Dissertation

Moleculargenetic analysis of the genes of fucosyltransferase 1, 2 and 3 in Styrian blood donors

vorgelegt von

Mag. rer. nat. Eva Maria Matzhold

zur Erlangung des akademischen Grades

„Doctor scientiae medicae“

(Dr. scient. med.)

an der Medizinischen Universität Graz

durchgeführt an der

Univ. Klinik für Blutgruppenserologie und Transfusionsmedizin

unter der Betreuung von

Univ. Prof. Dr. Wolfgang Helmberg

Jänner 2010

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am 11.01.2010

This thesis is based on the following publication in Transfusion journal:

Eva Maria Matzhold et al. Identification of 14 new alleles at the fucosyltransferase 1, 2 and 3 loci in Styrian blood donors, Austria, Transfusion 2009 Oct;49(10):2097-108. Epub 2009 Jun 30.

Acknowledgment

Many thanks to Wolfgang Helmberg, Dr., Professor of Medicine, for supervising this work and for many constructive discussions.

I want to thank Gerhard Lanzer, Professor of Medicine, Head of the department of blood group serology and transfusion medicine for enabling me to do this work at his department.

I should like to express great thanks to Maria Vadon, Dr. of Medicine, and to Thomas Wagner, Professor of Medicine, for their helpful advice and support.

Further I am very grateful to my colleagues Camilla Drexler, Dr. of Medicine and Silvia Ulrich, Mag. rer. nat, who had supported me in many ways.

Many thanks to the team of the HLA laboratory for their technical assistance.

Eugenia Lamont is gratefully acknowledged for editorial assistance.

Finally I want to thank my husband Stephan for encouraging and supporting me at anytime.

Content

1 Introduction

1.1 Biology and Function of the fucosyltransferase genes 1, 2 and 3

The genes for fucosyltranserase 1 (FUT1:H), 2 (FUT2: secretor) and 3 (FUT3: Lewis) are involved in the synthesis of ABO (H) and Lewis blood group antigens. These are expressed on cell surfaces and other tissues including respiratory and digestive mucosa and body secretions.

Biochemically the ABH and Lewis blood group determinants are oligosaccharides, carried by specific types of carbohydrate chains including type 1 (galactose beta1-3 N-acetyl-D-glucosamine) and type 2 (galactose beta1-4 N-acetyl-D-glucosamine) (figure 1). They are synthesized by sequential addition of sugar units by glycosyltransferases, including fucosyltransferases 1, 2 and 3.[1]

1.2 FUT1 and FUT2 and ABO (H) antigen synthesis

FUT1 and FUT2 encode α(1,2)- fucosyltransferases that transfer a terminal fucose residue in an α(1,2)- linkage onto a pre-existing galactose precursor substance to form the H antigen.[2-5]

Then the H determinant is modified by the action of the ABO gene encoded α(1,3)-N-acetylgalactosaminyltransferase or α(1,3)-galactosyltransferase, respectively to form the A or B antigen, respectively. The transferase encoded by the O allele is unfunctional and catalytically inactive. As a result the H antigen remains unchanged on red cells and secretions, determining the blood group O. [5-7]

Different types of H antigen are synthesized. The FUT1 gene locus determines the synthesis of type 2 H antigen which is preferentially expressed in erythroid tissues and vascular endothelial cells. FUT2 regulates the expression of type 1 and type 2 H antigen in secretory tissues like epithelial cells of the digestive and respiratory tracts as precursor of soluble ABH antigens.[5,6,8,9] Hence in individuals who have at least one functional FUT2 allele, their ABH antigens are not only detected on their cell surfaces, but also in their body fluids including saliva. Nonsecretors are homozygous for two inactive FUT2 alleles (se).

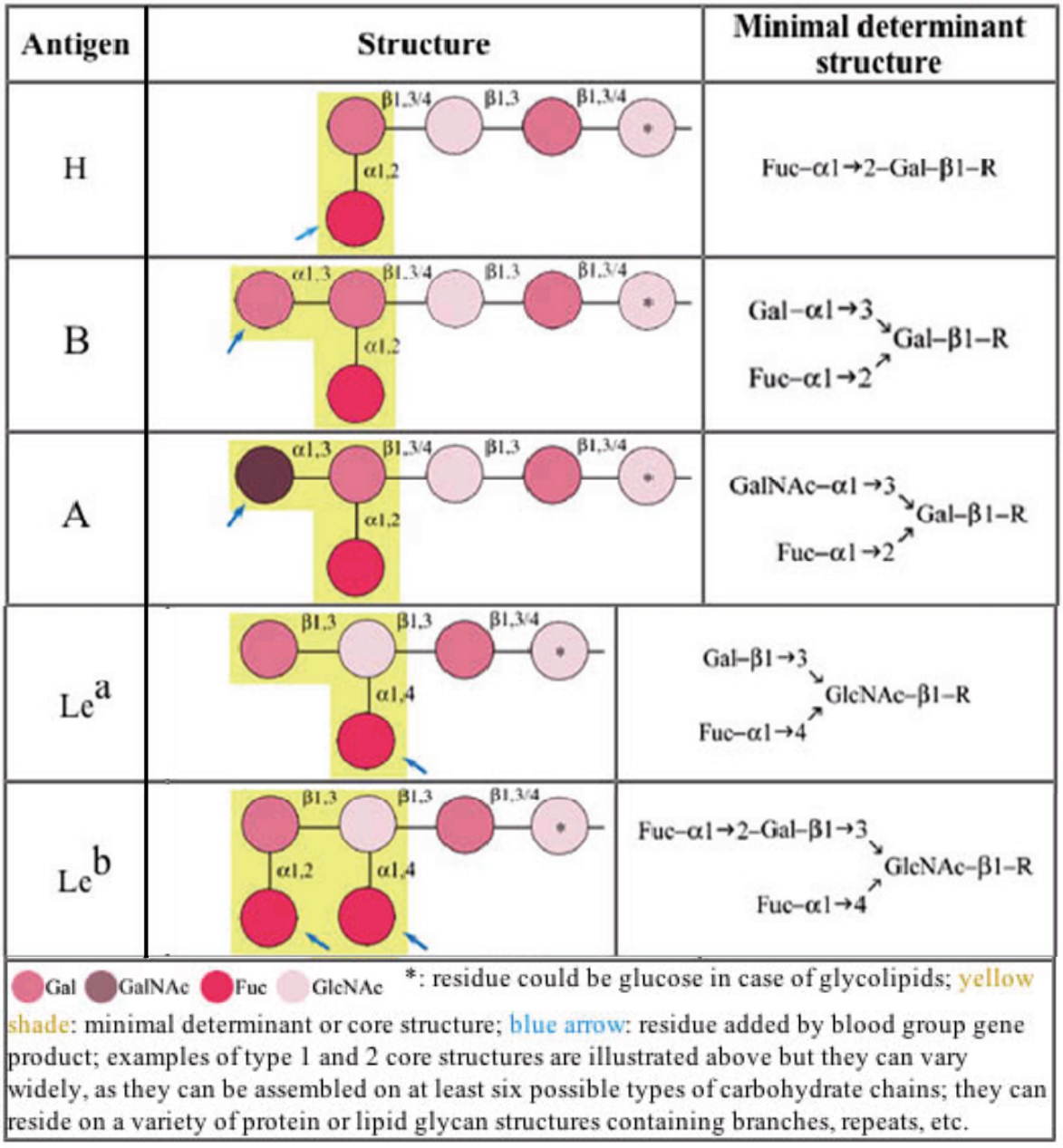

Figure 1: Monosaccharide based structure of ABH and Lewis antigens. (modified from BGMUT / NCBI blood group gene mutation database).[10]

