichungssystem bezeichnet  $f_{\text{homo+hetero}}$  die Frequenz des Genotyps  $hh$  *sese* mit identischen und unterschiedlichen  $h$ -Allelen.  $f_{\text{homo}}$  bezeichnet die Frequenz des Genotyps  $hh$  *sese* mit identischen  $h$ -Allelen, d. h. einen Bombay ( $0_h$ )-Genotyp, der durch die Homozygotie für ein privates  $h$ -Allel bedingt ist. In unseren Experimenten bestimmten wir  $f_{\text{homo+hetero}}$  sowie  $f_{\text{homo}}$  experimentell. Die Kalkulationen

wurden zunächst durchgeführt unter der vereinfachenden Annahme, daß alle  $h$ -Allele mit der gleichen Frequenz vorkommen, d. h. es gelte

$$q_1 = q_2 = \dots = q_n ,$$

wobei  $n$  die Zahl der unterschiedlichen  $h$ -Allele bezeichnet, die in der Population vorkommen. Die Lösungen des Gleichungssystems (Formeln 1.1 und 1.2) bestimmten die zwei Variablen  $\alpha$  und  $f_{hse}$ . Unter den genannten Voraussetzungen gilt

$$\sum (q_i)^2 = n \left( \frac{\sum_{i=1}^n q_i}{n} \right)^2 = \frac{\left( \sum_{i=1}^n q_i \right)^2}{n} . \text{ Damit läßt sich Formel 1.2 umformen}$$

zu (Formel 1.2a)  $f_{\text{homo}} = (1-a) \frac{\left( \sum_{i=1}^n q_i \right)^2}{n} + a \sum_{i=1}^n q_i$  . Danach wurden die Terme der

Formel 1.2a von den Termen in Formel 1.1 subtrahiert (Formel 1.1 - Formel 1.2a), wodurch der Term

$$a \sum_{i=1}^n q_i$$

aus dem Gleichungssystem eliminiert wird:

$$f_{\text{homo+hetero}} - f_{\text{homo}} = (1-a) \left( \sum_{i=1}^n q_i \right)^2 - (1-a) \frac{\left( \sum_{i=1}^n q_i \right)^2}{n}$$

Die weitere algebraische Umformung liefert

$$\text{(Formel 1.3)} \quad \sum_{i=1}^n q_i = \sqrt{\frac{n}{n-1} \cdot \frac{f_{\text{homo+hetero}} - f_{\text{homo}}}{1-a}} ,$$

wobei das Ergebnis der Quadratwurzel - ein bekanntes  $\alpha$  vorausgesetzt - unmittelbar das gesuchte  $f_{hse}$  liefert. Um  $\alpha$  zu bestimmen, wurde  $f_{hse}$  in Formel 1.1 durch die Quadratwurzel der Formel 1.3 substituiert. Die weitere Umformung liefert schließlich eine quadratische Gleichung der Form

$$\alpha^2 a + \alpha b + c = 0$$

mit den bekannten Lösungen für  $\alpha$

$$\text{(Formel 1.4)} \quad a = \frac{1}{2a} \left( -b \pm \sqrt{b^2 - 4ac} \right)$$

wobei 
$$a = \frac{n}{n-1} [f_{\text{homo+hetero}} - f_{\text{homo}}] \quad ; \quad b = \left( f_{\text{homo+hetero}} - \frac{n}{n-1} [f_{\text{homo+hetero}} - f_{\text{homo}}] \right)^2$$

und 
$$c = - \left( f_{\text{homo+hetero}} - \frac{n}{n-1} [f_{\text{homo+hetero}} - f_{\text{homo}}] \right)^2 \text{ gilt.}$$

Mit der Formel 1.4 kann der mittlere Konsanguinitätskoeffizient  $\alpha$  anhand der experimentellen Daten und einem vorgegebenen  $n$  berechnet werden. Ist  $\alpha$  aus der Formel 1.4 bekannt, läßt sich  $f_{hse}$  entsprechend Formel 1.3 für vorgegebene  $n$  einfach bestimmen.

Wir berechneten  $f_{hse}$  mit  $n = 19$ , was die kleinste statistisch anzunehmende Anzahl von  $h$ -Allelen in der Population darstellt, wie in den Ergebnissen abgeleitet wurde. Um unsere Ergebnisse abzusichern, führten wir die Kalkulationen auch unter der wahrscheinlicheren Annahme von  $n = 100$  sowie  $n = 1.000.000$  durch. Weiterhin wurden Kalkulationen für variable Anteile von unterschiedlichen  $h$ -Allelen durchgeführt, zum Beispiel mit

$$q_1 = 0.6 f_{hse} \quad \text{und} \quad \sum_{i=2}^n q_i = 0.4 f_{hse} \quad .$$

Diese Berechnungen zeigten, daß die Abschätzung der Frequenz des Haplotyps  $hse$ ,  $f_{hse}$ , sich um weniger als 18% ändert, wenn variable  $n$  und variable Anteile von unterschiedlichen  $h$ -Allelen angenommen werden (Daten nicht gezeigt).

Der Bombay ( $0_h$ )-Phänotyp ist bedingt durch den Genotyp  $hh\ sese$  (siehe Abschnitt 1.2.1). Wir bestimmten die Frequenz der  $h$ -Allele,  $f_h$ , in der Bevölkerung, indem wir das Ergebnis  $f_{hse}$  des Gleichungssystems durch die bekannte Frequenz der  $se$ -Allele,  $f_{se}$ , von 0,4516 in Deutschland [83] dividierten:

$$f_h = f_{hse}/f_{se}$$

Die Kalkulation von  $f_h$  war das primäre Ziel der vorliegenden Arbeit. Das 95% Vertrauensintervall nach der Poisson-Verteilung für den Bombay ( $0_h$ )-Phänotyp in der Population ist 1:88.226 - 1:1.758.205 (Tabelle 3.1). Mit diesen Angaben zum Vertrauensintervall der Bombay ( $0_h$ )-Phänotypfrequenz und  $\alpha = 0,00116$  wurde mit der Formel 1.1 das obere und untere Limit des 95% Vertrauensintervalls für  $f_{hse}$  und daraus folgend für die Frequenz der  $h$ -Allele,  $f_h$ , berechnet.



### 3. ERGEBNISSE

#### 3.1 Frequenzen von seltenen Blutgruppen-Phänotypen

Aufgrund der routinemäßig durchgeführten Untersuchung auf irreguläre anti-erythrozytäre Antikörper bei Blutspendern erwarteten wir, verschiedene sehr seltene Blutgruppen-Phänotypen zu identifizieren, da Individuen mit der Bombay ( $O_h$ )-Blutgruppe und dem p-Phänotyp generell starke natürliche Antikörper (anti-H und anti-Tj(a) = anti-PP<sub>1</sub>P<sup>k</sup>) aufweisen. Bei Individuen mit den ebenfalls sehr seltenen Rhesus-Phänotypen Rh<sub>null</sub>- und D-Deletionen (zum Beispiel D--, D.., DC-, Dc-, Dc<sup>w</sup>-) können bestimmte Rhesus-Antigene fehlen, was zu Diskrepanzen in der routinemäßig durchgeführten C/c- und E/e-Antigenaustestung geführt hätte.

Wir kalkulierten die Frequenzen der seltenen Blutgruppen-Phänotypen und deren 95% Vertrauensintervalle (Tabelle 3.1). Bei der Kalkulation dieser Frequenzen wurden nur jene Blutspender berücksichtigt, deren seltener Blutgruppen-Phänotyp zufällig anlässlich einer Blutspende gefunden wurde. Alle Blutspender, die ihren seltenen Blutgruppen-Phänotyp vor der Erstspende kannten, wurden von der Frequenzauswertung ausgeschlossen. Die Bombay ( $O_h$ )-Blutgruppe tritt nur in Individuen auf, die keine funktionalen *H*- und *SE*-Allele besitzen. Die Kalkulation der Frequenz dieser nicht-funktionalen *h*-Allele ist Gegenstand der vorliegenden Arbeit.

**Tabelle 3.1. Frequenz von seltenen Blutgruppen-Phänotypen<sup>1</sup>**

Phänotyp	Frequenz des Phänotyps		Anzahl der nachgewiesenen Blutspender <sup>3</sup>
	Abschätzung	95% Vertrauensbereich <sup>2</sup>	
Bombay ( $O_h$ )	1:312.081	1:88.226 <sup>4</sup> - 1:1.758.205 <sup>5</sup>	2
Rh <sub>null</sub>	1:552.261	1:101.329 <sup>4</sup> - 1:10.234.252 <sup>6</sup>	1
D-Deletionen <sup>7</sup>	0	1:184.333 <sup>5</sup> - 0 <sup>6</sup>	0 <sup>3</sup>
p	1:5.117.126	1:208.332 <sup>5</sup> - 1:5.117.126 <sup>6</sup>	0 <sup>3</sup>

<sup>1</sup> Die Anzahl der Probanden war  $n = 624.163$  für Bombay ( $O_h$ ) und p sowie  $n = 552.261$  für die seltenen Rhesus-Phänotypen.

<sup>2</sup> Kalkuliert unter der Annahme einer Poisson-Verteilung, einer hypergeometrischen Verteilung oder unter Berücksichtigung der Gesamtzahl von bekannten Individuen (siehe Methoden).

<sup>3</sup> Nachgewiesen im Rahmen der systematischen Untersuchung von Blutspendern. Zusätzlich sind in Baden-Württemberg eine weitere Person mit Bombay ( $O_h$ )-Blutgruppe und zwei Personen mit p-Phänotyp bekannt (Stand Februar 1994).

<sup>4</sup> Hypergeometrische Verteilung.

<sup>5</sup> Poisson-Verteilung.

<sup>6</sup> Abgeleitet von der Gesamtzahl der bekannten Individuen in Baden-Württemberg.

<sup>7</sup> D-Deletionen umfassen D--, D.., DC-, Dc-, DC<sup>w</sup>-.

Unter der - unzulässigen - Annahme, daß die Frequenz der *h*-Allele direkt aus der Frequenz der Bombay ( $O_h$ )-Probanden zu berechnen wäre, ergäbe sich eine *h*-Allelfrequenz von 1:559 [117]. Wie in dieser Arbeit gezeigt wird, kann die Kalkulation nicht unter dieser Annahme vorgenommen werden. Vielmehr muß die Konsanguinität in der Bevölkerung berücksichtigt werden. Bei der Kalkulation der Populationsfrequenz von Allelen der seltenen Phänotypen wie Bombay ( $O_h$ ),  $Rh_{null}$  und *p* hat die Konsanguinität einen erheblichen Einfluß und wird durch den mittleren Konsanguinitätskoeffizienten  $\alpha$  charakterisiert.

### 3.2 Analyse nicht-funktionaler Allele (*h*-Allele) im *FUT1*-Gen

Zwei Probanden mit der Bombay ( $O_h$ )-Blutgruppe wurden bei der systematischen Suche unter Blutspendern gefunden. Drei weitere Probanden mit H-defizientem Blutgruppen-Phänotyp wurden uns zugewiesen. Mit Standardmethoden der Serologie wurde der Phänotyp (Tabelle 3.2) und mittels SSP-PCR der *ABO*- und *SE*-Genotyp (Tabelle 3.3) der Probanden bestimmt. Die Ergebnisse stimmten überein. Erwartungsgemäß konnte bei den Probanden mit Le(a+ b-)-Phänotyp genotypisch ein funktionales *Se*-Allel ausgeschlossen werden (siehe Abschnitt 1.2.3). Probanden 1, 2 und 3 zeigten den typischen Bombay ( $O_h$ )-Phänotyp. Die molekulare *SE*-Bestimmung deutete darauf hin, daß Proband 4 ein funktionales Sekretorgen trägt. Da er phänotypisch Le(a- b-) war, konnte durch die Erythrozytenserologie der Sekretorstatus nicht geklärt werden. Das starke anti-H im Serum des Probanden 4 spricht gegen die Expression des H-Antigens. Der Proband 5 war Le(a- b+) und besaß dementsprechend - wie aus der Genetik der Blutgruppen bekannt - ein funktionales *Se*-Allel. Beim Probanden 5 war das anti-H im Serum vergleichsweise schwach ausgebildet, was auf die Expression der *Se*-

**Tabelle 3.2. Serologische Charakteristika der Probanden<sup>1</sup>**

Proband	Antigene der Erythrozyten					Antikörper im Serum		
	A	B	H	Le(a)	Le(b)	anti-A	anti-B	anti-H
1	-	-	-	++++	-	++++	++++	++++
2	-	-	-	++++	-	++++	++++	++++
3	-	-	-	++++	-	++++	++++	++++
4	-	-	-	-	-	++++	++++	++++
5	-	-	-	-	++++	++++	++++	+

<sup>1</sup> mittels serologischer Standardtechniken bestimmt

**Tabelle 3.3. *AB0/Se*-Genotyp der Probanden\***

Proband	Genotyp	
	<i>AB0</i> -Gen	<i>Se</i> -Gen
1	$O^1/O^1$	<i>se/se</i>
2	$O^1/O^1$	<i>se/se</i>
3	$O^1/O^1$	<i>se/se</i>
4	$O^1/O^1$	<i>Se/Se</i>
5	$A^1/O^1$	<i>Se/se</i>

\* Die Genotypisierung erfolgte mittels Allel-spezifischer Oligonukleotiden (ASO) in einer SSP-PCR Methode (durchgeführt in Kooperation mit Christoph Gassner und Diether Schönitzer, Innsbruck; Methode in Blood 1996 [43] sowie Manuskript in Vorbereitung).

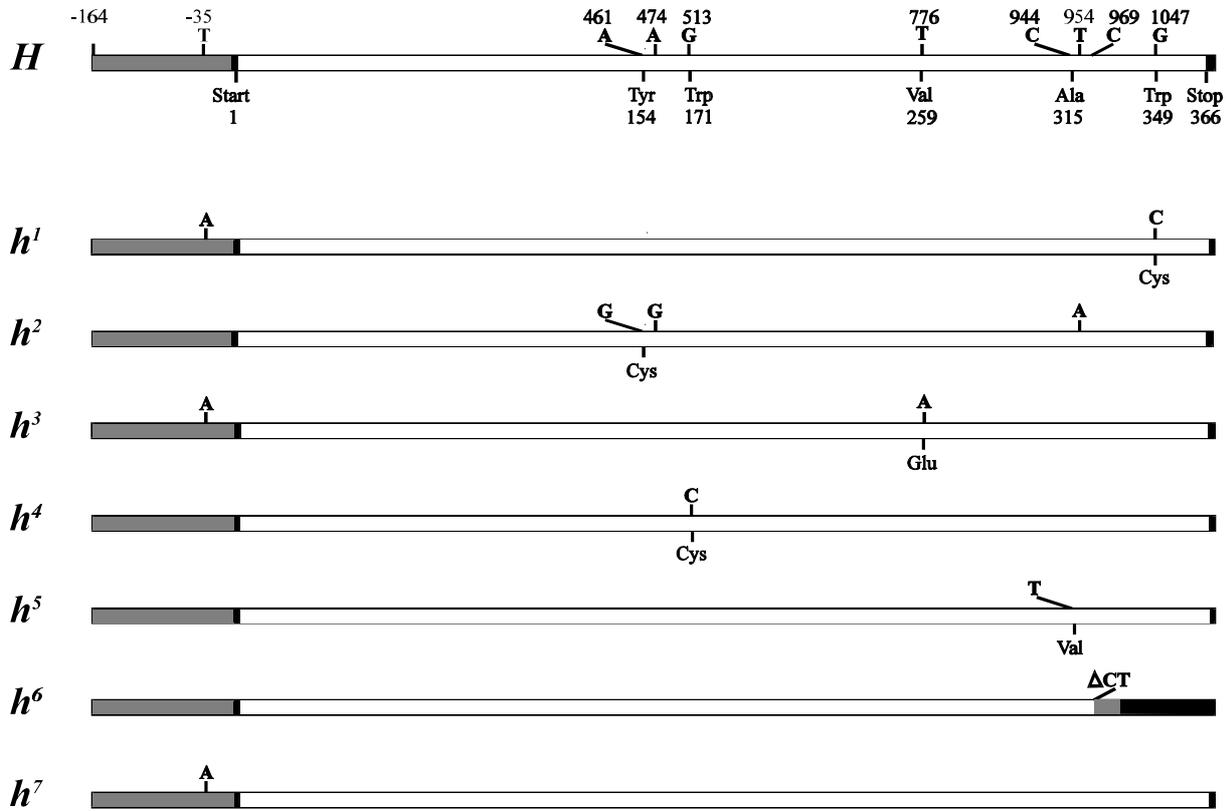
Transferase in nicht hämatopoetischen Geweben zurückzuführen ist. Dies entspricht dem erwarteten Blutgruppen-Phänotyp eines H-defizienten Sekretors.