Fuc: Fucose; Gal: D- Galactose; GalNAc: N-acetyl-D-galactosamine; GlcNAc: N-actyl-D-glucosamine;

1.2.1 Bombay and para-Bombay phenotype

Inactive alleles of FUT1 (h) and FUT2 (se), when homozygous, are responsible for either the Bombay or para-Bombay blood group phenotype. Since H antigen is the precursor substrate for both A and B antigens, in Bombay individuals neither A nor B antigen is present on erythrocytes and in secretions. However, in addition to anti-A

and anti-B, anti-H is detected in the plasma. This phenotype was first seen in an Indian individual in Bombay who was deficient of H antigens on red cells, but had antibodies in plasma reacting with all the cells exhibiting the normal red cell ABO phenotype.[11]

Due to the presence of a silenced FUT1 gene together with an active FUT2 gene, Para-Bombay individuals lack H antigens of erythrocytes but can secrete ABH antigens into saliva.[12] Bombay and para-Bombay phenotypes are very rare and vary from an estimation of 1:1000000 in Europe with 1 in 312,081 Germans to 1:1000 on Reunion Island.[10,13,14]

1.3 FUT3

The FUT3 gene encodes an α(1,3/1,4)- fucosyltransferase that is responsible for the last step in the biosynthesis of Lewis antigens. The Lewis determinants are structurally related to the ABO (H) determinants and assembled by sequential addition of specific monosaccharides onto terminal saccharide precursor chains on glycolipids or glycoproteins. Either type 1 or type 2 precursor oligosaccharides can be modified by the Lewis enzyme, resulting in the formation of α(1,3) – or α(1,4)-linkage, respectively.[15,16]

Lewis a and Lewis b antigens are the products when type 1oligosaccharides are used as substrates. The Lewis antigens are adsorbed by the RBCs through circulating glycolipids, determining their Lewis phenotype. They can also be detected in plasma, saliva and other secretions.[6]

1.3.1 Lewis phenotype

The Lewis phenotype of an individual is determined by the genotype at the FUT2 and FUT3 loci (figures 1 and 2). For the production of Lewis b (Le^b) antigen predominantly type 1 H epitope is used as substrate, which requires the activity of FUT2.[6,17] Hence, individuals phenotyped as Le(a-b+) usually have the corresponding secretor status in saliva. In these individuals the FUT2 enzyme is highly active and transforms most of the precursor substance into H type 1, which is in turn transformed into Le^b by the Lewis enzyme.

In contrast, the Le(a+b-) phenotype is found in ABH-nonsecretors with at least one functional FUT3 allele, but homozygous inactive FUT2.

The Le(a+b+) phenotype is almost absent in Caucasians while it occurs with a frequency of 22-25% in the Asian population. This phenotype is suggested to result from a weakening mutation in the FUT2 gene (A385T). Thus the expression of H antigen is decreased; a proportion of precursor chains remains un-substituted and can be transformed into Le^a by the Lewis enzyme. So both, Le^a and $Le^{b,}$ are abundant in secretions and are detected on RBCs.[18,19]

Individuals homozygous for an inactivating mutation in the Lewis gene are typed as Le(a-b-), regardless of the ABH secretor type.

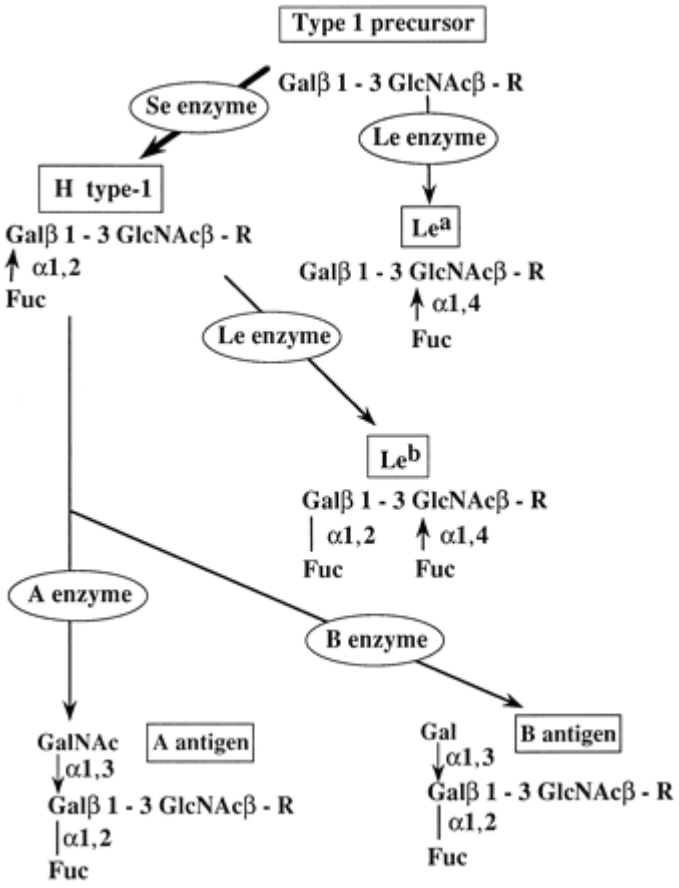

Figure 2: Biosynthetic pathways involved in synthesis of Lewis blood group antigens and secretion of ABH antigens into saliva as it is publicized in the work of Kudo et al. 1996.[18] Se enzyme: FUT2; Le enzyme: FUT3; Fuc: Fucose; A anzyme: N-Acetylgalactosyl-transferase; B enzyme: Galactosyltransferase; Gal: D- Galactose, GalNAc: N-acetyl-D-galactosamine, GlcNAc: N-actyl-D-glucosamin

1.4 The Genes

FUT1 and FUT2 genes are located 35-kb apart on chromosome 19q13.3, sharing almost 70 percent nucleotide sequence identity. [2,8,20,21] In each of the genes the protein coding region is found within a single exon. [22] The fucosyltransferase protein has, in general, the typical topology of type II membrane proteins and is organised in the NH2- terminal hydrophobic cytoplasmic domain, a hydrophobic Golgi-membrane spanning domain (transmembrane, TM) and the COOH-terminal catalytic domain, residing in the lumen of the cellular Golgi apparatus.[23]

FUT1 cDNA predicts a protein consisting of 365-amino acid polypeptide.[3]

In the FUT2 gene there are two isoforms of the protein which can be transcribed due to the presence of two in-frame methionine initiation codons. They differ in the absence or presence of eleven amino acids N-terminal but both, the long and the short isoform comprising 343 and 332 amino acids, respectively, are reported to have the same catalytic activity.[4]

An inactive FUT2 pseudogene (Sec1) which is located 5´of FUT2 within a 30 kb region having a high sequence homology to FUT2, has been described.[4]

Sec1 was reported to be involved in recombinational events, resulting in FUT2 fusion genes.[24,25] As reported, there is a nonfunctional fusion gene which consists of the 5´region of the pseudogene and the 3´region of FUT2, generated by an unequal crossover between the two genes.