Die *h*-Allele aller Probanden wurden molekularbiologisch charakterisiert. Wir amplifizierten die gesamte Länge der kodierenden Sequenz (CDS) und einen Teil der 5' nicht-translatierten Sequenz (UTS) im Intron 1 mittels PCR. Wir führten eine Sequenzierung der PCR-Amplifikate nach Subklonierung in einen Plasmidvektor sowie eine direkte Sequenzierung ohne Subklonierungsschritt durch. Die Sequenzierungsstrategie und die verwendeten Primer sind in den Methoden beschrieben (siehe Abbildung 2.1).

### 3.2.1 Mutationen in *h*-Allelen

In den fünf Probanden wurden sieben verschiedene *h*-Allele identifiziert (Abbildung 3.1). Fünf dieser *h*-Allele wiesen Nukleotidsubstitutionen auf, die für nicht-konservative Aminosäureaustausche in der H-Transferase kodierten („missense“-Mutationen). In einem Allel wurde eine Deletion von zwei Nukleotiden nachgewiesen („frame shift“-Mutation), die zu einem vorzeitigen Stopkodon in der kodierenden Sequenz führte. Die CDS des Allels  $h^7$  war identisch zu der *H*-Sequenz. Wir fanden bei diesem Allel auch keine Veränderungen in den Spleißstellen des einzigen Introns im *FUT1*-Gen (Daten nicht gezeigt).

Drei der sieben beobachteten *h*-Allele wiesen eine T → A Substitution im Intron an der Position -35 relativ zum Startkodon auf. Wir sequenzierten diese Region in drei zufällig ausgewählten Blutspendern (Genotyp *HH*) und fanden, daß eine Person homozygot für diese Substitution und eine andere heterozygot war (Daten nicht gezeigt). Hieraus folgt, daß diese Variante häufig in der Bevölkerung



**Abbildung 3.1. Schematische Darstellung der Nukleotid- und der abgeleiteten Aminosäuresequenz für das funktionale *FUT1*-Gen (*H*-Gen) und sieben sporadische nicht-funktionale *h*-Allele.** Nur die Abweichungen vom funktionalen *H*-Gen in der Nukleotid- beziehungsweise Aminosäuresequenz sind angegeben.

- : Intron beziehungsweise Region der abweichenden Aminosäuresequenz im *h*<sup>6</sup>-Allel
- : Exon 2
- : 3' nicht-translatierte Sequenz (UTS)

Die abweichende Aminosäuresequenz im *h*<sup>6</sup>-Allel ist bedingt durch eine „frame shift“-Mutation in den Kodons 323 und 324 (GTC/TAC → GTA/C), die ein vorzeitiges Stopkodon an der Position 333 verursacht. Die Nukleotid- und Aminosäure-Sequenzdaten sind in Datenbanken abgelegt (siehe Abbildung 3.2). Eine vergleichbare schematische Darstellung ist im Internet „online“ verfügbar (siehe Abbildung 3.3).

vorkommt und nicht genetisch mit dem Auftreten von *h*-Allelen verknüpft ist. Alle fünf Probanden dieser Studie trugen unterschiedliche *h*-Allele (Tabelle 3.4), die man ähnlich den „privaten“ Antigenen in der Blutgruppenserologie auch als „private“ Allele bezeichnen könnte.

Weiterhin fanden wir kein *h*-Allel, das identisch gewesen wäre mit den *h*-Allelen [56,58], die zum Zeitpunkt des Abschlusses der Sequenzanalyse bekannt

**Tabelle 3.4. *Hh* Genotypisierung\***

Proband	Genotyp
1	$h^1h^1$
2	$h^2h^7$
3	$h^3h^3$
4	$h^4h^4$
5	$h^5h^6$

\* Die *h*-Allelkennzeichnung bezieht sich auf die Abbildung 3.1. Der *Hh*-Genotyp wurde bestimmt durch die Nukleotid-Sequenzierung der ganzen Länge der kodierenden Sequenz (CDS). Die Bestimmung der CDS wurde in PCR-Produkten mittels internen Primern durchgeführt mit einem automatischen Sequenziergerät. Die Sequenzierungsstrategie ist in Abbildung 2.1 schematisch dargestellt. Die PCR-Amplifikate wurden nach der Subklonierung in einem Plasmidvektor untersucht. Die nachgewiesenen Mutationen wurden durch direkte Sequenzierung (ohne Subklonierungsschritt) von unabhängigen PCR-Produkten bestätigt.

waren (Januar 1996). Bis zu diesem Zeitpunkt waren lediglich Ergebnisse zu einem Probanden mit Bombay ( $0_h$ )-Phänotyp in der Literatur veröffentlicht, der homozygot für eine Mutation (Tyr → ter) in Kodon 316 war, und zu einem weiteren Probanden („para Bombay“-Phänotyp), der heterozygot für Allele war mit Substitutionen im Kodon 164 (Leu → His) und in Kodon 276 (Glu → ter) [58]. Bis Mai 1997 wurde eine ganze Reihe weiterer *h*-Allele meist in „para Bombay“-Phänotypen aus Ostasien beschrieben (siehe Tabelle 4.1), die dort häufiger beobachtet werden.

Die *h*-Allelsequenz-Daten einschließlich der vollen Länge der CDS [67]

und von 164 Nukleotiden des Introns wurden abgelegt in den EMBL-, Genbank- und DNA Data Bank of Japan (DDBJ)-Nukleotid Sequenz Datenbanken unter der Zugriffsnummer Z69587. Der Eintrag in den Nukleotid Sequenz Datenbanken gibt neben den Nukleotidsequenzen auch die Translation und die verschiedenen Annotationen (Genmerkmale und Literaturverweise) wieder (Abbildung 3.2). Über das National Center for Biotechnology Information (zum Beispiel NCBI WWW Entrez PubMed Server; <http://www3.ncbi.nlm.nih.gov/PubMed/>) ist eine schematische Repräsentation der Nukleotid-Informationen im Internet „online“ abrufbar (Abbildung 3.3).

### Abbildung 3.2. Eintrag in den verschiedenen Nukleotid Sequenz Datenbanken. Zugriffsnummer Z69587.

```

LOCUS      HSFUT1MUT      1268 bp      DNA      PRI      01-MAY-1997
DEFINITION H.sapiens FUT1 gene.
ACCESSION  Z69587
NID        g1616598
KEYWORDS   Fucosyltransferase; FUT1 gene.
SOURCE     human.
  ORGANISM Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE  1 (bases 1 to 1268)
  AUTHORS  Wagner,F.F. and Flegel,W.A.
  TITLE    Polymorphism of the h allele and the population frequency of
            sporadic nonfunctional alleles
  JOURNAL  Transfusion 37 (3), 284-290 (1997)
  MEDLINE  97240210
REFERENCE  2 (bases 163 to 1268)
  AUTHORS  Larsen,R.D., Ernst,L.K., Nair,R.P. and Lowe,J.B.
  TITLE    Molecular cloning, sequence, and expression of a human
            GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA
            that can form the H blood group antigen
  JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 87 (17), 6674-6678 (1990)
  MEDLINE  90370848
REFERENCE  3 (bases 1 to 1268)
  AUTHORS  Flegel,W.A.
  TITLE    Direct Submission
  JOURNAL  Submitted (09-FEB-1996) Flegel W. A., University of Ulm,
            Transfusion Medicine, Helmholtzstrasse 10, Ulm, Germany, D-89081

FEATURES             Location/Qualifiers
   source             1..1268
                     /organism="Homo sapiens"
                     /germline
   intron             <1..162
                     /citation=[1]
                     /cons_splice=(5'site:NO,3'site:YES)
                     /number=1
   variation          130
                     /citation=[1]
                     /frequency="0.5"
                     /replace="a"
   exon               163..>1268
                     /citation=[2]
                     /number=2
   gene               165..1262
                     /gene="FUT1"
   CDS                165..1262
                     /gene="FUT1"
                     /note="phenotype: Bombay blood group; coding sequence
                     identical with the wild type FUT1 gene"
                     /citation=[1]
                     /codon_start=1
                     /pseudo
                     /product="nonfunctional variant h7 of fucosyltransferase"
                     /db_xref="PID:e314316"
   CDS                165..1262
                     /gene="FUT1"
                     /standard_name="H transferase"
                     /EC_number="2.4.1.69"

```

## Abbildung 3.2. Fortsetzung

```

/allele="nonfunctional mutants; related sequence: M35531"
/citation=[1]
/citation=[2]
/codon_start=1
/product="fucosyltransferase"
/db_xref="PID:e314315"
/db_xref="PID:g1648868"
/translation="MWLRSHRQLCLAFLLVCVLSVIFFLHIHQDSFPHGLGLSILCPD
RRLVTPPVAIFCLPGTAMGPNASSSCPQHPASLSGTWTVYPNGRFGNQMGQYATLLAL
AQLNGRRRAFILPAMHAALAPVFRITLPVLAPEVDSRTPWRELQLHDWMSEYADLRDP
FLKLSGFPCSWTFFHHLREQIRREFTLHDHLREEAQSVLQQLRLGRTGDRPRTFVGVH
VRRGDYLVQVMPQRWKGVVGDSAYLRQAMDWFRARHEAPVFFVVTSSNGMEWCKENIDTSQ
GDVTFAGDQGEATPWKDFALLTQCNHTIMTIGTGFGWAAAYLAGGDTVYLANFTLPDSE
FLKIFKPEAAFLPEWVGINADLSPLWTLAKP"
allele 625
/gene="FUT1"
/allele="causes amino acid substitution Tyr to Cys at
position 154"
/citation=[1]
/frequency="0.0001"
/phenotype="Bombay blood group"
/replace="g"
/product="nonfunctional mutant h2 of fucosyltransferase
gene"
allele 638
/gene="FUT1"
/allele="silent mutation"
/citation=[1]
/frequency="0.0001"
/phenotype="Bombay blood group"
/replace="g"
/product="nonfunctional mutant h2 of fucosyltransferase
gene"
allele 677
/gene="FUT1"
/allele="causes amino acid substitution Trp to Cys at
position 171"
/citation=[1]
/frequency="0.0001"
/phenotype="Bombay blood group"
/replace="c"
/product="nonfunctional mutant h4 of fucosyltransferase
gene"
allele 940
/gene="FUT1"
/allele="causes amino acid substitution Val to Glu at
position 259"
/citation=[1]
/frequency="0.0001"
/phenotype="Bombay blood group"
/replace="a"
/product="nonfunctional mutant h3 of fucosyltransferase
gene"
allele 1108
/gene="FUT1"
/allele="causes amino acid substitution Ala to Val at
position 315"
/citation=[1]
/frequency="0.0001"
/phenotype="Bombay blood group"
/replace="t"

```

## Abbildung 3.2. Fortsetzung

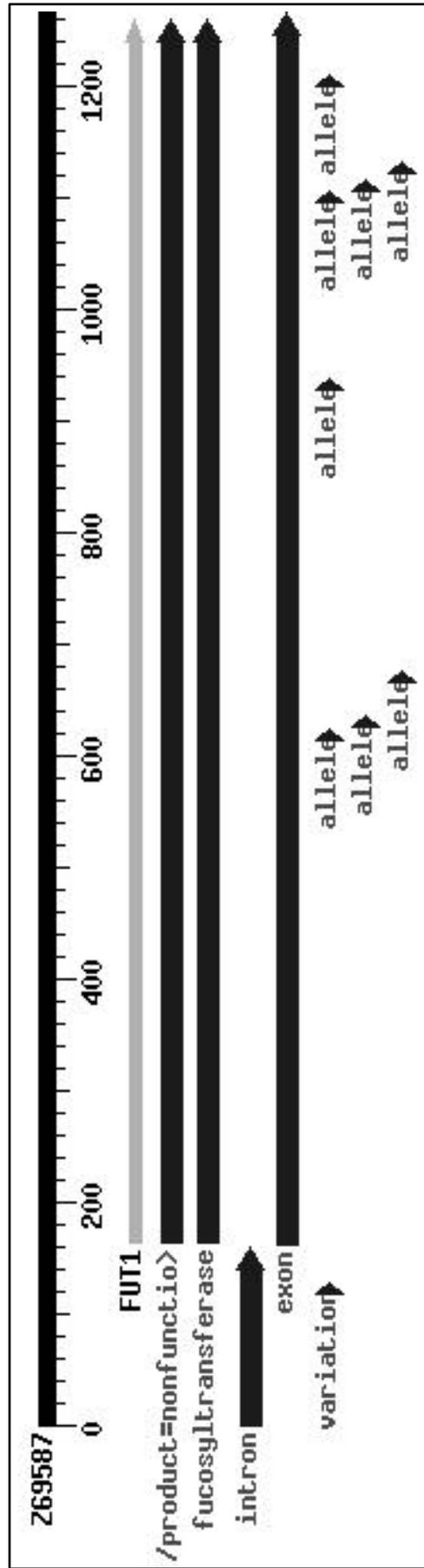
```

        /product="nonfunctional mutant h5 of fucosyltransferase
        gene"
allele 1118
        /gene="FUT1"
        /note="silent mutation"
        /citation=[1]
        /frequency="0.0001"
        /phenotype="Bombay blood group"
        /replace="a"
        /product="nonfunctional mutant h2 of fucosyltransferase
        gene"
allele 1133..1134
        /gene="FUT1"
        /note="two nucleotides deletion causing frameshift
        mutation"
        /citation=[1]
        /frequency="0.0001"
        /phenotype="Bombay blood group"
        /replace=""
        /product="nonfunctional mutant h6 of fucosyltransferase
        gene"
allele 1211
        /gene="FUT1"
        /note="causes amino acid substitution Trp to Cys at
        position 349"
        /citation=[1]
        /frequency="0.0001"
        /phenotype="Bombay blood group"
        /replace="c"
        /product="nonfunctional mutant h1 of fucosyltransferase
        gene"
BASE COUNT      212 a      442 c      333 g      281 t
ORIGIN
  1 aggctctccc tacctcagcc ccaggagtct aaacctttaa cttcctcttt cctctgggact
 61 aaggagtgct gcaccccagg cgctccctt accccacatc cctcctcagc ctcccctcct
121 cagcctcagt gcatttgcta attcgccttt cctcccctgc agccatgtgg ctccggagcc
181 atcgtcagct ctgcctggcc ttctgctag tctgtgtcct ctctgtaatc ttcttctctc
241 atatccatca agacagcttt ccacatggcc taggcctgtc gatcctgtgt ccagaccgcc
301 gcctggtgac acccccagtg gccatcttct gcctgccggg tactgcatg ggcccaacg
361 cctcctcttc ctgtccccag caccctgctt ccctctccgg cacctggact gtctacccca
421 atggccggtt tggtaatcag atgggacagt atgccacgct gctggctctg gccagctca
481 acggccgcog ggcctttatc ctgcctgcca tgcatgcgc cctggccccg gtattccgca
541 tcaccctgcc cgtgctggcc ccagaagtgg acagccgcac gccgtggcgg gagctgcagc
601 ttcacgactg gatgtcggag gagtacgcgg acttgagaga tcctttcctg aagctctctg
661 gcttcccctg ctcttgactt ttcttccacc atctccggga acagatccgc agagagttea
721 ccctgcacga ccaccttcgg gaagaggcgc agagtgtgct gggtcagctc cgctgggccc
781 gcacagggga ccgcccgcgc accttgtcg gcgtccacgt gcgccgtggg gactatctgc
841 aggttatgcc tcagcgtgag aagggtgtgg tgggcgacag cgcctacctc cggcaggcca
901 tggactgggt ccgggcacgg cacgaagccc ccgtttctgt ggtcaccagc aacggcatgg
961 agtgggtgtaa agaaaacatc gacacctccc agggcgatgt gacgtttgct ggcgatggac
1021 aggaggctac accgtggaaa gactttgcc tgctcacaca gtgcaaccac accattatga
1081 ccattggcac ctteggcttc tgggctgctt acctggctgg cggagacact gtctacctgg
1141 ccaacttcac cctgccagac tctgagttcc tgaagatctt taagccggag gcggccttcc
1201 tgcccagatg ggtgggcatt aatgcagact tgtctccact ctggacattg gctaagcctt
1261 gagagcca

```

**Abbildung 3.3. Schematische Darstellung der Nukleotid-Informationen in der NCBI-Datenbank**

Neben der Datenbank Zugriffsnummer Z69587 sind die circa 1,2 kb des Datenbankeintrags als Lineal dargestellt. Darunter sind die Genbezeichnung (FUT1), das Produkt (nicht-funktionales Allel einer Fucosyltransferase) sowie die Sequenzlängen des einzigen Introns und des Exons 2 wiedergeben. Die Positionen der Mutationen für die eine beschriebene Variation bzw. die verschiedenen Allele sind im unteren Teil schematisch angegeben.