Other non-functional fucosyltransferases 2 are based on complete deletion of the FUT2 gene. Since the The 5´and 3´flanking regions of the FUT2 gene contain numerous Alu repeats, an alu-mediated gene recombination mechanism is suggested as reason for FUT2 inactivation [26,27].

There also exists an Sec1-FUT2-Sec1 hybrid allele which has resulted from an interlocus gene conversion.[24]

Due to the localization of FUT1, FUT2 and Sec1 and the results of sequence comparisons it is suggested that these genes originated from two successive duplications. Thereby FUT1 was generated first as ancestor of *FUT*2 and *Sec*1 which have been generated in a second round.[28-30]

The gene encoding FUT3 is placed on the short arm of chromosome 19 (19p13.3). Similarly to FUT1 and FUT2 the protein is encoded within a single exon and comprises 361 amino acids [22]

1.5 The alleles

Alleles of blood group antigen encoding genes are listed in the NCBI´s Blood Group Antigen Gene Mutation Database (dbRBC/BGMUT). There 40 FUT1 gene polymorphisms, primarily responsible for the rare Bombay- or para-Bombay phenotype have been reported.[10,13,14] Recently, two novel enzyme-inactivating mutations were described in an Austrian individual who was homozygous for a new nonfunctional h allele.[31] The allele carried two adjacent single nucleotide polymorphisms (SNPs), 785G>A and 786C>A, resulting in the replacement of serine by lysine in the protein and so causing enzyme inactivation.

FUT2 and FUT3 genes are described as being highly polymorphic.[1] To date, 49 gene variants of FUT2 and 46 of FUT3 have been listed in the dbRBC/BGMUT.[10]

The FUT2 428G>A (W143stop) nonsense mutation, when homozygous, is mainly responsible for the nonsecretor phenotype in European and African populations (47% and 42%, respectively).[4,32] In the FUT3 gene, polymorphisms at nucleotide positions 59, 202, 314 and 1067 are known to explain 90-95% of the Lewis negative phenotype in Caucasians.[33-36]

1.6 Association of ABH and Lewis antigens with diseases

It is reported, that the expression of ABH antigens can change during embryonic development, cell maturation and malignant transformation.[5,37-40] In previous studies it was suggested, that the fucosylated glycans that are the products of FUT1 and FUT2 may serve as ligands in cell adhesion or as receptors for certain microorganisms. The secretor status may play a role in susceptibility to infections, such as Helicobacter pylori, HIV-1 or Norwalk virus.[41-45] Non-secretors who do not express the fucosyltransferase 2 and H type 1 or Le^b antigens, have been shown to be less susceptible or even resistant to infection with Noro Virus [44,45] Also Lewis phenotypes have been discussed to be associated with the prevalence of different diseases, including asthma bronchiale, non-insulin-dependent diabetes mellitus or coronary artery disease.[36,46-48] Alpha 1,3 fucosyltransferases, including FUT3, 6 and 7 are thought to be involved in the regulation of prostate cancer cell trafficking. Thereby Sialyl Lewis x on prostate cancer cells may promote metastasis by mediating the binding to microvascular endothelial via an E-selectin dependend adhesion amechanisms.[49,50]

1.7 Aim of the study

In Europe, various sporadic FUT1 mutations, mainly nonfunctional alleles responsible for the Bombay – or para Bombay phenotype, have been described.

Rather prevalent and population specific polymorphisms including enzyme-inactivating variations have been reported for the FUT2 and FUT3 gene.[32,51-53]

We aimed to investigate the genetic variations and allele frequencies of FUT1, FUT2 and FUT3 coding regions and flanking sequences in Styrian blood donors.

2 Study Design and Methods

The study was approved by the Ethics Committee of the Medical University of Graz and 100 unrelated Styrian blood donors (Caucasians) were included.

The patients were recruited during routine blood donation events on different days, in different villages of Styria.

ABO and Lewis blood groups (BG) were determined with standard serologic techniques and gel matrix techniques (MicroTyping system, DiaMed ID, Cressier, Schwitzerland). Genomic DNA was prepared from peripheral blood leukocytes by magnetic particle technology using a DNA isolation system (GenoM-6, Qiagen, Hilden, Germany).

2.1 Amplification of the coding regions of FUT1, FUT2 and FUT3

The entire coding regions and some 100 base pairs of adjacent intron and 3´untranslated regions (UTR) of FUT1, 2 and 3 were amplified by polymerase chain reaction (PCR). We designed appropriate primers (table 1) for both PCR amplification and sequencing, using an online primer design tool (https://ecom.mwgdna.com). All primers were synthesized by Eurofins MWG Operon, Ebersberg, Germany.

Genomic DNA was amplified in a total volume of 25 µl, using a 2xSuperHot PCR Master Mix (Bioron, Ludwigshafen, Germany), 0.2 µM per L primers and 150 ng DNA. Thermal cycling was carried out in a GeneAmp PCR System 9700 (Applied Biosystems, California, USA): one minute at 94°C, 35 cycles of denaturing for 30 seconds at 92°C, annealing for 50 seconds at 64°C (FUT1) or 56°C (FUT2, FUT3), and elongation for 40 seconds at 72°C. PCR products were analyzed by

electrophoresis in a two percent (w/v) agarose/Tris EDTA gel containing 70 µg ethidium bromide and visualized under UV light.

Table 1: Primer sequences for fucosyltransferase 1, 2 and 3

Primer name	Sequence (5´to 3´)	Purpose	Size of PCR product (bp)
FUT1_Amp_F	CTTACCCCACATCCCTCCTCAGCC	FUT1 amplification, sequencing	1248
FUT1_Amp_R	GAT GCC AGG CCT CTG AAG CCA CG	FUT1 amplification, sequencing	
FUT2_Amp_F	TGCCAAGTATTTACACACCTGAAG	FUT2 amplification, sequencing	1317
FUT2_Amp_R	GATTTCTGTTACTTGCAGCCCA	FUT2 amplification, sequencing	
FUT3_Amp_F	GAAACAGGAATAATAGCAGCTCCTC	FUT3 amplification, sequencing	1357
FUT3_Amp_R	TAGCAGGCAAGTCTTCTGGA	FUT3 amplification, sequencing	
FUT1_F1_R	GGC GTG CGG CTG TCC ACT TC	FUT1 sequencing	
FUT1_F2_F	GATGGG ACAGTATGCCACGCTGC	FUT1 sequencing	
FUT1_F2_R	CGT GCC CGG AAC CAG TCC ATG	FUT1 sequencing	
FUT1_F3_F	CAGCTCCGCCTGGGCCGC	FUT1 sequencing	
FUT2_F1_R	AGTCGTTCAGGTGGTAGTTCTGC	FUT2 sequencing	
FUT2_F2_F	TGTACGCCCTGGCCAAGA	FUT2 sequencing	
FUT2_F2_R	TCAATGTTCTCCCGACACCA	FUT2 sequencing	
FUT2_F3_F	CCATGTCATGCCAAAAGTGTG	FUT2 sequencing	
FUT3_F1_R	TCCAAGTTGAACCAGATCCAG	FUT3 sequencing	
FUT3_F2_F	ACAGGCAGACATGGTCATCGT	FUT3 sequencing	
FUT3_F2_R	TAGTTGCTTCTGCTGGGGC	FUT3 sequencing	
FUT3_F3_F	CAAGTTCTACCTGGCCTTCG	FUT3 sequencing	

The genes for fucosyltransferase 1 (FUT1), fucosyltransferase 2 (FUT2) and fucosyltransferase 3 (FUT3) were analysed. PCR products were screened for genetic variations by direct sequencing of three overlapping fragments (F1-F3). Forward (F) and reverse (R) primers used to amplify the coding regions plus some adjacent 100 base pairs and for sequencing reactions are listed. The size of PCR product when the respective amplification primers (Ampf_F and Amp_R) are used is indicated.

2.2 Sequencing analysis and genotyping

FUT1, FUT2 and FUT3 PCR products were screened for gene variations by direct sequencing of three overlapping fragments (F1-F3) using the dideoxynucleotid chain termination method. The amplified DNA was enzymatically cleaned up with reagent (ExoSAP-IT, USB, Staufen, USA) according to the manufacturer's protocol. Sequence PCR was performed using a cycle sequencing kit (BigDye Terminator v1.1, Applied Biosystems, Foster City, USA). A mastermix with appropriate forward or reverse primers at a final concentration of 0.5µM per L was prepared. Thermocycler protocols were as follows: 60 seconds at 96°C, 25 cycles of 10 seconds at 92°C, and

240 seconds at 60°C (FUT1). For FUT2 and FUT3 sequencing PCR, 10 seconds at 56°C were inserted as second step.

To remove access primers and nucleotides, the sequence-PCR products were purified with a gel filtration method, according to manufacturer's instructions. (Sephadex™ G-50 Superfine, GE Healthcare Bio-Sciences AB, Uppsala, Sweden).

After sample processing (ABI PRISM 3130XL) the sequences were analyzed with SeqScape Version 2.5 computer software (Applied Biosystems). As reference alleles for the target gene sequences of FUT1, FUT2 and FUT3, AC009002.5, AC008888.7, AY870341 were used, respectively.

Regarding FUT2, the positions in this publication refer to the short isoform of the FUT2 protein comprising 332 amino acids.

2.3 Identification of haplotypes (alleles)

For the allelic haplotyping of individuals heterozygous at more than one variant site in FUT1, FUT2 or FUT3 gene, PCR with sequence-specific primers (SSPs) was performed. We designed SSPs that differ by genotypically specific mutations on the last nucleotide at the 3´-end (table 2). For PCR, mastermix protocols were used as described above. Thermocycler protocols were modified according to the different annealing temperatures, optimized with gradient PCR. After analyzing the PCR products on agarose gel, each of the amplified alleles served as template for sequencing reaction as described above. In addition haplotypes were inferred from genotype data using PyPop computer software (Python for Population Genetics).[54] PyPop was designed for population genetic analyses. It implements a number of population genetic tests and statistics for both single locus and multiple loci analysis.

Table 2: Allele specific primers for haplotyping of fucosyltransferase 1, 2 and 3

Primer name	Variant site	Sequence (5´to 3´) of the SSP	Second primer	Annealing temperature
FUT1_SSP-35A_F	-35W	TCCCCTCCTCAGCCTCAGA	FUT1_F1_R	65°C
FUT1_SSP-35T_F	-35W	TCCCCTCCTCAGCCTCAGT	FUT1_F1_R	65°C
FUT1_SSP_9C_R	9Y	AGCTGACGATGGCTCCGG	FUT1_Amp_F	67°C
FUT1_SSP_9T_R	9Y	AGCTGACGATGGCTCCGA	FUT1_Amp_F	65°C
FUT2_SSP-106C_F	-139S	ACACACCTGA AGTAGAAGCAC	FUT2_Amp_R	54°C
FUT2_SSP-106G_F	-139S	ACACACCTGA AGTAGAAGCAG	FUT2_Amp_R	61°C
FUT2_SSP_1011T_R	3´UTR 1011Y	AGA AGG AGA AAA GGT CTC AAAGA	FUT2_Amp_F	63°C
FUT2_SSP_1011C_R	3´UTR 1011Y	AGA AGG AGA AAA GGT CTC AAAGG	FUT2_Amp_F	65°C
FUT3_SSP_41G_F	41R	CACAATGGCCAT GGCG	FUT3_Amp_R	62°C
FUT3_SSP_41A_F	41R	GCCACAATGGCCATGGCA	FUT3_Amp_R	65°C
FUT3_SSP_47C_F	47Y	AATGGCCATGGCGCCGCTC	FUT3_Amp_R	66°C
FUT3_SSP_47G_F	47Y	AATGGCCATGGCGCCGCTG	FUT3_Amp_R	66°C
FUT3_SSP_59T_F	59K	CGC TGT CTG GCC GCA CT	FUT3_F2_R	65.6°C
FUT3_SSP_59G_F	59K	CGC TGT CTG GCC GCA CG	FUT3_F2_R	65.6°C
FUT3_SSP_202C_F	202Y	ACCC TCC TGA TCC TGC TAC	FUT3_Amp_R	64°C
FUT3_SSP_202T_F	202Y	ACCC TCC TGA TCC TGC TAT	FUT3_Amp_R	62°C
FUT3_SSP_314C_R	314Y	CCAGTGGTGCACGATGACCG	FUT3_Amp_F	65°C
FUT3_SSP_314T_R	314Y	CAGTGGTGCACGATGACCA	FUT3_Amp_F	64°C
FUT3_SSP_1067T_R	1067W	AGG TGA ACC AAG CCG CTA	FUT3_Amp_F	64°C
FUT3_SSP_1067A_R	1067W	AGG TGA ACC AAG CCG CTT	FUT3_Amp_F	64°C

PCR was performed with sequence specific primers (SSPs) to haplotype individuals heterozygous at more than one variant site in fucosyltransferase 1 (FUT1), fucosyltransferase 2 (FUT2) or fucosyltransferase 3 (FUT3) gene. The sequences and annealing temperatures of the SSPs are listed. SSPs were designed that differ by genotypically specific mutations on the last nucleotide at the 3´-end. Annealing temperatures were optimized with gradient PCR. Heterozygosis for two nucleotides is indicated according to the IUB code (W:A/T; Y:C/T; S:G/C; R:G/A; K:G/T)[55]. F: Forward, R: Reverse;

2.4 Statistics

Statistical analysis included the calculation of SNP- and haplotype frequencies.