### 3.2.2 Homozygotie für private *h*-Allele

Mit einer Familienstudie wurde bewiesen, daß Proband 1 homozygot ist für das  $h^1$ -Allel (Abbildung 3.4). Wir konnten mittels direkter Sequenzierung in den mütterlichen und väterlichen Linien das Vorhandensein des  $h^1$ -Allels (Trp → Cys im Kodon 349) nachweisen.

Die Probanden 3 und 4 waren offensichtlich homozygot für das  $h^3$ - beziehungsweise  $h^4$ -Allel, was ebenfalls mittels direkter Sequenzierung bestimmt wurde (Abbildung 3.5). Ein formaler Beweis der Homozygotie hätte auch bei diesen Probanden Familienuntersuchungen erfordert. Prinzipiell könnten bei Probanden 3 und 4 Deletionen vorliegen, die Teile oder das ganze *FUT1*-Gen umfassen. Da für die populationsgenetische Analyse (siehe Abschnitt 3.4) jedoch lediglich Proband 1 und 2 berücksichtigt wurden, konnte auf die weitergehende Verifizierung der Homozygotie bei Probanden 3 und 4 verzichtet werden.

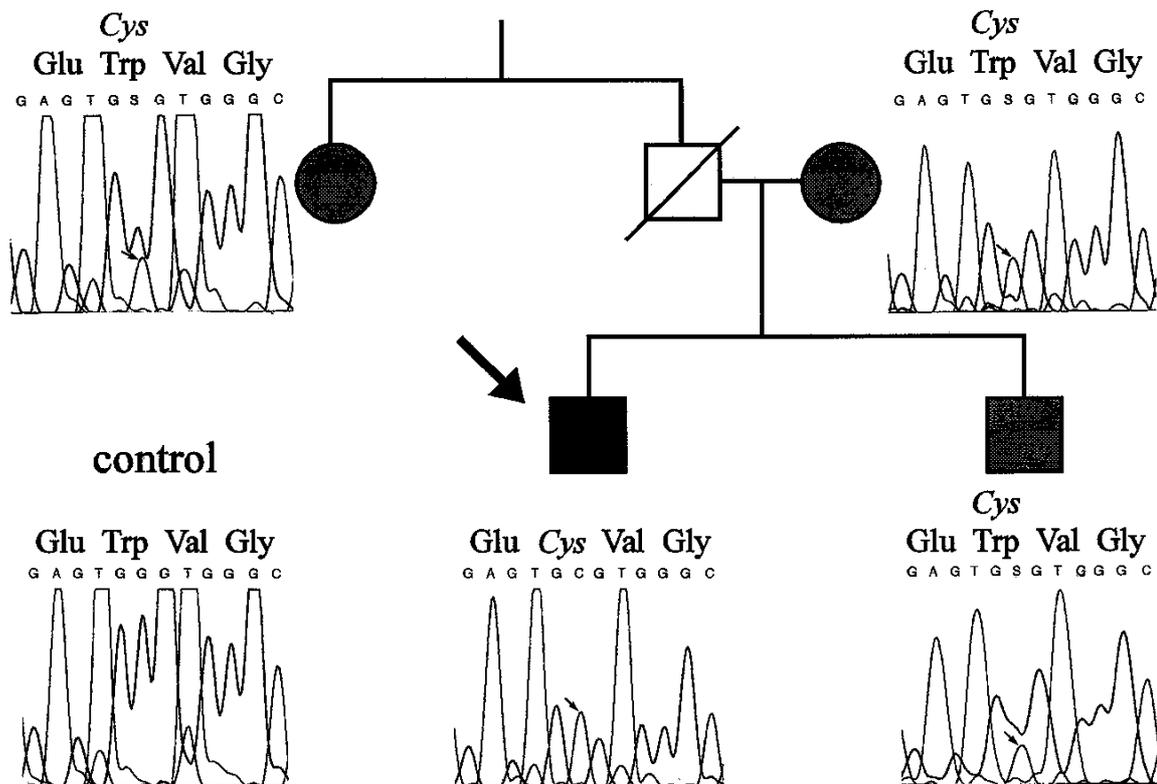
## 3.3 Statistische Betrachtung zum *h*-Polymorphismus

### 3.3.1 Hardy-Weinberg Equilibrium bei seltenen Allelen

Wir beobachteten in der Studie sieben unterschiedliche *h*-Allele in fünf Individuen mit dem Genotyp *hh*. Kein Allel wurde in mehr als einem Probanden gefunden. Wenn man von der Annahme einer guten Durchmischung (Panmixie) im Hardy-Weinberg Equilibrium ausgeht, wäre die Wahrscheinlichkeit dieser Beobachtung

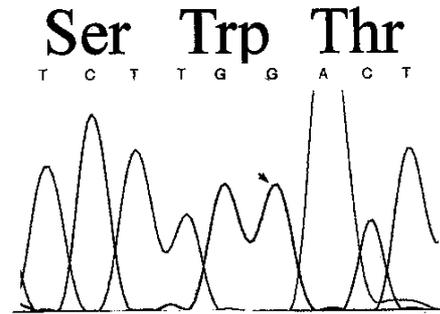
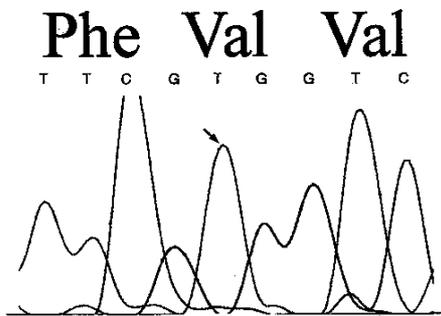
$$1/(9 \times 7 \times 5) = 0,0032 \text{ .}$$

Daraus folgt, daß die Nullhypothese der Panmixie mit der Irrtumswahrscheinlichkeit von  $p = 0,0032$  zurückgewiesen werden kann. Dies bestätigt die aus der Populationsbiologie bekannte Folgerung, daß das Hardy-Weinberg Equilibrium in großen Populationen für vergleichsweise seltene Allele nicht erfüllt ist, die zum Beispiel die Häufigkeit von *h*-Allelen aufweisen. Vielmehr beobachtet man aufgrund einer in natürlichen Populationen immer vorhandenen, geringgradigen Verwandtschaft, die durch den mittleren Konsanguinitätskoeffizient  $\alpha$  (englisch: average inbreeding factor  $\alpha$ ) beschrieben wird [12], oft eine Homozygotie für seltene Allele.

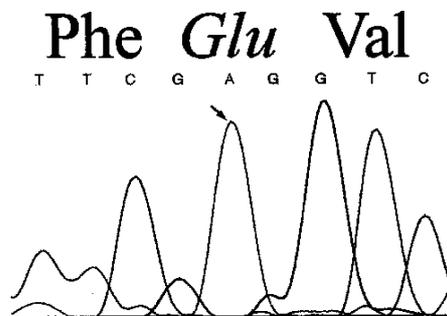


**Abbildung 3.4. Stammbaum des Probanden 1 (großer Pfeil), der homozygot für das  $h^1$ -Allel ist.** Die Sequenzanalyse zeigt die amplifizierten Nukleotidpositionen 981 bis 992 im *FUT1*-Gen. Die kleinen Pfeile deuten die Punktmutation an der Position 986 an ( $G \rightarrow C$ ; S bezeichnet G+C) in der Position des dritten Nukleotids im Tryptophan (Trp)-Kodon, die eine nicht-konservative Aminosäuresubstitution durch Zystein (Cys) kodiert. Alle Familienmitglieder des Probanden 1 tragen das prävalente funktionale *H*-Allel und das nicht-funktionale  $h^1$ -Allel in heterozygoter Form. Sie weisen an der markierten Position (kleine Pfeile) deshalb Signale für zwei Nukleotide auf (G: schwarzer Gipfel; C: blauer Gipfel). Die Kontrollperson ist homozygot für das *H*-Allel und weist deswegen - genauso wie Proband 1 - an dieser Position nur ein Signal auf.

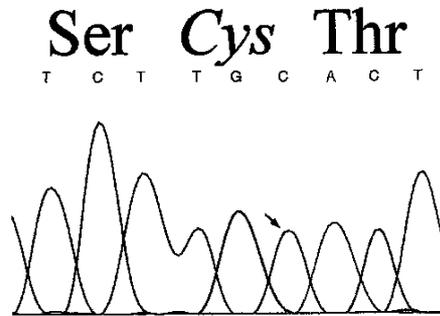
## Control



## Proband 3



## Proband 4



**Abbildung 3.5. Sequenzanalyse der Punktmutationen  $h^3$  (Proband 3) und  $h^4$  (Proband 4).** Die amplifizierte Nukleotide von Position 772 bis 780 für das  $h^3$ -Allel und von 509 bis 517 für das  $h^4$ -Allel sind dargestellt. Die Pfeile bezeichnen die Positionen der beiden Punktmutationen. In der dargestellten direkten Sequenzierung ist eine Heterozygotie an den entsprechenden Nukleotidpositionen nicht nachweisbar. Deswegen tritt an den markierten Positionen jeweils nur das Signal für ein Nukleotid auf.

### 3.3.2 Unabhängige Mutationen als Ursache von *h*-Allelen

Da die Beobachtung von *h*-Allelen in homozygoter Form durch die zu erwartende Konsanguinität in der Population bedingt ist, haben wir tatsächlich sieben unabhängige *h*-Allele in unserer Studie bei fünf Probanden nachgewiesen. Zum Zeitpunkt der molekularen und populationsgenetischen Analyse (Januar 1996) waren lediglich drei weitere unabhängige Beobachtungen bei zwei Probanden in der Literatur beschrieben [58]. Folglich waren damals insgesamt zehn *h*-Allele unabhängig voneinander beobachtet worden (entsprechend zehn Chromosomen mit *FUT1*-Genen).

Unter der Annahme, daß *k* Allele gleicher Frequenz in der Population vorkommen, wäre die Wahrscheinlichkeit, *k* Allele in *n* Beobachtungen zu finden,

$$k \times (k - 1) \times (k - 2) \times \dots \times (k - n + 1) / k^n .$$

Gäbe es nur die zehn beobachteten Allele in der weißen Bevölkerung, wäre die Wahrscheinlichkeit, daß alle bis dahin beobachteten Allele unterschiedlich sind, lediglich  $p = 0,000045$  gewesen. Damit  $p > 0,05$  gilt, müßte es mindestens neunzehn verschiedene Allele in der Bevölkerung geben, die alle für unterschiedliche Aminosäuresubstitutionen, Stopkodons oder „frame shift“-Mutationen kodierten. Diese Abschätzung von mindestens neunzehn Allelen stellt allerdings nur die untere Grenze der tatsächlichen Anzahl von *h*-Allelen dar. Wir vermuteten daher [115], daß mit großer Wahrscheinlichkeit die tatsächliche Zahl in der Population wesentlich höher ist. Diese Vermutung war beim Schreiben dieses Manuskripts bereits bewiesen, da nunmehr insgesamt 16 *h*-Allele in der Literatur veröffentlicht [58,115,118,129] und zudem drei weitere *h*-Allele in Abstraktform beschrieben sind [56]. Eine Zusammenfassung der bekannten *h*-Allele findet sich in der Diskussion (siehe Tabelle 4.1).

## 3.4 Populationsgenetische Analyse

### 3.4.1 Mittlerer Konsanguinitätskoeffizient $\alpha$

Die Bombay ( $O_h$ )-Phänotypfrequenz in unserer Population war 1 in 312.081 [117]. Proband 1 der beiden Blutspender dieser Frequenzstudie (Proband 1 und 2) war nachweislich homozygot für ein sporadisches *h*-Allel. Diese Beobachtung stimmt

überein mit der Homozygotie von zwei Probanden (Proband 1 und 3) von vier untersuchten Probanden deutscher Abstammung (Proband 1, 2, 3 und 5). Hieraus leiten sich

$$f_{\text{homo+hetero}} = 1/312.081$$

und

$$f_{\text{homo}} = (1/312.081) \times 0,5 = 1/624.162$$

ab.

Wir lösten das Gleichungssystem (Formel 1.1 und 1.2 im Methodenteil), indem wir diese experimentell bestimmten Populationsfrequenzen zugrunde legten, und berechneten  $\alpha = 0,00116$  und  $f_{hse} = 1:769$ .

### 3.4.2 Frequenz sporadischer *h*-Allele in der Population

Da die Frequenz des *se*-Allels in der deutschen Bevölkerung mit  $f_{se} = 0,4516$  bekannt war [83], konnten wir die gesuchte Frequenz der *h*-Allele in der deutschen Bevölkerung bestimmen mit

$$f_h = 1:347 \text{ beziehungsweise } f_h = 0,0029.$$

Das 95% Vertrauensintervall der Bombay ( $O_h$ )-Phänotypfrequenz war 1:88.226 - 1:1.758.205 (siehe Tabelle 3.1). Unter Zugrundelegung dieses Intervalls und  $\alpha = 0,00116$  wurde die obere und untere Grenze des 95% Vertrauensintervalls für die Frequenz der *h*-Allele in der deutschen Bevölkerung mit 1:185 - 1:824 abgeleitet.

## 4. DISKUSSION

### 4.1 Häufigkeit sporadischer nicht-funktionaler Allele im *FUT1*-Gen

In dieser Studie konnten wir belegen, daß die nicht-funktionalen Allele im *FUT1*-Gen in der weißen Bevölkerung durch viele unterschiedliche Mutationsereignisse charakterisiert sind. Die Ursache des sehr seltenen Bombay ( $O_h$ )-Blutgruppen-Phänotyps ist deswegen ein „Hintergrund“ von vielen verschiedenen nicht-funktionalen Allelen, die zusammen mit dem vorherrschenden funktionalen *H*-Allel im *FUT1*-Gen auftreten. Solche seltenen nicht-funktionalen Allele können in jedem Gen infolge von ständig stattfindenden Mutationen vorhanden sein.

Wir bestimmten in einer großen weißen Bevölkerung, daß die Frequenz nicht-funktionaler *h*-Allele in der Summe 1 in 347 ( $f_h = 0,0029$ ; 95% Vertrauensintervall: 1:185 - 1:824) beträgt. Obwohl diese Frequenz bisher nur für das *FUT1*-Gen untersucht wurde, kann man verallgemeinern, daß nicht-funktionale Allele mit ähnlichen Frequenzen in vielen anderen Genen auftreten, sofern sie Mutationen ohne wesentlichen Selektionsdruck repräsentieren. Dies kann insbesondere für die Gene anderer Glykosyltransferasen als der H-Transferase zutreffen, die zum Beispiel die Blutgruppensysteme ABO, P und Lewis bilden.

Beim *FUT1*-Gen haben wir bewiesen, daß der Anteil von zufällig mutierten sporadischen nicht-funktionalen Allelen unter allen auftretenden Null-Allelen überwiegt. Diese Verteilung von Null-Allelen mag für viele andere Gene ebenfalls zutreffen, die Tumor-assoziiert auftreten oder Erbkrankheiten verursachen. In diesem Fall können solche Gene prinzipiell nicht verlässlich mit dem Nachweis von einzelnen oder einigen wenigen Null-Allelen untersucht werden.