Non-random association of alleles at different loci (linkage disequilibrium FUT1-FUT2) and within individuals (departure from Hardy-Weinberg equilibrium) was tested with PyPop.[54,56] The degree of linkage disequilibrium (LD) was evaluated with the standardized disequilibrium coeffizient D´. D is the difference between the observed and the expected proportion of haplotypes with specific alleles at two loci. D´ranges from -1 (perfect negative allelic association) to +1 (perfect positive allelic association). [57]

The impact of newly identified amino acid substitutions on the protein was predicted with the online-tool PolyPhen (http://coot.embl.de/PolyPhen/)[58]. The prediction is based on straightforward empirical rules that are applied to the sequence, with phylogenetic and structural information characterizing the substitution.

2.5 Allele designation

In this publication we use the allele designation of dbRBC/NCBI.[10] The nomenclature follows the principle of allele nomenclature of HLA alleles:[59] each allele coding for a unique amino acid sequence receives a subsequent number in a first number block, alleles differing only by silent mutations are separated in the second number block, and alleles differing in non-coding regions (UTR, introns) are differentiated in the third number block by a subsequent number. Number blocks are separated by a period (“.”).

3 Results

Three hundred genotypes of three genes were successfully determined in 100 subjects. SNP frequencies were in Hardy-Weinberg equilibrium (HW) with the exclusion of 202T>C in the coding region of FUT3, where there is a departure from HW by a loss of heterozygosis. Detected polymorphisms in each gene are indicated in figure 3. Identified alleles are listed in table 3 (FUT1), 4 (FUT2) and 5 (FUT3). FUT genotypes comprising new nonsynonymous SNPs and the corresponding blood group phenotypes are listed in the order of the sample ID in table 6.

To the best of our knowledge, the SNPs and alleles we identified as novel have not yet been described for human fucosyltransferase genes. The sequences for newly identified alleles were submitted to EMBL Nucleotide Sequence Database (table 7) (http://www.ebi.ac.uk/embl/). If our sequences contained non-redundant information extending previously described alleles, the data were additionally uploaded to EMBL. Following submission to EMBL, the data were incorporated into dbRBC/NCBI.[10]

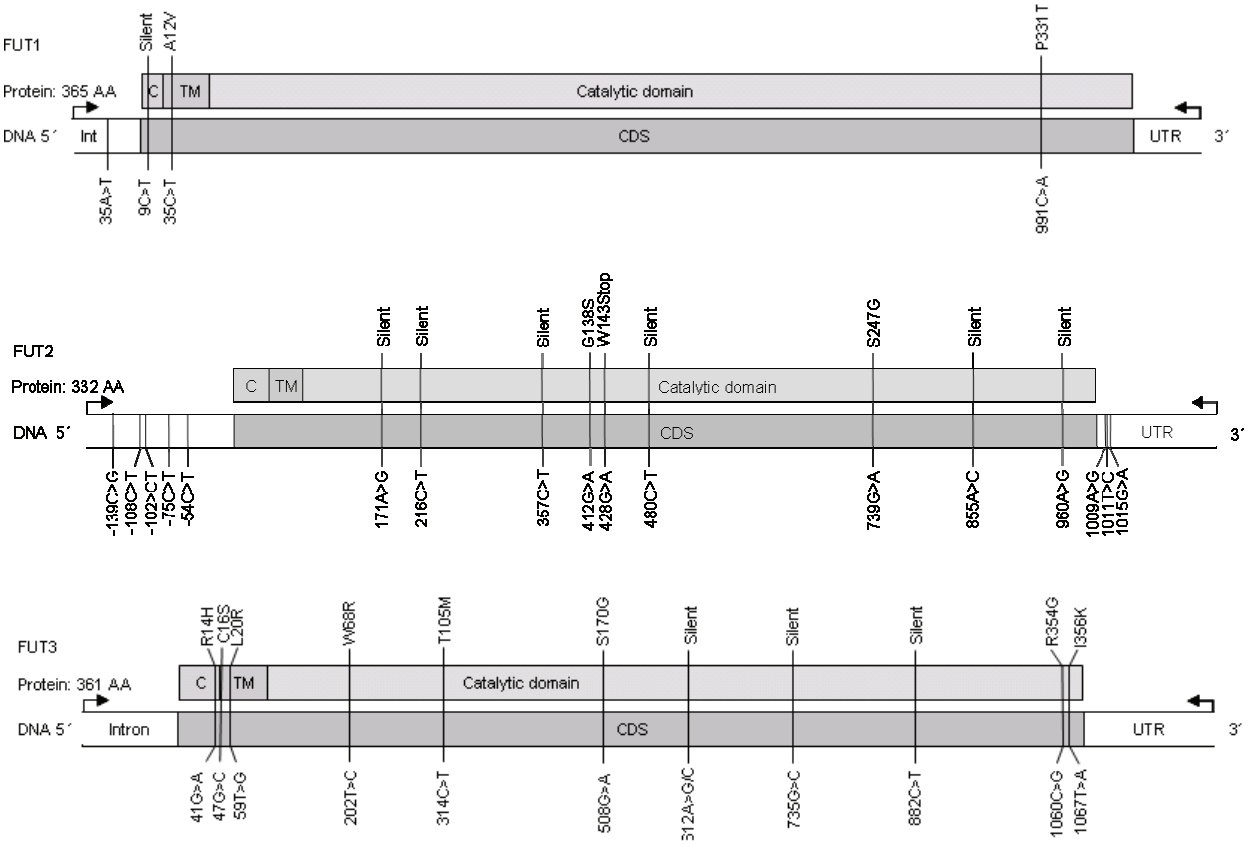

Fig. 3.The protein coding sequence (CDS), adjacent 5´ and 3´untranslated region (UTR) of FUT1, FUT2 and FUT3 genes are shown. Flanking primer binding sites are indicated by arrows. According to UniProt Protein database the corresponding domains of the protein are organized in the NH2- terminal hydrophobic cytoplasmic domain (C), a Golgi-membrane spanning domain (transmembrane, TM) and the COOH-terminal catalytic domain, residing in the lumen of the Golgi apparatus.[23] The sequence variations found are indicated by base substitutions in the DNA and the corresponding amino acid / silent substitutions in the protein. The positions in FUT2 refer to the short isoform of the FUT2 protein, comprising 332 amino acids.UniProt Protein database Accession numbers are P19526 (FUT1), Q10981 (FUT2) and P21217 (FUT3).

3.1 FUT1

In the FUT1 gene we identified five different alleles (table 3). Two of them were new: H*01.02.01 and H*39.01.01, characterized by the novel mutations 9C>T (silent), and 991C>A (P331T), respectively. 991C>A (ID 41, BG B) results in amino acid substitution in the catalytic domain of the translated protein and was predicted by Polyphen analysis to affect the protein function or structure with high probability. The individual carrying the new H*39.01.01 allele also was heterozygous for the FUT1 reference allele H*01.01.01.