### 4.2 Polymorphismus nicht-funktionaler Allele im *FUT1*-Gen

In sechs der sieben in dieser Studie gefundenen *h*-Allele konnten „missense“-Mutationen nachgewiesen werden, die für nicht-konservative Aminosäuresubstitutionen kodierten, beziehungsweise Nukleotiddeletionen, die zu einem veränderten Leserahmen führten. Der überwiegende Teil der Mutationen fand sich im 3' Teil der CDS. Diese Befunde stehen in Übereinstimmung mit den anderen inzwischen beschriebenen *h*-Allelen im *FUT1*-Gen [56,58,118,129] (Tabelle 4.1). Wir vermuten, daß diese Aminosäureaustausche im H-Transferaseprotein die Ursachen der feh-

lenden H-Transferaseaktivität darstellen. Der formale experimentelle Beweis des ursächlichen Zusammenhangs würde funktionelle Tests erfordern [67,118] mit Proteinexpression in Zelllinien und Messung der enzymatischen Aktivität. Der Beweis der Ursache war jedoch zum Nachweis des Polymorphismus und für die Schlußfolgerungen dieser Studie nicht erforderlich. Zudem wird der ursächliche Zusammenhang in der Literatur oft ohne experimentellen Beweis akzeptiert, wenn eine Assoziation mit einem Phänotyp oder einer Krankheit nachweisbar ist [16,35].

Es ist auffällig, daß drei der fünf „missense“-Mutationen in den *h*-Allelen unserer Studie für die Aminosäure Zystein kodieren. Bekanntermaßen trägt die Brückenbildung zwischen Zysteinen wesentlich zur Ausbildung der Tertiärstruktur von Proteinen bei. Die aberrant aufgetretenen Zysteine könnten zu veränderten Brückenbildungen und zu Konformationsänderungen des H-Transferaseproteins führen. Unsere gehäufte Beobachtung von Zystein-Substitutionen hätte auf eine Besonderheit beim *FUT1*-Gen hinweisen können. Da bei den *h*-Allelen der anderen Untersuchungen allerdings keine solche Häufung vorkam, gehen wir davon aus, daß unsere Beobachtung zufällig war.

Das *h*<sup>7</sup>-Allel ist von besonderem Interesse, weil es die kodierende Sequenz (CDS) des funktionalen *H*-Allels aufweist, jedoch in einem H-defizienten Individuum nachgewiesen wurde. Diese Beobachtung ist bisher für kein anderes Transferasegen beschrieben worden. Lediglich in Abstraktform [56] findet sich der Hinweis, daß im „klassischen“ Bombay (*O<sub>h</sub>*)-Phänotyp unter Indern mittels DGGE- und SSCP-Techniken ebenfalls keine Mutationen in der CDS des *FUT1*-Gens nachweisbar waren. Weitere Studien sollten belegen, ob es sich bei dem *h*<sup>7</sup>-Allel, das wir bei einem Probanden deutscher Abstammung fanden, um das Allel des „klassischen“ Bombay (*O<sub>h</sub>*)-Phänotyps der indischen Bevölkerung handelt. Zur Klärung der fehlenden Funktionalität des *h*<sup>7</sup>-Allels wäre die Charakterisierung insbesondere der 5´ und 3´ nicht-translatierten cDNA (UTS) sowie der Promotorregion des *h*<sup>7</sup>-Allels sinnvoll. Auch könnte eine kürzlich beschriebene Spleißvariante des *FUT1*-Transkripts eine Erklärung liefern [60]. Bei dieser alternativen Expression in einigen Geweben trägt auch das - von uns nicht untersuchte - Exon 1 zur CDS bei, sodaß prinzipiell Mutationen in diesem Bereich des *FUT1*-Gens Bedeutung für den Verlust der Funktionalität haben könnten.

**Tabelle 4.1. Zusammenstellung der bekannten nicht-funktionalen Allele im *FUT1*-Gen (Stand Mai 1997)**

beobachtete Mutationen		Phänotyp	Population	Homozygotie	experimentell gesichert	OMIM Varianten	Literaturstelle
Nukleotid	Aminosäure						
C35T/A980C	Ala12Val/Asn327Thr	„para Bombay“	taiwanesisch	ja	nein	-	Yu et al. 1997
T460C	Tyr154His	„para Bombay“	taiwanesisch	nein	nein	-	Yu et al. 1997
T460C/G1042A	Tyr154His/Glu348Lys	„para Bombay“	japanisch	ja	ja	-	Wang et al. 1997
A461G	Tyr154Cys	Bombay	deutsch	nein	nein	-	<i>diese Studie</i>
T491A	Leu164His	„para Bombay“	USA	nein	nein	.0002	Kelly et al. 1994
G513C	Trp171Cys	Bombay	türkisch	ja	nein	-	<i>diese Studie</i>
C658T	Arg220Cys	„para Bombay“	taiwanesisch	nein	nein	-	Yu et al. 1997
T776A	Val259Glu	Bombay	deutsch	ja	nein	-	<i>diese Studie</i>
k. A.	Trp267Cys	k. A.	k. A.	k. A.	nein	-	unveröffentlicht
k. A.	Asp278Asn	k. A.	k. A.	k. A.	nein	-	unveröffentlicht
C944T	Ala315Val	H-def. Sekretor	deutsch	nein	nein	-	<i>diese Studie</i>
G1047C	Trp349Cys	Bombay	deutsch	ja: IBS	nein	-	<i>diese Studie</i>
AG547/548del	182 frame shift	„para Bombay“	taiwanesisch	nein	nein	-	Yu et al. 1997
C826T	Glu276ter	„para Bombay“	USA	nein	nein	.0003	Kelly et al. 1994
TT880/881del	294 frame shift	„para Bombay“	taiwanesisch	nein	nein	-	Yu et al. 1997
C948G	Tyr316ter	Bombay	USA	ja: IBS	ja	.0001	Kelly et al. 1994
CT969/970del	323 frame shift	H-def. Sekretor	deutsch	nein	nein	-	<i>diese Studie</i>
G990del	330 frame shift	k. A.	k. A.	k. A.	nein	-	unveröffentlicht
keine Veränderungen in der CDS		Bombay	deutsch	nein	nein	-	<i>diese Studie</i>
keine Veränderungen in der CDS		Bombay	indisch	ja	nein	-	unveröffentlicht

k. A. - keine Angaben. H-def. Sekretor - H-defizienter Sekretor. ja: IBS - Homozygotie und Identität beider Allele der Probanden durch Familienstudien gesichert; ja - Homozygotie wegen fehlendem Nachweis eines zweiten Allels wahrscheinlich. Eine Identität der Allele durch Abstammung (IBD) ist bisher bei keinem molekular charakterisierten *h*-Allel abschließend gesichert worden. Lediglich bei zwei Mutationen wurde bisher experimentell in Expressionsstudien gesichert, daß das mutierte Allel keine Fukosyltransferase-Aktivität aufweist. OMIM - Online Mendelian Inheritance in Man (Datenbank); die Striche weisen darauf hin, daß für die meisten Allele bisher noch keine Numerierung als „Allelic Variants“ in der OMIM Datenbank unter dem Genlocus „\*211100 FUT1“ erfolgt ist. frame shift - ab den angegebenen Kodons sind die Leserahmen verändert, was zu einem vorzeitigen Ende der Translation führt. Die unveröffentlichten Daten sind einem Abstrakt von Johnson et al. [118] entnommen. In dem Abstrakt werden fünf weitere Mutationen in CDS ohne genauere Spezifikationen berichtet, die offensichtlich zum Teil in Yu et al. [129] veröffentlicht wurden. Alle Probanden des „klassischen“ Bombay (O<sub>n</sub>)-Phänotyps indischer Abstammung in Johnson et al. waren anscheinend homozygot für ein Allel, bei dem sich keine Veränderung in der CDS nachweisen ließ (Geoff Daniels, persönliche Mitteilung).

Seit mehr als 20 Jahren sind Hinweise auf eine mögliche serologische Heterogenität der Bombay ( $O_h$ )-Phänotypen veröffentlicht [81]. Die verschiedenen von uns nachgewiesenen  $h$ -Allele könnten eine Ursache für serologische Unterschiede bei H-defizienten Phänotypen bilden. Um diese Hypothese zu prüfen, werden zur Zeit in einer Kooperation mit P. D. Issitt die molekular charakterisierten Bombay ( $O_h$ )-Phänotypen mit serologischen Daten korreliert.

#### 4.3 Mittlerer Konsanguinitätskoeffizient $\alpha$ in der deutschen Bevölkerung

Der überraschend hohe Anteil von Bombay ( $O_h$ )-Phänotypen, die durch die Homozygotie für „private“  $h$ -Allele verursacht sind, ist durch die geringe Konsanguinität erklärt, die selbst in großen und gut durchmischten Populationen auftritt. Diese Schlußfolgerung wird unterstützt durch die früher veröffentlichte Beobachtung [7,99], daß der Bombay ( $O_h$ )-Phänotyp in Indien in solchen Familien vermehrt vorkommen, in denen aus religiösen Gründen Verwandtschaftsehen erlaubt sind. Der Vergleich mit den veröffentlichten Daten aus anderen Populationen ergibt, daß der in dieser Studie bestimmte mittlere Konsanguinitätskoeffizient  $\alpha$  für die baden-württembergische Population gut in die bekannten Daten für Europa paßt (Tabelle 4.2).

**Tabelle 4.2. Mittlere Konsanguinitätskoeffizienten  $\alpha$  in verschiedenen Populationen<sup>1</sup>**

Population	Koeffizient $\alpha$
USA	0,00009
Belgien	0,0005
Frankreich	0,0011
Deutschland <sup>2</sup>	0,00116
Spanien	0,00203
Japan	0,0046
Indien	0,019

<sup>1</sup> Nach [12] Tabelle 7.3; Seite 350f; abgeschätzt aus der Rate blutsverwandter Ehen in Standesregistern. Für deutsche Bevölkerungen lagen in Lehrbüchern bisher keine Abschätzungen vor.

<sup>2</sup> Nach den Ergebnissen und Berechnungen dieser Studie durch molekulargenetische Analyse experimentell bestimmt.

Bisher wurde der mittlere Konsanguinitätskoeffizient  $\alpha$  üblicherweise indirekt abgeschätzt aus dem Vorkommen von Verwandtschaftsehen, wie sie zum Beispiel in Kirchenregistern der römisch-katholischen Kirche im Rahmen von sogenannten Dis-

pensionsaufzeichnungen dokumentiert sind [12]. Dieses Verfahren unterschätzt den Koeffizienten  $\alpha$  geringfügig, aber systematisch, da die meisten Verwandtschaftsverhältnisse der Urgroßeltern und der noch älteren Vorfahren unberücksichtigt bleiben. Andererseits werden manche Nachkommen bei der Berechnung berücksichtigt, obwohl sie in Wahrheit nicht aus der Verbindung stammen.

Unser Ansatz war insofern neuartig, als wir eine direkte Kalkulation des mittleren Konsanguinitätskoeffizienten  $\alpha$  aus experimentellen Ergebnissen vornehmen konnten, ohne auf vermeintlich bekannte Verwandtschaftsverhältnisse zurückgreifen zu müssen.

Allerdings weist unser Ergebnis für den Konsanguinitätskoeffizienten  $\alpha$  ebenfalls nur eine eingeschränkte Genauigkeit auf, da es auf einer kleinen Probandenzahl und einem einzelnen Genlocus beruht. Dennoch zeigt die Zusammenstellung in Tabelle 4.2, daß der von uns bestimmte Koeffizient  $\alpha$  gut übereinstimmt mit den Abschätzungen für andere europäische Populationen. Die Bestimmung des mittleren Konsanguinitätskoeffizienten  $\alpha$  war nicht das primäre Ziel der vorliegenden Studie. Der Koeffizient  $\alpha$  wurde vielmehr im Rahmen des angewandten Gleichungssystems im Zusammenhang mit der Frequenzbestimmung der *h*-Allele abgeleitet. Daher kann der mittlere Konsanguinitätskoeffizient  $\alpha$  als Kontrolle für die Verlässlichkeit des eingesetzten populationsgenetischen Modells und des verwendeten Gleichungssystems betrachtet werden.

#### 4.4 Verlässlichkeit der Frequenzabschätzung von Null-Allelen in der Population

Unsere systematische Untersuchung von über 600.000 Blutspendern auf seltene Phänotypen (Bombay ( $O_h$ )-Phänotyp, D-Deletionen und *p*; siehe Tabelle 3.1) förderte neben den beiden Beobachtungen beim *FUT1*-Gen lediglich eine Beobachtung in der P-Blutgruppe zutage. Die molekulare Analyse und die Kalkulation der Null-Allelfrequenz wurde deswegen am Beispiel des *FUT1*-Gens durchgeführt. Vergleichbare Kalkulationen mit den *P*- und *Rhesus*-Genen würden niedrigere Frequenzen von Null-Allelen ergeben. Die willkürliche Auswahl eines Genlocus, der - im Rahmen der seltenen Ereignisse - die meisten Beobachtungen ergab, könnte so zu einer Überschätzung der Null-Allelfrequenz beigetragen haben.

Eine molekulare Analyse der Ursache von Null-Allelen im *P*-Gen ist zur Zeit nicht möglich, da dieses Blutgruppengen zusammen mit dem *Dombrock*-Gen das letzte der noch nicht klonierten Blutgruppengene darstellt. Der molekulare Mechanismus von Null-Allelen in dem Genlocus der beiden *Rhesus*-Gene ist nach dem derzeitigen Stand der molekularen Ergebnisse komplex, da ein molekular charakterisierter  $Rh_{\text{null}}$ -Phänotyp durch eine Veränderung in einem Begleitprotein verursacht ist [14]. Wegen der nicht abschließend geklärten Molekularbiologie von Null-Phänotypen der *P*- und *Rhesus*-Blutgruppen ist die Anwendbarkeit des Gleichungssystems nicht gesichert. Eine Berechnung für diese beiden Genloci wurde deswegen nicht durchgeführt.

Ein möglicher Fehler bei der Abschätzung findet sich nur angenähert mit der Quadratwurzel im Ergebnis der Null-Allelfrequenzen wieder (siehe Formel 1.3 und Abschnitt 2.4) und hat deswegen einen vergleichsweise geringen Effekt auf die gesuchte Größe. Weiterhin hängt die Relevanz der Null-Allelfrequenz für die Genotypisierungsstrategien lediglich von der Größenordnung ab. Die einzige uns bekannte Abschätzung von Null-Allelfrequenzen [65] ergab in *Drosophila melanogaster* am Beispiel von 25 Enzym-Genloci eine Frequenz von  $f = 0,0023$  und stimmt somit weitgehend mit der von uns beim Menschen bestimmten Null-Allelfrequenz ( $f = 0,0029$ ) überein. Wir nehmen deswegen an, daß die von uns nachgewiesene Frequenz recht verlässlich ist und mit kleineren Anpassungen für viele andere Genloci bei Vertebraten und Invertebraten zutrifft.

#### 4.5 Bedeutung sporadischer nicht-funktionaler Allele für die Genotypisierung am Beispiel des AB0-Blutgruppensystems

Die zur Zeit eingesetzten Verfahren zur Genotypisierung nutzen ganz überwiegend sogenannte „charakteristische“ beziehungsweise „diagnostische“ kurze (< 20 bp) Nukleotidsequenzen. Unter der Annahme einer Rate sporadischer nicht-funktionaler Allele unter den Transferasen des *AB0*-Gens, die ähnlich der im *FUT1*-Gen ist, implizieren unsere Ergebnisse, daß die Vorhersage des Phänotyps aufgrund der zur Zeit verwendeten Genotypisierungsverfahren in ungefähr einer von 300 Bestimmungen von Probanden mit der Blutgruppe 0 versagt (Tabelle 4.3). Die erwartete Rate von Genotypisierungsfehlern ist bei Individuen der Blutgruppe A oder B niedriger. Bei der Blutgruppe AB können grundsätzlich keine Fehltypisierungen aufgrund von nicht-funktionalen Allelen auftreten. Die mittlere kalkulierte Fehlerrate für die AB0-Blutgruppenbestimmung war 1 in 630.

**Tabelle 4.3. Rate der zu erwartenden *ABO*-Genotypisierungsfehler infolge von sporadischen nicht-funktionalen *ABO*-Allelen<sup>1</sup>**

wahrer Phänotyp	wahrer Genotyp <sup>2</sup>	nachgewiesener Genotyp	Phänotyp-Vorhersage	kalkulierte Frequenz in der Bevölkerung (%) <sup>3</sup>	Rate der Genotypisierungsfehler <sup>4</sup>
<b>0</b>	<i>00</i>	<i>00</i>	korrekt	40,83	<b>1 : 308</b>
	<i>a0</i> or <i>b0</i>	<i>A0</i> or <i>B0</i>	inkorrekt	0,1326	
	<i>aa</i> , <i>ab</i> , or <i>bb</i>	<i>AA</i> , <i>AB</i> , or <i>BB</i>	inkorrekt	0,0001	
<b>A</b>	<i>A0</i>	<i>A0</i>	korrekt	35,66	<b>1 : 3.340</b>
	<i>AA</i>	<i>AA</i>	korrekt	7,78	
	<i>Aa</i>	<i>AA</i>	korrekt	0,0449	
	<i>Ab</i>	<i>AB</i>	inkorrekt	0,0130	
<b>B</b>	<i>B0</i>	<i>B0</i>	korrekt	10,35	<b>1 : 846</b>
	<i>BB</i>	<i>BB</i>	korrekt	0,6561	
	<i>aB</i>	<i>AB</i>	inkorrekt	0,0130	
	<i>Bb</i>	<i>BB</i>	korrekt	0,0038	
<b>AB</b>	<i>AB</i>	<i>AB</i>	korrekt	4,52	<b>0</b>
<b>alle <i>ABO</i>-Blutgruppen</b>				100	<b>1 : 630</b>

<sup>1</sup> Die Fehlerraten der Genotypisierung im *ABO*-Gen mit üblichen Genotypisierungsverfahren wurden berechnet unter der Annahme, daß die Frequenz von nicht-funktionalen Allelen identisch zu der *h*-Allelfrequenz ist.

<sup>2</sup> *a* und *b* bezeichnen nicht-funktionale Varianten von *A*- und *B*-Allelen. Varianten der bekannten *O*-Allele führen generell nicht zu Fehlern in der Phänotypvorhersage, wenn die gegenwärtig üblichen Genotypisierungsmethoden angewandt werden.

<sup>3</sup> Berechnet unter der Berücksichtigung der veröffentlichten *ABO*-Allelfrequenzen in der deutschen Bevölkerung [117] und unter Zugrundelegung der *h*-Allelfrequenz für die *a*- und *b*-Allele (*A*-Allel Frequenz von 0,279; *B*, 0,081; *O*, 0,639; *a*, 0,000804; *b*, 0,0002334).

<sup>4</sup> Rate der Genotypisierungsfehler unter Individuen mit den verschiedenen *ABO*-Phänotypen.

Keine der verfügbaren Genotypisierungsmethoden kann gegenwärtig wegen des erheblichen Aufwands und der damit verbundenen Kosten die etablierten Phänotypisierungsverfahren in der Blutgruppenserologie ersetzen. Die vorliegende Studie belegt darüber hinaus, daß die Verlässlichkeit der gegenwärtigen Genotypisierungsverfahren aufgrund der unvermeidbaren Fehlerrate prinzipiell begrenzt ist. Verlässlichere Genotypisierungsstrategien müßten daher technisch grundsätzlich anderen Prinzipien folgen.

#### 4.6 Bedeutung bei Onkogenen

Nicht-funktionale Allele sind offensichtlich von erheblicher klinischer Bedeutung bei allen „Tumorsuppressor-Genen“ wie dem *Rb*-Gen, dem *p53*-Gen oder dem *BRCA1*-Brustkrebssuszeptibilitäts-gen. Die Frequenzen von Null-Allelen in dominanten krankheitsassoziierten Genen wären niedriger als im *FUT1*-Gen, wenn ein Selektionsdruck den Verlust solcher Allele in der Population bedingt. „Missense“-Mutationen, die für ein funktionales Protein kodieren und daher keinem Selektionsdruck unterliegen können (neutrale Allele), verbleiben im Allelpool der Population, während nicht-funktionale Allele - wegen des angenommenen Selektionsdrucks - eher eliminiert werden und verhältnismäßig seltener vorkommen. Deswegen könnten neutrale Allele einen erheblichen Anteil aller varianten Allele bei Onkogenen repräsentieren [15].

In diesem Zusammenhang ist das Beispiel des *BRCA1*-Gens interessant. Die Frequenz aller Tumor-assoziierten *BRCA1*-Mutationen wird auf zirka 1 in 833 in der Bevölkerung geschätzt [34]. Ganz ähnlich dem von uns untersuchten *FUT1*-Gen wurde kein prävalentes Allel unter den nachgewiesenen Null-Allelen beobachtet [17,63]. Ein solches prävalentes Null-Allel scheint allerdings in einer isolierten Population aufzutreten [25,106]. Eine Untersuchung zur Entscheidungsfindung („decision analysis“) [100] zeigte die großen Probleme infolge modifizierender, zur Zeit nicht einschätzbarer Einflußgrößen auf (variable Penetranz und Interaktion mit anderen Genen, Umwelt- und Ernährungsfaktoren), die eine Routineanwendung vor einer aufwendigen Klärung der Zusammenhänge ausschließen [5,47]. Die pathogenetische Relevanz von „missense“-Mutationen kann kaum bestimmt werden, solange kein funktionaler Test bekannt ist. Dies stellt ein erhebliches Problem für eine populationsweite Reihenuntersuchung von krankheitsassoziierten Allelen dar, da klinisch irrelevante Allele (neutrale Allele) nicht von krankheitsassoziierten Allelen (Null-Allele) zu unterscheiden sind und eine solide genetische Beratung vereiteln.

Ein weiteres Beispiel ist das *p53*-Gen. Das *p53*-Gen und sein Proteinprodukt werden sehr intensiv untersucht, seit beobachtet wurde, daß ungefähr 50% der menschlichen Malignome Mutationen in diesem Gen aufweisen. Veränderungen im Sinne eines Null-Phänotyps prädisponieren zu Malignomen [98]. Am häufigsten kommen „missense“-Mutationen vor, seltener Deletionen oder Stopkodon-Muta-

tionen mit vorzeitigem Translationsabbruch. Auch „knock out“-Mäuse mit einer homozygoten *p53*-Null-Mutation entwickeln gehäuft Malignome [20,21].

#### 4.7 Bedeutung bei erbgenetischen Fragestellungen

Null-Allele können bei erbgenetischen Untersuchungen nur dann zu Fehlern führen, wenn für ein Erbmerkmal die molekularen Ergebnisse von Probanden mit serologischen Befunden anderer Probanden verglichen werden. Solange molekulare Untersuchungsergebnisse, die mit identischen Techniken erhoben wurden, innerhalb eines zu prüfenden Abstammungsverhältnisses miteinander verglichen werden, ist ein Fehler aufgrund von Null-Allelen grundsätzlich ausgeschlossen.

Bei der Blutgruppe Duffy kann zum Beispiel die niedrige, aber nicht zu vernachlässigende Frequenz von Null-Allelen ( $f = 0,01910$ ) [52] einen Ausschluß verhindern, wenn der Nachkömmling reinerbig erscheint, jedoch potentieller Träger eines Null-Allels sein kann. Die Genotypisierung - auch im Bereich der Blutgruppen-gene - würde so einen reinerbigen Ausschluß ermöglichen, wo bislang aufgrund der serologischen Techniken Zweifel verblieben sind. Neben der Duffy-Blutgruppe kann dies von praktischer Relevanz bei verschiedenen anderen Blutgruppenmerkmalen wie zum Beispiel der Kell-Blutgruppe sein. Gerade bei Genloci mit niedriger Frequenz von Null-Allelen kann die Genotypisierung der Phänotyp-basierten Abstammungsuntersuchung überlegen sein, weil Genotypisierungsfehler, die bei der Phänotyp-Prädiktion problematisch wären, keinen Einfluß auf Gen-basierte Berechnungen nehmen müssen.

#### 4.8 Genotypisierungsstrategie in Gegenwart von sporadischen nicht-funktionalen Allelen

Viele Genotypisierungsstrategien beruhen auf der Bestimmung von einzelnen oder wenigen Nukleotiden, die für die Funktionalität und Phänotypexpression [126] wichtig oder genetisch mit einem Phänotyp verbunden sind [3,6]. Nukleotid austausche werden üblicherweise nachgewiesen mittels PCR oder „Southern blot“-Technik in Verbindung mit Allel-spezifischen Assays, wie zum Beispiel Restriktionsfragment Längenpolymorphismus (RFLP) [88,104,125], Allel-spezifischen Oligonukleotid- (so genannten ASO-) [22,105] und Peptidnukleotidproben [10], sequenzspezifischen Primern (SSP) [23,72,110,122] und Nukleotidsequenzierungen von kurzen DNA-Abschnitten [96]. Diese Techniken erfassen nicht die volle Länge der CDS, sondern

nur einen kleinen Teil davon. Deswegen können sie grundsätzlich die meisten sporadisch auftretenden Mutationen nicht nachweisen, die einen Verlust der Funktion verursachen und, wie in dieser Studie gezeigt, in der Summe verhältnismäßig oft vorkommen. Solange keine alternativen Methoden zur Verfügung stehen, ist die Fehlerrate durch Null-Allele unvermeidbar. Dennoch sollte der mögliche Fehler in seiner Größenordnung bewußt sein und gegebenenfalls bei der Interpretation im Rahmen einer klinischen Anwendung berücksichtigt werden.

Der Einsatz von anderen Methoden der Genotypisierung, wie Einzelstrang Konformationspolymorphismus (SSCP) [92], „enzyme mismatch cleavage“ [128], Gelelektrophorese mit einem Gradienten denaturierender Bedingungen (DGGE) [24] oder andere Nachweismethoden von Heteroduplex-DNA [103], die viele bisher unbekannte Nukleotidaustausche in Abschnitten genomischer DNA oder cDNA finden, könnte zu einer Reduktion der Rate von Genotypisierungsfehlern bei allen Allelen mit aberranten kodierenden Sequenzen (CDS) führen. Zur Zeit sind solche Techniken noch sehr aufwendig, insbesondere wenn die Abklärung des zugrunde liegenden Polymorphismus abgeschlossen wird.

Allerdings sind auch die Techniken, die die volle Länge der CDS erfassen, nicht in der Lage, solche nicht-funktionalen Allele nachzuweisen, die durch Mutationen in nicht-kodierenden cDNA-Sequenzen oder in der Promotorregion verursacht sind. Solche Mutationen können eine normale Transkription des Gens verhindern oder zu einer Instabilität der mRNA führen. Das von uns beobachtete  $h^7$ -Allel ist ein Beispiel hierfür, da dieses  $h^7$ -Allel eine kodierende Sequenz (CDS) entsprechend dem funktionalen  $H$ -Allel aufwies.

#### 4.9 Qualitätsgesicherte Genotypisierung

Die Qualitätskontrolle für Genotypisierungsverfahren erfordert eine Quantifizierung der möglichen Fehlerrate für die zu untersuchenden Probanden. Aufgrund der erheblichen Unterschiede in der Frequenz und Art der genetischen Varianten zwischen verschiedenen Populationen, die zum Beispiel bei der molekularen Charakterisierung von Blutgruppengenen deutlich werden [116,117], sollte die Abschätzung der Fehlerrate für jede zu untersuchende Population getrennt erfolgen.

Sporadische nicht-funktionale Allele sind in beinahe allen autosomalen Genloci zu erwarten. Es wäre vorteilhaft für verlässliche Strategien zur Genotypisierung, wenn auch sporadische, bisher nicht beschriebene, nicht-funktionale Allele sicher

erkannt würden. Kompliziert wird die Genotypisierung, wie auch in dieser Studie beobachtet, durch Null-Allele, die sich in der kodierenden Sequenz nicht von den funktionalen Allelen unterscheiden. Besondere Bedeutung für die Genotypisierung haben nicht-funktionale Allele in Genloci mit prävalenten Null-Allelen und in Genen unter niedrigem Selektionsdruck, wenn eine Phänotyp-Vorhersage angestrebt wird. Weiterhin können sporadische Null-Allele die populationsweite Reihenuntersuchung auf Onkogene erheblich komplizieren.

Unter der Voraussetzung, daß das Auftreten von „silencer“-Genen zum Beispiel in trans-Position wie beim  $K_0$ -Phänotyp ausgeschlossen ist, könnten zwei Bedingungen zu einer Verbesserung der Verfahren führen: Eine zuverlässige Strategie zur Genotypisierung sollte eine mRNA semiquantitativ nachweisen und deren volle Übereinstimmung mit einer bekanntermaßen funktionalen mRNA belegen.

Die Präsenz einer mRNA kann in der Regel durch eine reverse Transkriptase-reaktion semiquantitativ getestet werden. Als Ausgangsmaterial könnte die mRNA aus Leukozyten oder in der Pränataldiagnostik aus Amnionzellen in Betracht kommen. Letztlich stellt die mRNA aus solchen Zellen - auch im Falle von Blutgruppen - nicht das Transkriptionsprodukt der Gewebe dar, in denen das Genprodukt typischerweise seine Funktion ausübt. Folglich könnten sich die Spleißmuster der normalen (klinisch relevanten) und der ektopischen (zum Beispiel leukozytären) Transkripte unterscheiden. In der Praxis hat dies jedoch keine Probleme verursacht, weil bisher - auch bei ektopischer Transkription - immer korrekt gespleißte Transkripte nachweisbar waren [16].

Die volle Übereinstimmung mit einer bekanntermaßen funktionalen mRNA ist schwierig zu beweisen. Da sie nach jetzigem Stand aus Gründen der Praktikabilität PCR-basiert erfolgen muß, sind zumindest die Primersequenzen nicht verlässlich nachweisbar. Weiterhin können einzelne Nukleotidabweichungen zur Zeit nur sicher durch die sehr aufwendige Sequenzierung belegt beziehungsweise ausgeschlossen werden. Wegen der erheblichen Bedeutung der Genotypisierung ist zu erwarten, daß geeignetere und einfachere Verfahren entwickelt werden. Hierbei erscheinen Heteroduplex-basierte Methoden wie die „enzyme mismatch cleavage“-Methode [128] oder die Sequenzierung durch Hybridisierung an Oligonukleotid-„Arrays“ [13] vielversprechend.

#### 4.10 Ausblick

Die Bestimmung eines Phänotyps oder „Immuntyps“, zum Beispiel der Antigene für Blutgruppenmerkmale, wird zunehmend durch die Genotypisierung ersetzt werden. Eines der sichersten diagnostischen Verfahren ist die ABO-Blutgruppenbestimmung, die mit serologischen Techniken die geringste zu erwartende Fehler rate in der medizinischen Diagnostik überhaupt aufweist. Eine Genotypisierungsstrategie, die eine auch nur annähernd vergleichbare Verlässlichkeit besitzt, würde die meisten herkömmlichen diagnostischen Verfahren bei weitem übertreffen. Insofern kann der Entwicklung von verlässlicheren Genotypisierungsstrategien am Beispiel von Blutgruppengeneten Modellcharakter für die Genotypisierung in der Medizin zukommen.

Bei speziellen Fragestellungen wird bereits jetzt die Genotypisierung zur Blutgruppenbestimmung durchgeführt. Die absehbaren raschen Fortschritte in der Miniaturisierung und Wirtschaftlichkeit der Genotypisierung lassen erwarten, daß in Zukunft die Genotypisierung auch für die Blutgruppenbestimmung breite Anwendung finden wird.