The most common allele, detected with a frequency of 66% was H*01.01.01. The second most prevalent allele, H*01.01.03, contained the non-coding -35A>T SNP

and was observed with a frequency of 32.5%. The previously known 35C>T SNP (A12V) in H*11.01.01 was detected in a heterozygous state in one individual. This SNP is not recognized to be FUT1-inactivating and is more common in Chinese individuals.[60]

Table 3: Allele specific sequence variations and frequencies of fucosyltransferase 1

Nucleotide position Reference	-35 A	9 C	35 C	991 C	N=200	F
Rs-number	838137		2071699			
Allele name						
H*01.01.01	A	C	C	C	132	0,66
H*01.01.03	T	C	C	C	65	0,325
H*01.02.01	A	**T**	C	C	1	0,005
H*11.01.01	A	C	T	C	1	0,005
H*39.01.01	A	C	C	**A**	1	0,005
Major SNP frequency	0.675	0.995	0.995	0.995		
Amino acid substitution	N.C.	Silent (L3L)	A12V	P331T		

Different FUT1 alleles that were identified in the study samples are listed in the order of their allele frequency. Indicated nucleotide positions are related to the coding sequence of the gene. The reference nucleotides correlate to H*01.01.01. The Rs-number is the reference single nucleotide polymorphism (SNP) accession ID, according to NCBI's SNP database. Sequence variations are indicated by shaded boxes. Newly identified alleles and SNPs are indicated by boldface. The major SNP frequency is based on the more common nucleotide in the gene. Impacts on the encoded protein are displayed under "amino acid substitution". N.C.: non-coding. N: number of alleles, F: allele frequency

3.2 FUT2

Analysis of the FUT2 gene revealed ten alleles, including five novel alleles (table 4). Three of them were defined by new variants: -54C>T(non-coding) in SE*25.01.03, 412G>A (G149S) in SE*34.01.01, and 1115G>A (non-coding) in SE*01.04.02.

PolyPhen analysis predicted that the nonsynonymous 412G>A mutation, located in the catalytic domain of the protein, would affect the function of the protein. The sample with this variant (ID 40) was typed as Le(a-b-), and was heterozygous for SE*25.01.01 (nonsecretor) and homozygous for the nonfunctional FUT3 allele LE*14.01.01.

Two alleles designated SE*25.01.02 and SE*33.02.01, with novel combinations of previously known SNPs, were detected.

SE*25.01.02 defined a novel nonsecretor allele, based on the presence of the frequent 428G>A FUT2-inactivating mutation. It was observed together with the more prevalent nonsecretor allele SE*25.01.01 in one sample (ID 31), which was typed as Le(a+b-).

Interestingly, SE*33.02.01 was composed of regions identical to the 5´ part of SE*01.04.01 and the 3´part of SE*25.01.01 (figure 2). The two relevant alleles were the most prevalent FUT2 alleles in our study population.

SE*33.01.01 was detected in a heterozygous state in four individuals (ID 14: Le^b; ID 45: Le negative, homozygous for nonfunctional LE*14.01.01; ID 89: Le^b; ID 96: Le negative, homozygous for nonfunctional FUT3 alleles LE*12.01.01 and LE*14.01.01). Further, the common 739G>A (G247S) polymorphism occurred with an SNP frequency of 55% (heterozygous) and 25% (homozygous).

As reported for Caucasians, the most prevalent FUT2 allele was the 428G>A containing SE*25.01.01 allele (nonsecretor, 49.5%).[32] Twenty-four of 100 Styrian blood donors were homozygous, and 53 were heterozygous for 428G>A, corresponding to the reported incidence of approximately 20% of nonsecretor individuals in Caucasian populations.[4]

The second common allele (Se*01.04.01, 29.5%) is represented by the previously reported 357C>T (silent) polymorphism. We also observed the SE*01.05.01 allele (357T, 480G, both silent), with a frequency of 9.5%. This allele had previously been described as being present in Caucasians and Africans.[32]

The reference allele (SE*01.01.01) was detected with a frequency of 7%.

Based on weighted frequencies, an overall linkage disequilibrium between pairwise combinations of FUT1 and FUT2 alleles was apparent from a disequilibrium coefficient[57] of D´=0.52139 ($p<0.001$). The FUT1 H*01.01.01 allele (reference) was linked with the FUT2 SE*25.01.01 (nonsecretor) allele (observed: 90, expected:65).

Table 4: Allele specific sequence variations and frequencies of fucosyltransferase 2

Nucleotide Reference	-139 C	-108 C	-102 G	-75 C	-54 C	171 A	216 C	357 C	412 G	428 G	480 C	739 G	855 A	960 A	1009 A	1011 T	1115 G	N 200	F
Rs-number	67957 4		58823794	58899327		59109937	681343	281377		1800459	1800027	602662				603985			
Allele name																			
SE*25.01.01	G	C	C	T	C	G	T	C	G	A	C	A	A	G	G	C	G	99	0.495
SE*01.04.01	C	C	G	C	C	A	C	T	G	G	C	G	A	A	A	T	G	59	0.295
SE*01.05.01	C	C	G	C	C	A	C	T	G	G	T	G	A	A	A	T	G	19	0.095
SE*01.01.01	C	C	G	C	C	A	C	C	G	G	C	G	A	A	A	T	G	14	0.07
SE*33.02.01	C	C	G	C	C	A	C	T	G	G	C	A	A	G	G	C	G	4	0.02
SE*01.03.01	C	C	G	C	C	A	C	T	G	G	C	G	C	A	A	T	G	1	0.005
SE*01.04.02	C	C	G	C	C	A	C	T	G	G	C	G	A	A	A	T	**A**	1	0.005
SE*25.01.02	G	T	C	T	C	G	T	C	G	A	C	A	A	G	G	C	G	1	0.005
SE*25.01.03	G	C	C	T	**T**	G	T	C	G	A	C	A	A	G	G	C	G	1	0.005
SE*34.01.01	C	C	G	C	C	A	C	T	**A**	G	C	G	A	A	A	T	G	1	0.005
Major SNP frequency	0,505	0,995	0,505	0,505	0,995	0,505	0,505	0,575	0,995	0,505	0,905	0,525	0,995	0,525	0,525	0,525	0,995		
Aminoacid substitution	N.C.	N.C.	N.C.	N.C.	N.C	Silent A57A	Silent Y72Y	Silent N119N	G138S	**W143X**	Silent H160H	G247S	Silent A285A	Silent T320T	N.C..	N.C.	N.C.		

Different FUT2 alleles that were identified in our samples are listed in the order of their allele frequency.
Nucleotide positions indicated are related to the coding sequence of the gene. The reference nucleotides correlate to SE*01.01.01. The Rs-number is the reference single nucleotide polymorphism (SNP) accession ID, according to NCBI's SNP database. Sequence variations are indicated by shaded boxes. Newly identified alleles and SNPs are indicated by boldface. The major SNP frequency is based on the more common nucleotide in the gene and can differ from the reference nucleotide frequency. Impacts on the encoded protein are displayed under “amino acid substitution”.
W143X: Tryptophan is substituted with a stop codon resulting in premature termination of the protein translation and defining a non-secretor allele. The positions indicated refer to the short isoform of the FUT2 protein, comprising 332 amino acids.
N.C.: non-coding. N: number of alleles, F: allele frequency

3.3 FUT3

In the FUT3 gene, seven of thirteen different alleles found were novel (table 5).
Five of them resulted from new mutations located in the coding region:
41G>A (R14H) defined LE*17.01.01 (ID 1, Le^b) and was present in a heterozygous state with the nonfunctional LE*14.01.01 allele.
1060C>G (R354G) in LE*19.01.01 (ID 73, positive for Lewis b) is accompanied by the reference allele LE*01.01.01.
The new mutations 735G>C (LE*01.04.01) and 882C>T (LE*01.03.01) are silent, each of them heterozygously present per individual.
The allele LE*14.02.01 was characterized by the novel arrangement of previously known SNPs (ID 85, Le^b, heterozygous for LE*01.01.01). Thereby the 612A>C (silent) mutation was joined by the common 202C and 314T. At nucleotide position 612 a third variant (612 A>G, silent) was identified, characterizing the novel LE*01.02.01 allele.
Further, the new LE*20.01.01 allele was defined by an isolated appearance of the single SNP 202C (W68R) (ID 72, Le negative, heterozygous for the nonfunctional LE*14.01.01).
The most prevalent allele found in 65% was the LE*01.01.01 (reference). The second most common allele was the nonfunctional LE*14.01.01 (202T>C, 314 C>T), which was detected with a frequency of 21%.