Durch die klinische Blutgruppenbestimmung sind die umfangreichsten verfügbaren Datenbasen zu Phänotypen in natürlichen Populationen erstellt worden. Diese Datenbasen können für die molekulargenetische Analyse von Null-Allelen genutzt werden, wie in dieser Arbeit gezeigt wurde. Andere Allelien, wie zum Beispiel stille Mutationen und Genkonversionen, können ebenso anhand dieser Datenbasen in der Transfusionsmedizin analysiert werden. Solche Untersuchungen werden zum Grundlagenverständnis der Variabilität von Allelen und der chromosomalen Dynamik bei Genen unter niedrigem Selektionsdruck in natürlichen Populationen beitragen.

## 5. ABKÜRZUNGEN

ASO	Allel-spezifisches Oligonukleotid
bp	Basenpaar
cDNA	komplementäre DNA
CDS	kodierende Sequenz
DGGE	Gelelektrophorese mit einem Gradienten denaturierende Bedingungen
DNA	Desoxyribonukleinsäure
<i>h</i> -Allel	nicht-funktionales Allel des <i>FUT1</i> -Gens
IBD	„identical by descent“ - identisch durch Abstammung
IBS	„identical by state“ - identisch aufgrund der Nukleotidsequenz
IgG	Immunglobulin G
IL-1	Interleukin 1
kb	Kilobasen
LDL	„low-density“ Lipoproteine
LPS	Lipopolysaccharid, Endotoxin
mRNA	„messenger RNA“ - Boten RNA
ODN	Oligodesoxynukleotid
OMIM	Online Mendelian Inheritance in Man
ORF	„open reading frame“ - offener Leserahmen
PCR	Polymerase Kettenreaktion
RFLP	Restriktionsfragment Längenpolymorphismus
RNA	Ribonukleinsäure
RT	Reverse Transkriptase
SDS	„Sodium Dodecyl Sulfat“ - Natriumlaurylsulfat
SDS-PAGE	SDS-Polyacrylamid Gelelektrophorese
SSCP	Einzelstrang Konformationspolymorphismus
SSP	Sequenz-spezifischer Primer
SSP-PCR	Polymerase Kettenreaktion mit Sequenz-spezifischen Primern
TNF $\alpha$	Tumor Nekrose Faktor
UTS	„untranslated sequence“ - nicht-translatierte Sequenz
WWW	World Wide Web



## 6. ZUSAMMENFASSUNG UND ÜBERSICHT EIGENER ARBEITEN

### 6.1 Zusammenfassung

Für die Genotypisierung werden gegenwärtig häufig PCR-gestützte Techniken benutzt, die kurze, sogenannte charakteristische beziehungsweise diagnostische Sequenzen nachweisen. Mit diesen Techniken kann der Phänotyp regelmäßig nicht korrekt vorhergesagt werden, wenn sporadische nicht-funktionale Allele angetroffen werden. Die Häufigkeit solcher sporadischer aberranter Allele und ihre dadurch bedingte Bedeutung für die Genotypisierung waren für kein Gen beim Menschen unter niedrigem Selektionsdruck bekannt.

In der vorliegenden Arbeit wurde die Frequenz der sehr seltenen Bombay ( $O_h$ )-Blutgruppe (Genotyp  $hh\ sese$ ) in einer systematischen Suche unter mehr als 600.000 Blutspendern in Baden-Württemberg bestimmt. Die Ergebnisse dieser Studie zusammen mit molekularbiologischen Untersuchungen wurden ausgewertet, um die Populationsfrequenz nicht-funktionaler Allele am Beispiel des Gens zu bestimmen, das für die  $\alpha(1,2)$ Fukosyltransferase kodiert ( $FUT1$  oder  $H$ ).

Sieben unterschiedliche  $h$ -Allele wurden in fünf unverwandten Individuen gefunden. Die nachgewiesenen  $h$ -Allele unterschieden sich bei allen untersuchten Probanden. Drei der Probanden waren homozygot für ihr privates Allel. Es wurde kein prävalentes  $h$ -Allel nachgewiesen. Im einzelnen wurden eine „frame shift“- und fünf „missense“-Mutationen gefunden, die wahrscheinlich die Null-Phänotypen verursachten. Die kodierende Sequenz eines  $h$ -Allels war identisch mit der Sequenz des funktionalen  $H$ -Allels. Der mittlere Konsanguinitätskoeffizient  $\alpha$  war 0,00116. Die Frequenz der nicht-funktionalen Allele im  $FUT1$ -Gen wurde mit 1:367 in einer großen deutschen Population berechnet (95% Konfidenzintervall: 1:185 - 1:824).

Der Bombay ( $O_h$ )-Phänotyp wird in der weißen Bevölkerung durch diverse, sporadische nicht-funktionale Allele verursacht. Es gibt kein prävalentes  $h$ -Allel. Unser Ansatz erlaubte zum ersten Mal eine direkte Kalkulation des Koeffizienten  $\alpha$  für eine Population ohne Berücksichtigung von Verwandtschaftsverhältnissen. Unter der Annahme ähnlicher Häufigkeiten von nicht-funktionalen Allelen in anderen Glykosyltransferase-Genen, wie zum Beispiel in den Genen der AB0-Blutgruppe, können die gegenwärtigen Genotypisierungsstrategien Fehlerraten mit einer Häufigkeit von einem in ungefähr 300 Individuen der Blutgruppe 0 aufweisen. Sporadische

neutrale Allele können ebenso ein erhebliches Problem für die populationsweite Genotypisierung vieler krankheitsassoziierter Gene darstellen.

## 6.2 Summary

### **FREQUENCY OF SPORADIC NON-FUNCTIONAL ALLELES AND THEIR RELEVANCE FOR GENOTYPING EXEMPLIFIED BY THE POLYMORPHISM OF THE *FUT1* BLOOD GROUP GENE**

Current PCR-based strategies for phenotype prediction often fail when sporadic non-functional alleles are encountered. The population frequency of such mutations was not known for any gene under low selection pressure and may be best examined in blood group systems lacking prevalent non-functional alleles. We recently determined the frequency of the very rare Bombay blood group ( $O_h$ , genotype  $hh$  sese) in a systematic survey of more than 600,000 white individuals. Using this survey in conjunction with additional blood samples, we determined the population frequency of non-functional alleles of the *FUT1* gene encoding the  $\alpha(1,2)$ fucosyltransferase. Seven different  $h$  alleles in five unrelated individuals were found, three of whom were homozygous for unique  $h$  alleles. There was no prevalent  $h$  allele. We observed five missense and one frameshift mutations that were the presumptive causes of the null phenotype; the coding sequence of one  $h$  allele was identical to the  $H$  sequence. The average inbreeding factor  $\alpha$  was 0.00116. The frequency of non-functional alleles at the *FUT1* gene locus was calculated as 1:347 in a large white population (95% confidence interval: 1:185 - 1:824). The Bombay blood group phenotype in whites is due to diverse, sporadic non-functional alleles without any prevalent allele. Assuming similar rates of non-functional alleles in glycosyltransferase genes like *ABO*, current genotyping strategies may fail as often as once in about 300 individuals of blood group O. Sporadic neutral alleles may also pose a serious obstacle for population-wide screening of many disease-associated genes.

### 6.3 Immunantwort auf bakterielle Toxine

#### 6.3.1 Endotoxin-Inhibition durch humane Lipoproteine

Die Toxizität von Lipopolysacchariden (LPS, Endotoxin) ist zum großen Teil durch die Aktivierung von Monozyten beziehungsweise Makrophagen und die Freisetzung von Monokinen (Interleukin 1, Tumor Nekrose Faktor) vermittelt [37,38]. Seit langem ist die Bindung von LPS an Serum-Lipoproteine bekannt [36]. Durch die Messung von IL-1, TNF $\alpha$  und IL-6 konnte gezeigt werden, daß die Interaktion von verschiedenen Lipopolysacchariden und Lipid A mit humanem Serum die Aktivierung von humanen Monozyten verhindern kann. Die inaktivierten LPS-Mengen betragen etwa 10 ng pro ml humanes Serum [33]. Diese LPS-inaktivierende Kapazität des Serums war eine Funktion der humanen Lipoproteine. Andere Serumkomponenten wie natürlich auftretende Anti-LPS IgG-Antikörper, Komplement oder freie Serumlipide hatten keinen signifikanten Effekt. Weiterführende Experimente belegten die Inaktivierung von LPS durch „low-density“ Lipoproteine (LDL) [27]. Die „high-density“ Lipoproteine (HDL), an die LPS bevorzugt bindet [36,111], inaktivierten LPS in wesentlich geringerem Umfang. Die „very low-density“ Lipoproteine (VLDL) inaktivierten LPS nicht. Die Lipidanteile der Lipoproteine waren nicht unbedingt erforderlich für die LPS-Inaktivierung, weil gereinigtes humanes Apolipoprotein A-I (ohne seinen physiologischen Lipidanteil) ebenfalls eine Inhibition der LPS-induzierten Monokin-freisetzung bewirkte [27].

Diese in vitro-Ergebnisse wurden durch in vivo-Untersuchungen bei Patienten mit familiärer Hypercholesterinämie im Rahmen von klinischen LDL-Apharesen bestätigt [119]. Vor der LDL-Entfernung durch Immunapharese inaktivierten die Patientenserum zirkulierende LPS in dem Testsystem um etwa fünfmal mehr LPS in dem Testsystem als nach der LDL-Entfernung. Diese LPS-inaktivierende Kapazität, die im Rahmen der LDL-Apharesen entfernt wurde, konnte weitgehend in dem LDL-reichen Eluat der Immunabsorptionssäulen wiedergewonnen werden. Wir schlußfolgerten aus den in vitro-Daten und den klinischen Untersuchungen, daß die Inhibition der LPS-induzierten Monozytenaktivierung durch humanes Serum zu einem großen Teil von humanem LDL abhängig ist.

### 6.3.2 Aktivierung humaner Monozyten durch *Clostridium difficile* Toxin B

An humanen Blutmonozyten konnte nachgewiesen werden, daß *Clostridium difficile* Toxin B in einer Dosis von  $10^{-12}$  M die Freisetzung von Interleukin 1, Tumor Nekrose Faktor und Interleukin 6 induziert [29]. Durch geeignete Kontrollen (Messung der LPS-Kontamination, Toxin B-Inhibition durch polyklonale Antikörper und Toxin B-Hitzeinaktivierung) konnte ein Aktivierungsartefakt durch LPS-Kontamination sicher ausgeschlossen werden. Auffälligerweise war die getestete Toxin A-Präparation ungefähr 1.000fach weniger aktiv als Toxin B. Aufgrund der Präparationstechnik (Protein-Chromatographie) ist es wahrscheinlich, daß eine geringfügige Toxin B-Kontamination in der Toxin A-Präparation das aktive Agens war. Wir belegten, daß ein kurzzeitiger Kontakt der Monozyten mit Toxin B ausreichte, um die IL-1 Freisetzung zu bewirken. Mit Supravitalfärbungen konnte die bekannte Zelltoxizität auch für humane Monozyten belegt werden. Diese Toxizität und die induzierten proinflammatorischen Zytokine (IL-1 und TNF) können an dem Pathomechanismus der *Clostridium difficile*-Infektion und der pseudomembranösen Kolitis beteiligt sein.

### 6.4 Homolog eines humanen Leukämie-Onkogens in *Drosophila melanogaster*

*PBX* ist als *E2A/PBX*-Proteinhybrid eine wesentliche Ursache bei pädiatrischen B Zell-lymphoblastischen Leukämien [85,86]. Zunächst war die Funktion des Proto-Onkogens *PBX* unklar. Um Anhaltspunkte für die Proteinfunktion zu gewinnen, wurde das Gen in *Drosophila melanogaster* kloniert, das homolog zum humanen Proto-Onkogen *PBX* ist [30]. Anhand des Gen-Homologs (*Dpbx*; 1994 umbenannt: *exd*) konnte das Expressionsmuster in der frühen Embryonalentwicklung der Fruchtfliege bestimmt werden. Diese Ergebnisse wurden kurze Zeit später unabhängig bestätigt [93]. Unsere Arbeit bewies, daß das *exd*-Genprodukt wichtige Funktionen in der Transkriptionsregulation aufweisen muß, und deutete auf eine Repression des *exd*-Gens durch ein Gen aus dem Bithorax-Komplex hin [30] - eine Vorhersage, die inzwischen experimentell gesichert ist [112]. Der Nachweis des *exd*-Gens belegte, daß es eines der in der Evolution am stärksten konservierten Proto-Onkogene ist.

## 6.5 Molekularbiologie und Populationsgenetik von Blutgruppengen

Die vorhandenen Ergebnisse zu menschlichen Blutgruppen [83] stellen den umfangreichsten Datensatz dar, der für Phänotypen in natürlichen Populationen verfügbar ist. Die entsprechenden Datenbasen bieten einen guten Ausgangspunkt für Untersuchungen bezüglich der genetischen Diversität in natürlichen Populationen [26]. Wir nutzten die umfangreiche Datenbasis der Blutgruppen-Phänotypen, die uns zur Verfügung stand [117], für die Charakterisierung von Proteinhybriden beim *Rhesus*-Haplotyp und zur molekulargenetischen Analyse von Null-Allelen beim *FUT1*-Gen. Die Genotypisierung als Ersatz der serologischen Blutgruppenbestimmung wird bereits in der Pränataldiagnostik angewandt.

### 6.5.1 *Rhesus*-Gene

Mittels effizienter serologischer Untersuchungsmethoden, die in unserem Labor etabliert wurden, konnten wir bei über 60.000 Blutspendern die Frequenz von RhD-Varianten bestimmen [31], in deren RhD-Protein einzelne Epitope fehlen und die deswegen als „partial RhD“ bezeichnet werden. Unter unverwandten Blutspendern wurden etwa 100 Blutproben mit partialem RhD gefunden. Das häufigste partielle RhD war D Kategorie VII. Weiterhin konnten wir die Frequenzen von D<sup>VI</sup>, D<sup>V</sup>, D<sup>IV</sup>, DNU und Rh33 bestimmen [4,31,117]. Insgesamt 68 D<sup>VII</sup> Blutproben wurden mittels Serologie und Durchflußzytometrie untersucht und im Rahmen des „3rd International Workshop on Monoclonal Antibodies against Human Red Blood Cells“ (Nantes, Frankreich, September 1996) vorgestellt [31]. Unter diesen Blutproben wurden 28 Proben zufällig ausgewählt und mittels molekularbiologischen Methoden untersucht. Die Sequenzierung der Basen 280 bis 335, in denen der D<sup>VII</sup> Phänotypspezifische Nukleotidaustausch an Position 329 auftritt, belegte zusammen mit den Ergebnissen einer sequenzspezifischen PCR, daß alle D<sup>VII</sup> Proben durch die gleiche molekulare Ursache bedingt sind [28]. Unsere Ergebnisse stimmten mit den Berichten der anderen internationalen Arbeitsgruppen überein, die ihre Ergebnisse von insgesamt zehn D<sup>VII</sup> Proben darstellten [11].

Die molekulare Analyse von 23 Blutproben mit D Kategorie VI bestätigte zwei bereits beschriebene molekulare Ursachen (Austausch der Exons 4/5 beziehungsweise 4-6). Daneben fanden wir eine neue molekulare Ursache. Die von uns beschriebene *D Kategorie VI Typ III* stellt ein Hybridgen dar, bei dem die Exons 3 bis 6

im *RHD*-Gen durch die homologen Exons des *RHCE*-Gens im Sinne eines Genkonversionsereignisses substituiert sind [116]. Dieser dritte D Kategorie VI-Typ unterschied sich von den beiden bekannten D Kategorie VI-Typen durch eine normale Antigendichte [32] und war somit immunhämatologisch von den beiden bekannten D Kategorie VI-Typen zu differenzieren. Dieser Unterschied mag Bedeutung haben bei der Frage, ob Blutpräparate der Kategorie D<sup>VI</sup> bei Rhesus negativen Empfängern eine Immunisierung hervorrufen könnten; dies ist zumindest bei dem von uns beschriebenen D Kategorie VI Typ III eher anzunehmen, weil dieser Phänotyp eine weitgehend normale Antigendichte aufweist.

Die Charakterisierung der molekularen Heterogenität der *RHD*-Allele in der deutschen Bevölkerung hat einen unmittelbaren Einfluß auf die Methodik der molekularen Typisierung von Rhesus D, zum Beispiel in der Pränataldiagnostik. Populationsbasierte molekulare Daten können zur Formulierung von verbesserten molekularen *RHD*-Typisierungsstrategien beitragen. Unsere serologischen und molekularen Ergebnisse haben bereits ihre praktische Anwendung in den Richtlinien der Bundesärztekammer gefunden, die weitgehend den von uns veröffentlichten Vorschlag zur Rhesus D-Typisierung und -Transfusionsstrategie [117] als Richtlinie für die Blutgruppenserologie [121] und als Leitlinie für die Transfusionstherapie [114] berücksichtigt haben.

### 6.5.2 *FUT1*-Gen

Wie in der Einleitung dargelegt, eignet sich das *FUT1*-Gen gut zur Frequenzbestimmung von Null-Allelen in großen Populationen, was der Gegenstand dieser Habilitationsschrift ist. Wesentliche Teile der in dieser Arbeit beschriebenen Ergebnisse sind veröffentlicht in *Infusionsther. Transfusionsmed.* 1995; 22:285ff [117] und *Transfusion* 1997; 37:284ff [115].

## 7. LITERATUR

1. Aloysia M, Gelb AG, Fudenberg H, Hamper J, Tippett P und Race RR (1961) The expected "Bombay" group OhA1 and OhA2. *Transfusion* 1: 212-217
2. Anstee DJ (1995) Blood group antigens defined by the amino acid sequences of red cell surface proteins. *Transfus Med* 5: 1-13
3. Arce MA, Thompson ES, Wagner S, Coyne KE, Ferdman BA und Lublin DM (1993) Molecular cloning of RhD cDNA derived from a gene present in RhD-positive, but not RhD-negative individuals. *Blood* 82: 651-655
4. Avent ND, Jones JW, Liu W, Scott ML, Voak D, Flegel WA, Wagner FF und Green C (1996) Molecular bases of D-variants DNU and DII: localisation of residues critical for epD3, 4 and 9 expression. *Transfus Med* 6: 21S (Abstr)
5. Beckmann MW, Niederacher D, Goecke TO, Boddien-Heidrich R, Schnürch H-G und Bender HG (1997) Hochrisikofamilien mit Mamma- und Ovarialkarzinomen. *Dtsch Ärzteblatt* 94: B-139-145
6. Bennett PR, Le Van Kim C, Colin Y, Warwick RM, Cherif-Zahar B, Fisk NM und Cartron JP (1993) Prenatal determination of fetal RhD type by DNA amplification. *N Engl J Med* 329: 607-610
7. Bhatia HM und Sathe MS (1974) Incidence of 'Bombay' (Oh) phenotype and weaker variants of A and B antigen in Bombay (India). *Vox Sang* 27: 524-532
8. Bhende YM, Deshpande CK, Bhatia HM, Sanger R, Race RR, Morgan WTJ und Watkins WM (1952) A "new" blood group character related to the ABO system. *Lancet* 1: 903-904
9. Bottema CDK, Ketterling RP, Li S, Yoon H-P, Phillips JA und Sommer SS (1991) Missense mutations and evolutionary conservation of amino acids: evidence that many of the amino acids in factor IX function as "spacer" elements. *Am J Hum Genet* 49: 820-838
10. Carlsson C, Jonsson M, Norden B, Dulay MT, Zare RN, Noolandi J, Nielsen PE, Tsui L-C und Zielenski J (1996) Screening for genetic mutations. *Nature* 380: 207 (Leserzuschrift)
11. Cartron J-P (1996) Rh DNA - coordinator's report. *Transfus Clin Biol* 3: 491-495
12. Cavalli-Sforza LL und Bodmer WF (1971) *The Genetics of Human Populations*. W. H. Freeman, San Francisco
13. Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS und Fodor SPA (1996) Accessing genetic information with high-density DNA arrays. *Science* 274: 610-614
14. Cherif-Zahar B, Raynal V, Le Van Kim C, D'Ambrosio A-M, Bailly P, Cartron JP und Colin Y (1993) Structure and expression of the RH locus in the Rh-deficiency syndrome. *Blood* 82: 656-662
15. Collins FS (1996) BRCA1 - lots of mutations, lots of dilemmas. *N Engl J Med* 334: 186-188
16. Cooper DN und Krawczak M (1993) *Human Gene Mutation*. Bios Scientific Publishers, Oxford

17. Couch FJ, DeShano ML, Blackwood MA, Calzone K, Stopfer J, Campeau L, Ganguly A, Rebbeck T und Weber BL (1997) BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer. *N Engl J Med* 336: 1409-1415
18. Daniels GL, Anstee DJ, Cartron JP, Dahr W, Issitt PD, Jorgensen J, Kornstad L, Levene C, Lomas-Francis C, Lubenko A, Mallory D, Moulds JJ, Okubo Y, Overbeeke M, Reid ME, Rouger P, Seidl S, Sistonen P, Wendel S, Woodfield G und Zelinski T (1995) Blood group terminology 1995. *Vox Sang* 69: 265-279
19. Davey RJ, Tourault MA und Holland PV (1978) The clinical significance of anti-H in an individual with the Oh (Bombay) phenotype. *Transfusion* 18: 738-742
20. Donehower LA, Godley LA, Aldaz CM, Pyle R, Shi Y-P, Pinkel D, Gray J, Bradley A, Demina D und Varmus HE (1995) Deficiency of p53 accelerates mammary tumorigenesis in Wnt-1 transgenic mice and promotes chromosomal instability. *Genes Devel* 9: 882-895
21. Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CAJ, Butel JS und Bradley A (1992) Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 356: 215-221
22. Eiermann TH, Fakler J, Müller CR, Ballas M und Goldmann SF (1991) HLA-DPB1 oligonucleotide typing of a Southwest German Caucasian population. *Tissue Antigens* 38: 193-198
23. Faas BHW, Simsek S, Bleeker PMM, Overbeeke MAM, Cuijpers HTM, von dem Borne AEGK und van der Schoot CE (1995) Rh E/e genotyping by allele-specific primer amplification. *Blood* 85: 829-832
24. Fischer SG und Lerman LS (1983) DNA fragments differing by single base-pair substitutions separated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci USA* 80: 1579-1583
25. FitzGerald MG, MacDonald DJ, Krainer M, Hoover I, O'Neil E, Unsal H, Silva-Arrieto S, Finkelstein DM, Beer-Romero P, Englert C, Sgroi DC, Smith BL, Younger JW, Gerber JE, Duda RB, Mayzel KA, Isselbacher KJ, Friend SH und Haber DA (1996) Germ-line BRCA1 mutations in Jewish and non-Jewish women with early-onset breast cancer. *N Engl J Med* 334: 143-149
26. Flegel WA (im Druck) Transfusion. In: Delves PJ und Roitt IM (Hrsg.) *Encyclopedia of Immunology*. 2. Auflage: Academic Press, San Diego
27. Flegel WA, Baumstark MW, Weinstock C, Berg A und Northoff H (1993) Prevention of endotoxin-induced monokine release by human low-and high-density lipoproteins and by apolipoprotein A-I. *Infect Immun* 61: 5140-5146
28. Flegel WA, Hillesheim B, Kerowgan M, Kasulke D und Wagner FF (1996) Lack of heterogeneity in the molecular structure of RHD category VII. *Transfusion* 36: 50S (Abstr)
29. Flegel WA, Müller F, Däubener W, Fischer H-G, Hadding U und Northoff H (1991) Cytokine response by human monocytes to *Clostridium difficile* toxin A and toxin B. *Infect Immun* 59: 3659-3666

30. Flegel WA, Singson AW, Bang AG, Margolis JS, Posakony JW und Murre C (1993) Dpbx, a new homeobox gene related to the human proto-oncogene pbx1: molecular structure and developmental expression. *Mech Dev* 41: 155-161
31. Flegel WA und Wagner FF (1996) The frequency of RHD protein variants in Caucasians. *Transfus Clin Biol* 3: 10s (Abstr)
32. Flegel WA und Wagner FF (1996) RHD epitope density profiles of RHD variant red cells analyzed by flow cytometry. *Transfus Clin Biol* 3: 429-431
33. Flegel WA, Wölpl A, Männel DN und Northoff H (1989) Inhibition of endotoxin-induced activation of human monocytes by human lipoproteins. *Infect Immun* 57: 2237-2245
34. Ford D, Easton DF und Peto J (1995) Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. *Am J Hum Genet* 57: 1457-1462
35. French DL und Coller BS (1997) Hematologically important mutations: Glanzmann thrombasthenia. *Blood Cells Molecul Dis* 23: 39-51
36. Freudenberg MA, Bog-Hansen TC, Back U und Galanos C (1980) Interaction of lipopolysaccharides with plasma high-density lipoprotein in rats. *Infect Immun* 28: 373-380
37. Freudenberg MA und Galanos C (1991) Tumor necrosis factor alpha mediated the lethal activity of killed gram-negative and gram-positive bacteria in D-galactosamine-treated mice. *Infect Immun* 59: 2110-2115
38. Freudenberg MA, Keppler D und Galanos C (1986) Requirement for lipopolysaccharide-responsive macrophages in galactosamine-induced sensitization to endotoxin. *Infect Immun* 51: 891-895
39. Frickhofen N, Arnold R, Hertenstein B, Wiesneth M und Young NS (1992) Parvovirus B19 infection and bone marrow transplantation. *Ann Hematol* 64 (suppl): A121-A124
40. Frydman M, Etzioni A, Eidlitz-Markus T, Avidor I, Shechter Y, Orlin JB und Gershoni-Baruch R (1992) Rambam-Hasharon syndrome of psychomotor retardation, dwarfism, defective leukocyte motility, and Bombay phenotype. *Am J Med Genet* 44: 297-302
41. Frydman M, Vardimon D, Shalev E und Orlin JB (1996) Prenatal diagnosis of Rambam-Hasharon syndrome. *Prenat Diagn* 16: 266-269
42. Garratty G (1995) Blood group antigens as tumor markers, parasitic/bacterial/viral receptors, and their association with immunologically important proteins. *Immunol Invest* 24: 213-232
43. Gassner C, Schmarda A, Nussbaumer W und Schönitzer D (1996) ABO glycosyltransferase genotyping by polymerase-chain reaction using sequence-specific primers. *Blood* 88: 1852-1856
44. Gerard G, Vitrac D, Le Pendu J, Muller A und Oriol R (1982) H-deficient blood groups (Bombay) of Reunion Island. *Am J Hum Genet* 34: 937-947
45. Gomez L und Chrispeels MJ (1994) Complementation of an *Arabidopsis thaliana* mutant that lacks complex asparagine-linked glycons with the human cDNA encoding N-acetylglucosaminyltransferase I. *Proc Natl Acad Sci USA* 91: 1829-1833

46. Grunnet N, Steffensen R, Bennett EP und Clausen H (1994) Evaluation of histo-blood group ABO genotyping in a Danish population: frequency of a novel O allele defined as O2. *Vox Sang* 67: 210-215
47. Healy B (1997) BRCA genes - bookmaking, fortunetelling, and medical care. *N Engl J Med* 336: 1448-1449
48. Hirano K, Kawa S, Oguchi H, Kobayashi T, Yonekura H, Ogata H und Homma T (1987) Loss of Lewis antigen expression on erythrocytes in some cancer patients with high serum CA19-9 levels. *J Natl Canc Inst* 79: 1261-1266
49. Ho M, Chelly J, Carter N, Danek A, Crocker P und Monaco AP (1994) Isolation of the gene for McLeod syndrome that encodes a novel membrane transport protein. *Cell* 77: 869-880
50. Horuk R, Chitnis CE, Darbonne WC, Colby TJ, Rybicki A, Hadley TJ und Miller LH (1993) A receptor for the malarial parasite *Plasmodium vivax*: the erythrocyte chemokine receptor. *Science* 261: 1182-1184
51. Hrubisko M, Laluha J, Mergancova O und Zakovicova S (1970) New varinats in the ABOH blood group system due to interaction of recessive genes controlling the formation of H antigen in erythrocytes: the 'Bombay'-like phenotypes OHm, OBHm and OABHm. *Vox Sang* 19: 113-122
52. Hummel K (1992) *Biostatistical Tables for Paternity Opinions*. Johannes Krause, Freiburg im Breisgau
53. Ioffe E und Stanley P (1994) Mice lacking N-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proc Natl Acad Sci USA* 91: 728-732
54. Issitt PD (1985) *Applied Blood Group Serology*. 3. Auflage: Montgomery Scientific Publications, Miami
55. Issitt PD (1993) Null red blood cell phenotypes: associated biological changes. *Transfus Med Rev* 7: 139-155
56. Johnson PH, Mak MK, Leong S, Broadberry R, Duraisamy G, Gooch A, Lin-Chu M, Makar Y, Okubo Y, Samrt E, Koepsall E und Ewers M (1994) Analysis of mutations in the blood-group H gene in donors with H-deficient phenotypes. *Vox Sang* 67(suppl 2): 25 (Abstr)
57. Joyce P und Tavaré S (1995) The distribution of rare alleles. *J Math Biol* 33: 602-618
58. Kelly RJ, Ernst LK, Larsen RD, Bryant JG, Robinson JS und Lowe JB (1994) Molecular basis for H blood group deficiency in Bombay (Oh) and para-Bombay individuals. *Proc Natl Acad Sci USA* 91: 5843-5847
59. Kelly RJ, Rouquier S, Giorgi D, Lennon GG und Lowe JB (1995) Sequence and expression of a candidate for the human Secretor blood group  $\alpha(1,2)$ fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the nonsecretor phenotype. *J Biol Chem* 270: 4640-4649

60. Koda Y, Soejima M und Kimura H (1997) Structure and expression of H-type GDP-L-fucose: $\beta$ -D-galactoside 2- $\alpha$ -L-fucosyltransferase gene (FUT1): two transcription start sites and alternative splicing generate several forms of FUT1 mRNA. *J Biol Chem* 272: 7501-7505
61. Koda Y, Soejima M, Liu Y und Kimura H (1996) Molecular basis for secretor type  $\alpha$ (1,2)-fucosyltransferase gene deficiency in a Japanese population: a fusion gene generated by unequal crossover responsible for the enzyme deficiency. *Am J Hum Genet* 59: 343-350
62. Kolb G, Safi F, Beckh K und Beger HG (1997) Klinische Aussagekraft des Tumormarkers CA 19-9 unter besonderer Berücksichtigung des Lewis-Phänotyps. *Med Klin* 92: 228-232
63. Krainer M, Silva-Arrieta S, FitzGerald MG, Shimada A, Ishioka C, Kanamuru R, MacDonald DJ, Unsal H, Finkelstein DM, Bowcock A, Isselbacher KJ und Haber DA (1997) Differential contributions of BRCA1 and BRCA2 to early-onset breast cancer. *N Engl J Med* 336: 1416-1421
64. Kurtzman G, Frickhofen N, Kimball J, Jenkins DW, Nienhuis AW und Young NS (1989) Pure red-cell aplasia of 10 years' duration due to persistent parvovirus B19 infection and its cure with immunoglobulin therapy. *N Engl J Med* 321: 519-523
65. Langley CH, Voelker RA, Leigh Brown AJ, Ohnishi S, Dickson B und Montgomery E (1981) Null allele frequencies at allozyme loci in natural populations of *Drosophila melanogaster*. *Genetics* 99: 151-156
66. Langston AA, Malone KE, Thompson JD, Daling JR und Ostrander EA (1996) BRCA1 mutations in a population-based sample of young women with breast cancer. *N Engl J Med* 334: 137-142
67. Larsen RD, Ernst LK, Nair RP und Lowe JB (1990) Molecular cloning, sequence, and expression of a human GDP-L-fucose: $\beta$ -D-galactoside 2- $\alpha$ -L-fucosyltransferase cDNA that can form the H blood group antigen. *Proc Natl Acad Sci USA* 87: 6674-6678
68. Le Pendu J, Cartron JP, Lemieux RU und Oriol R (1985) The presence of at least two different H-blood-group-related  $\beta$ -D-Gal  $\alpha$ -2-L-fucosyltransferases in human serum and the genetics of blood group H substances. *Am J Hum Genet* 37: 749-760
69. Le Pendu J, Clamagirand-Mulet C, Cartron J-P, Gerard G, Vitrac D und Oriol R (1983) H-deficient blood groups of Reunion Island. III.  $\alpha$ -2-L-fucosyltransferase activity in sera of homozygous and heterozygous individuals. *Am J Hum Genet* 35: 497-507
70. Le Pendu J, Gerard G, Vitrac D, Juszczak G, Liberge G, Rouger P, Salmon C, Lambert F, Dalix A-M und Oriol R (1983) H-deficient blood groups of Reunion Island. II. Differences between Indians (Bombay phenotype) and whites (Reunion phenotype). *Am J Hum Genet* 35: 484-496
71. Le Pendu J, Oriol R, Juszczak G, Liberge G, Rouger P, Salmon C und Cartron JP (1983)  $\alpha$ -2-L-fucosyltransferase activity in sera of individuals with H-deficient red cells and normal H antigen in secretions. *Vox Sang* 44: 360-365