Table 5 Allele specific sequence variations and frequencies of fucosyltransferase 3

Nucleotide position Reference	41 G	47 G	59 T	202 T	314 C	508 G	612 A	735 G	882 C	1060 C	1067 T	N 200	F
Rs-number			28362459			28362464	28362465 (G)				3894326		
Allele name													
LE*01.01.01	G	G	T	T	C	G	A	G	C	C	T	130	0,65
LE*14.01.01 *	G	G	T	C	T	G	A	G	C	C	T	42	0,21
LE*12.01.01 *	G	G	G	T	C	G	A	G	C	C	A	13	0,065
LE*09.01.01 *	G	G	G	T	C	A	A	G	C	C	T	4	0,02
LE*17.01.01	**A**	G	T	T	C	G	A	G	C	C	T	1	0,005
LE*01.02.01	G	G	T	T	C	G	G	G	C	C	T	1	0,005
Le*03.01.01 *	G	G	G	T	C	G	A	G	C	C	T	2	0,01
LE*13.01.01 *	G	C	T	C	T	G	A	G	C	C	T	2	0,01
LE*01.03.01	G	G	T	T	C	G	A	G	**T**	C	T	1	0,005
LE*20.01.01 *	G	G	T	C	C	G	A	G	C	C	T	1	0,005
LE*19.01.01	G	G	T	T	C	G	A	G	C	**G**	T	1	0,005
LE*14.02.01 *	G	G	T	C	T	G	C	G	C	C	T	1	0,005
LE*01.04.01	G	G	T	T	C	G	A	**C**	C	C	T	1	0,005
Major SNP Frequency	0,995	0,99	0,905	0,77	0,775	0,98	0,99	0,995	0,995	0,995	0,935		
Aminoacid substituation	R14H	C16S	L20R	W68R	T105M	G170S	Silent S204S	Silent L245L	Silent D294D	R354G	I356K		

Different FUT3 alleles that were identified in our samples are listed in the order of their allele frequency. The nucleotide positions indicated are related to the coding sequence of the gene. The reference nucleotides correlate to LE*01.01.01. The Rs-number is the reference single nucleotide polymorphism (SNP) accession ID, according to NCBI's SNP database. Sequence variations are indicated by shaded boxes. Newly identified alleles and SNPs are indicated by boldface. The major SNP frequency is based on the more common nucleotide in the gene. Impacts on the encoded protein are displayed under “amino acid substitution”. Alleles including polymorphisms that are described to be enzyme inactivating are flagged by a star. N.C.: non-coding. N: number of alleles, F: allele frequency.

3.4 ABO- and Lewis blood-groups

With blood group analysis, 100 individuals were phenotyped as follows: 41 blood group A, 44 blood group O, 12 blood group B and 3 blood group AB.

The numbers of individuals with RBC phenotype Le(a+b-), Le(a-b+), and Le(a-b-) were 20, 66 and 13, respectively. One haemolysed sample was not tested for the Lewis phenotype, but ABO typing was available from previous results.

All FUT1 gentoypes were functional. FUT2 genotyping was found to be concordant with the serologic Lewis blood group results (table 6). All samples positive for Le^b were at least heterozygous for a previously described functional FUT2 allele.

The prevalence of Lewis genotypes suggesting a lack of FUT3 activity corresponded to the presence of Lewis negative phenotypes, which was 13%.

Six of the individuals serologically typed as Le(a-b-) were homozygous for the FUT3-inactivating allele LE*14.01.01 (202C, 314T), which is common in Caucasian Lewis negative individuals.[35,61] Three of them carried LE*14.01.01 together with the third most common allele, LE*12.01.01 (59G, 1067A), which also was described to be enzyme inactivating. One individual was homozygous for LE*12.01.01. One genotype comprised LE*14.01.01 and a third inactivating allele LE*09.01.01 (59G, 508A) that is common in Asian populations.[62] Another genotype identified in a Lewis-negative individual included the LE*14.01.01 allele, and the rare enzyme-inactivating allele LE*13.01.01 (47C, 202C, 314T).[63] As mentioned above, one Lewis-negative individual (ID 72) was identified to carry the FUT3-inactivating allele Le*14.01.01 as well as the new allele LE*20.01.01 (202C).

Table 6: Genotype and phenotype correlations

Sample ID	FUT1 alleles	FUT2 alleles	FUT3 alleles	ABO	Lewis
1	H*01.01.01	SE*25.01.01	LE*14.01.01	A	a-b+
	H*01.01.03	SE*01.04.01	**LE*17.01.01**		
14	H*01.01.01	SE*25.01.01	LE*01.01.01	B	a-b+
	H*01.01.03	**SE*33.02.01**	LE*01.01.01		
31	H*01.01.01	SE*25.01.01	LE*01.01.01	A	a+b-
	H*01.01.01	**SE*25.01.02**	LE*01.01.01		
40	H*01.01.01	SE*25.01.01	LE*14.01.01	B	a-b-
	H*01.01.03	**SE*34.01.01**	LE*14.01.01		
41	H*01.01.01	SE*01.04.01	LE*01.01.01	B	a-b+
	H*39.01.01	**SE*25.01.03**	LE*01.01.01		
45	H*01.01.01	SE*01.04.01	LE*14.01.01	O	a-b-
	H*01.01.03	**SE*33.02.01**	LE*14.01.01		
72	H*01.01.01	SE*25.01.01	**LE*20.01.01**	A	a-b-
	H*01.01.03	SE*25.01.01	LE*14.01.01		
73	H*01.01.01	SE*01.01.01	LE*01.01.01	O	a-b+
	H*01.01.03	SE*25.01.01	**LE*19.01.01**		
85	H*01.01.01	SE*01.01.01	LE*01.01.01	O	a-b+
	H*01.01.03	SE*25.01.01	**LE*14.02.01**		
89	H*01.01.01	SE*01.04.01	LE*01.01.01	O	a-b+
	H*01.01.01	**SE*33.02.01**	LE*01.01.01		
96	H*01.01.03	SE*01.01.01	LE*12.01.01	A	a-b-
	H*01.01.03	**SE*33.02.01**	LE*14.01.01		

Individuals hetero- or homozygous for new alleles of fucosyltransferase 1 (FUT1), fucosyltransferase 2 (FUT2) or fucosyltransferase 3 (FUT3), including new nonsynonymous mutations are indicated. The two alleles defining the genotype of each locus, and the corresponding ABO blood group and Lewis status are listed. Alleles detected for the first time are indicated by boldface.
H*01.01.01, SE*01.01.01 and LE*01.01.01correlate to respective reference alleles.
Previously described nonfunctional alleles (SE*25.01.01, LE*12.01.01 and LE*14.01.01), as well as the 428G>A (nonsense) carrying alleles are indicated by shaded boxes.