72. Le Van Kim C, Mouro I, Brossard Y, Chavinie J, Cartron JP und Colin Y (1994) PCR-based determination of Rhc and RhE status of fetuses at risk for Rhc and RhE haemolytic disease. *Br J Haematol* 88: 193-195
73. Leo A, Strittmatter M, Rowe G und Roelcke D (1996) Anti-Co(a) in a Colton(a+) typed patient with chronic renal insufficiency. *Infusionsther Transfusionsmed* 23(suppl 3): 36 (Abstr)
74. Levine P, Robinson E, Celano M, Briggs O und Falkenburg L (1955) Gene interaction resulting in suppression of blood group substance B. *Blood* 10: 1100-1108
75. Levine P, Uhlir M und White JA (1961) Ah, an incomplete suppression of A resembling Oh. *Vox Sang* 6: 561-567
76. Mallinson G, Burgess G, Hook S, Murphy M, Poole J und Anstee DJ (1995) A caucasian individual with the Fy(a-b-) phenotype has a novel defect in the Fy gene. *Transfus Med* 5(suppl 1): 19 (Abstr)
77. Maly P, Thall AD, Petryniak B, Rogers CE, Smith PL, Marks RM, Kelly RJ, Gersten KM, Cheng G, Saunders TL, Camper SA, Camphausen RT, Sullivan FX, Isogai Y, Hindsgaul O, von Andrian UH und Lowe JB (1996) The  $\alpha(1,3)$ fucosyltransferase Fuc-TVII controls leukocyte trafficking through an essential role in L-, E-, and P-selectin ligand biosynthesis. *Cell* 86: 643-653
78. Masson P, Palsson B und Andren-Sandberg A (1990) Cancer-associated tumour markers CA19-9 and CA50 in patients with pancreatic cancer with special reference to the Lewis blood cell status. *Br J Cancer* 62: 118-121
79. Metzler M, Gertz A, Sarkar M, Schachter H, Schrader JW und Marth JD (1994) Complex asparagine-linked oligosaccharides are required for morphogenic events during post-implantation development. *EMBO J* 13: 2056-2065
80. Mollicone R, Candelier J-J, Reguigne I, Couillin P, Fletcher A und Oriol R (1994) Molecular genetics of  $\alpha$ -L-fucosyltransferase genes (H, Se, Le, FUT4, FUT5 and FUT6). *Transfus Clin Biol* 1: 91-97
81. Moores PP, Issitt PD, Pavone BG und McKeever BG (1975) Some observations on "Bombay" bloods, with comments on evidence for the existence of two different Oh phenotypes. *Transfusion* 15: 237-243
82. Moulds JM, Nowicki S, Moulds JJ und Nowicki BJ (1996) Human blood groups: incidental receptors for viruses and bacteria. *Transfusion* 36: 362-374
83. Mourant AE, Kopec AC und Domaniewska-Sobczak K (1976) The distribution of the human blood groups and other polymorphisms. 2. Auflage: Oxford University Press, London
84. Mudad R und Telen MJ (1996) Biologic functions of blood group antigens. *Curr Opin Hematol* 3: 473-479
85. Murre C, Schonleber McCaw P und Baltimore D (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56: 777-783

86. Murre C, Schonleber McCaw P, Vaessin H, Caudy M, Jan LY, Jan YN, Cabrera CV, Buskin JN, Hauschka SD, Lassar AB, Weintraub H und Baltimore D (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 58: 537-544
87. Nishihara S, Nakazato M, Kudo T, Kimura H, Nndo T und Harimatsu H (1993) Human  $\alpha$ -1,3 fucosyltransferase (FucT-VI) gene is located at only 13 Kb 3' to the Lewis type fucosyltransferase (FucT-III) gene on chromosome 19. *Biochem Biophys Res Comm* 190: 42-46
88. Olsson ML und Chester MA (1995) A rapid and simple ABO genotype screening method using a novel B/O<sup>2</sup> versus A/O<sup>2</sup> discriminating nucleotide substitution at the ABO locus. *Vox Sang* 69: 242-247
89. Olsson ML und Chester MA (1996) Frequent occurrence of a variant O1 gene at the blood group ABO locus. *Vox Sang* 70: 26-30
90. Oriol R (1990) Genetic control of the fucosylation of ABH precursor chains: evidence for new epistatic interactions in different cells and tissues. *J Immunogen* 17: 235-245
91. Oriol R, Danilovs J und Hawkins BR (1981) A new genetic model proposing that the Se gene is a structural gene closely linked to the H gene. *Am J Hum Genet* 33: 421-431
92. Orita M, Suzuki Y, Sekiya T und Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using polymerase chain reaction. *Genomics* 5: 874-879
93. Rauskolb C, Peifer M und Wieschaus E (1993) extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene pbx1. *Cell* 74: 1101-1112
94. Reguigne-Arnould I, Couillin P, Mollicone R, Faure S, Fletcher A, Kelly RJ, Lowe JB und Oriol R (1995) Relative positions of two clusters of human  $\alpha$ -L-fucosyltransferases in 19q (FUT1-FUT2) and 19p (FUT6-FUT3-FUT5) within the microsatellite genetic map of chromosome 19. *Cytogenet Cell Genet* 71: 158-162
95. Reid ME und Lomas-Francis C (1997) *The Blood Group Antigen Facts Book*. Academic Press, San Diego
96. Rouillac C, Le Van Kim C, Beolet M, Cartron JP und Colin Y (1995) Leu110Pro substitution in the RhD polypeptide is responsible for the DVII category blood group phenotype. *Am J Hematol* 49: 87-88
97. Rouquier S, Lowe JB, Kelly RJ, Fertitta AL, Lennon GG und Giorgi D (1995) Molecular cloning of a human genomic region containing the H blood group  $\alpha$ (1,2)fucosyltransferase gene and two H locus-related DNA restriction fragments. Isolation of a candidate for the human Secretor blood group locus. *J Biol Chem* 270: 4632-4639
98. Saitoh S, Cunningham J, De Vries EMG, McGovern RM, Schroeder JJ, Hartmann A, Blaszyk H, Wold LE, Schaid D, Sommer SS und Kovach JS (1994) p53 gene mutations in breast cancers in midwestern US women: null as well as missense-type mutations are associated with poor prognosis. *Oncogene* 9: 2869-2875

99. Sathe M und Bhatia HM (1976) Bombay (Oh) phenotype in two generations in an Indian family. *Vox Sang* 30: 312-314
100. Schrag D, Kuntz KM, Garber JE und Weeks JC (1997) Decision analysis - effects of prophylactic mastectomy and oophorectomy on life expectancy among women with BRCA1 or BRCA2 mutations. *N Engl J Med* 336: 1465-1471
101. Schrem A und Flegel WA (1996) Comparison of solid-phase antibody screening tests with pooled red cells in blood donors. *Vox Sang* 71: 37-42
102. Shechter Y, Etzioni A, Levene C und Greenwell P (1995) A Bombay individual lacking H and Le antigens but expressing normal levels of  $\alpha$ -2- and  $\alpha$ -4-fucosyltransferases. *Transfusion* 35: 773-776
103. Stoerker J, Hurwitz C, Rose NC, Silberstein LE und Highsmith WE (1996) Heteroduplex generator in analysis of Rh blood group alleles. *Clin Chem* 42: 356-360
104. Stroncek DF, Konz R, Clay ME, Houchins JP und McCullough J (1995) Determination of ABO glycosyltransferase genotypes by use of polymerase chain reaction and restriction enzymes. *Transfusion* 35: 231-240
105. Struewing JP, Abeliovich D, Peretz T, Avishai N, Kaback MM, Collins FS und Brody LC (1995) The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nat Genet* 11: 198-200
106. Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Brody LC und Tucker MA (1997) The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *N Engl J Med* 335: 1401-1408
107. Telen MJ (1995) Erythrocyte blood group antigens: not so simple after all. *Blood* 85: 299-306
108. Telen MJ (1996) Erythrocyte blood group antigens: polymorphisms of functionally important molecules. *Semin Hematol* 33: 302-314
109. Tournamille C, Colin Y, Cartron JP und Le Van Kim C (1995) Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet* 10: 224-228
110. Ugozzoli L und Wallace RB (1992) Application of an allele-specific polymerase chain reaction to the direct determination of ABO blood group genotypes. *Genomics* 12: 670-674
111. Ulevitch RJ, Johnston AR und Weinstein DB (1979) New function for high density lipoproteins. *J Clin Invest* 64: 1516-1524
112. van Dijk MA und Murre C (1994) extradenticle raises the DNA binding specificity of homeotic selector gene products. *Cell* 78: 617-624
113. Voelker RA, Schaffer HE und Mukai T (1980) Spontaneous allozyme mutations in *Drosophila melanogaster*: rate of occurrence and nature of the mutants. *Genetics* 94: 961-968
114. Vorstand und Wissenschaftlicher Beirat der Bundesärztekammer (1995) Leitlinien zur Therapie mit Blutkomponenten und Plasmaderivaten. Deutscher Ärzte-Verlag, Köln
115. Wagner FF und Flegel WA (1997) Polymorphism of the h allele and the population frequency of sporadic nonfunctional alleles. *Transfusion* 37: 284-290

116. Wagner FF, Gassner C, Müller TH, Schönitzer D, Schunter F und Flegel WA (1997) Three molecular structures cause Rhesus D category VI phenotypes with distinct immunohematologic features. Zur Veröffentlichung eingereicht.
117. Wagner FF, Kasulke D, Kerowgan M und Flegel WA (1995) Frequencies of the blood groups ABO, Rhesus, D category VI, Kell, and of clinically relevant high-frequency antigens in South-Western Germany. *Infusionsther Transfusionsmed* 22: 285-290
118. Wang B, Koda Y, Soejima M und Kimura H (1997) Two missense mutations of H type  $\alpha(1,2)$ fucosyltransferase gene (FUT1) responsible for para-Bombay phenotype. *Vox Sang* 72: 31-35
119. Weinstock C, Ullrich H, Hohe R, Berg A, Baumstark M, Northoff H und Flegel WA (1992) Low-density lipoproteins inhibit endotoxin to activate monocytes. *Arterioscler Thromb* 12: 341-347
120. Wissenschaftlicher Beirat der Bundesärztekammer und Bundesgesundheitsamt (1992) Richtlinien zur Blutgruppenbestimmung und Bluttransfusion. Deutscher Ärzte-Verlag, Köln
121. Wissenschaftlicher Beirat der Bundesärztekammer und Paul-Ehrlich-Institut (1996) Richtlinien zur Blutgruppenbestimmung und Bluttransfusion (Hämotherapie). Deutscher Ärzte-Verlag, Köln
122. Wolter LC, Hyland CA und Saul A (1993) Rhesus D genotyping using polymerase chain reaction. *Blood* 82: 1682-1683
123. Wright S (1922) Coefficients of inbreeding and relationship. *Amer Nat* 56: 330-338
124. Yamamoto F, McNeill PD, Yamamoto M, Hakomori S-i, Bromilow IM und Duguid JKM (1993) Molecular genetic analysis of the ABO blood group system: 4. Another type of O allele. *Vox Sang* 64: 175-158
125. Yamamoto F-I, Clausen H, White T, Marken J und Hakomori S-i (1990) Molecular genetic basis of the histo-blood group ABO system. *Nature* 345: 229-233
126. Yamamoto F-I, McNeill PD und Hakomori S-i (1992) Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem Biophys Res Comm* 187: 366-374
127. Yazawa S, Kishi K, Akamatsu S, Oh-Kawara H, Seno T und Okubo Y (1996) A search for the Secretor gene nonsense mutation (G428 to A428) in Japanese nonsecretors. *Transfusion* 36: 286 (Leserzuschrift)
128. Youil R, Kemper BW und Cotton RGH (1995) Screening for mutations by enzyme mismatch cleavage with T4 endonuclease VII. *Proc Natl Acad Sci USA* 92: 87-91
129. Yu LC, Yang YH, Broadberry RE, Chen YH und Lin M (1997) Heterogeneity of the human H blood group  $\alpha(1,2)$ fucosyltransferase gene among para-Bombay individuals. *Vox Sang* 72: 36-40
130. Yunis EJ, Svardal JM und Bridges RA (1969) Genetics of the Bombay phenotype. *Blood* 33: 124-132



## 8. DANKSAGUNG

Mit großem Respekt und Dank erwähne ich die zahlreichen Patienten und Blutspender, die sich für blutgruppenserologische und molekularbiologische Untersuchungen zur Verfügung stellten. Ich danke Herrn Dr. med. Walter E. Hitzler, Mainz, Herrn Dr. med. Rainer Lynen, Göttingen, und Herrn Dr. med. Jürgen Mehler, Emmendingen, für die freundliche Überlassung von Blutproben zur molekularbiologischen Austestung. Herrn Dr. rer. nat. Christoph Gassner und Herrn Primarius Dr. med. Diether Schönitzer, Innsbruck, verdanke ich die Durchführung der *ABO*- und *SE*-Genotypisierung mittels SSP-PCR. Die immer fruchtbare und kollegiale Zusammenarbeit mit Herrn Dr. med. Franz F. Wagner hat manche der hier beschriebenen Untersuchungen erst ermöglicht.

Ganz besonderer Dank gilt Herrn Prof. Dr. med. Bernhard Kubanek für seine engagierte Förderung des wissenschaftlichen Nachwuchses im allgemeinen und meiner Arbeit im besonderen. Ihm verdanke ich einen wesentlichen Teil meiner ärztlichen und wissenschaftlichen Ausbildung. Seine kontinuierliche Unterstützung möchte ich besonders hervorheben und ihm darüber hinaus für die Bereitstellung von Laborräumen für wissenschaftliche Untersuchungen danken.

Besonderer Dank gilt meinen weiteren akademischen Lehrern Herrn Prof. Dr. med. Hinnak Northoff, Universitätsklinik Ulm und DRK-Blutspendezentrale Ulm, Herrn Prof. Dr. med. Bernhard Kornhuber, Universitätsklinik Frankfurt/Main, Herrn Prof. Cornelius Murre, PhD, University of California at San Diego, La Jolla CA, USA und Herrn Prof. Dr. med. Hermann Heimpel, Universitätsklinik Ulm.

Ich danke allen Mitarbeitern meines blutgruppenserologischen und immunhämatologischen Labors in der DRK-Blutspendezentrale Ulm, insbesondere Frau E. Hochgeladen, Frau S. Krotzin, Frau A. Kuner, Frau E. Schmid, Frau H. Wintersinger und Frau O. Zarupski für ihre hervorragende technische Assistenz. Ohne die geduldige sekretarielle Hilfe von Frau S. Grabherr wäre die Arbeit in den letzten Jahren sehr viel mühsamer gewesen. Ich danke ihr ebenso für ihre effiziente Hilfe beim Schreiben der vorliegenden Arbeit. Weitere sekretarielle Unterstützung erfuhr ich von Frau A. Jacobs und Frau H. Barro. Herrn H. Stegmann und Frau B. Töpert gilt mein Dank für die EDV-unterstützte Auswertung der Datenbasis für die blutspenderbezogenen Blutgruppenbestimmungen.

Der Deutschen Forschungsgemeinschaft, dem Deutschen Akademischen Austauschdienst, der Universität Ulm und dem DRK-Blutspendedienst Baden-Württemberg verdanke ich die finanzielle Unterstützung meiner bisherigen wissenschaftlichen Ausbildung und Arbeiten.