4 Discussion

This study investigated the genetic variations of fucosyltransferase 1, 2 and 3 in Styrian blood donors. In each of the three genes, we found various common alleles and SNPs, as well as some novel sequence variations. The frequencies of common SNPs were consistent with previously reported SNP frequencies in the Caucasian population, whereas the novel mutations appear to be rather sporadic mainly occurring with a frequency <1%.

Two novel mutations were detected in the coding region of FUT1. The 9C>T variant was observed in a heterozygous state in one individual (ID 30) who was phenotyped with BG A. Since it is silent no functional effect on the protein is expected even when present in a homozygous state.
In contrast, the second novel mutation identified, 991C>A (P331T, H*39.01.01), is located in the catalytic domain of the protein. Based on its position and prediction by PolyPhen analysis, 991C>A is most likely to inactivate the FUT1 enzyme. The replacement of the nonpolar amino acid proline by the polar amino acid threonine thereby implies alterations in the structure or binding properties of the protein. H*39.01.01, however, is recognised as being heterozygous in one sample (ID 41) within a functional genotype. Only functional analysis with invitro expression studies could elucidate the impact of 991C>A on the protein level.

Many sequence variations in different populations have been reported at the FUT2 locus.[1,4,18,19,25-27,29,32,51,64-76] In 100 Styrian blood donors, we found various common and novel variants in 10 different alleles.

A novel nonsynonymous SNP was detected in the coding region of the FUT2 gene. 412 G>A (G138S) is present in conjunction with the common silent mutation at nucleotide position 357, defining the new allele SE*34.01.01 (ID 40). Computer based analysis predicts that the mutation located in the catalytic domain of the protein, will have adverse effects on protein function. Since the individual is heterozygous for the common nonsecretor allele, and FUT2 activity can be recognized by the presence of Lewis b antigen, we focused on his Lewis phenotype. In this case, however, the serologically determined Lewis status Le(a-b-) is explained by the nonfunctional FUT3 genotype (homozygous for 202C and 314T), rather than

by the 412G>A variant in SE*34.01.01. The functional status of the new SE*34.01.01 allele so remains unclear. The identification of a new nonsecretor allele can not be excluded.

The newly identified SE*25.01.02 consists of a novel combination of previously known SNPs, including the known 428G>A FUT2 inactivating mutation. Thus, we have identified a second novel nonsecretor allele that is based on the 428G>A nonsense mutation.

Further notable is the discovery of the allele SE*33.02.01 in four of the blood donors. While seven SNPs in the 5´part of the new allele are identical to the SNPs in the 5´sequence of SE*01.04.01, four SNPs in the 3´part correspond to the 3´ sequence of SE*25.01.01. Intragenic recombination between SE*01.04.01 and SE*25.01.01, resulting in the formation of the new recombination-allele is suggested (figure 4). Since there is a homologue region of the two alleles between nucleotide positions 428 and the 739, recombination may have occurred somewhere between these two positions. Recombination breakpoints, however, cannot be located exactly. There is a high prevalence of the involved alleles in our study population (29.5%, 49.5%, respectively), and interallelic recombination may be likely.

With the exception of the nonsynonymous 739G>A SNP, all variants in the new allele are silent or noncoding.

The previously described 739G>A (S247G) [51,77] was always observed together with several other polymorphisms including the FUT2 inactivating 428G>A. The influence of the single mutation 739A on the protein activity, resulting in the replacement of two neutral amino acids, has not been studied so far. We detected the new recombination allele in a genotype (ID 14) showing homozygosis for 739A but heterozygosis for 428A (nonsense). The sample also was typed positive for Le^b antigen, indicating the presence of FUT2 activity. Thus, we suggest the new allele to be functional.

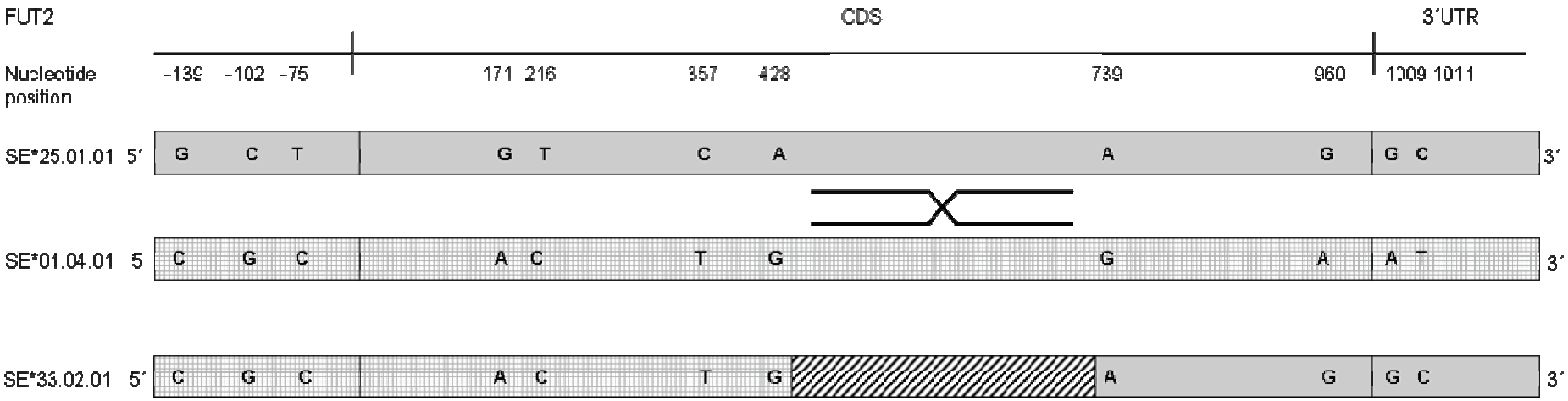

Fig. 4. An intragenic recombination between SE*01.04.01 and SE*25.01.01, resulting in the formation of the new recombination-allele SE*33.02.01 is considered. Polymorph nucleotide positions located in the coding sequence (CDS) and adjacent 5´ and 3´ untranslated region (UTR) are indicated. Seven SNPs in the 5´part of the new allele are identical to the SNPs in the 5´sequence of SE*01.04.01. Four SNPs in the 3´part correspond to the 3´ sequence of SE*25.01.01. Since there is a homologue region between nucleotide positions 428 and the 739 of the involved alleles, recombination breakpoints may be situated somewhere between these two nucleotide positions. The middle panel shows the physical process of recombination, which involves crossing over of DNA strands between the paired alleles. Nucleotide positions refer to the short isoform of the FUT2 protein, comprising 332 amino acids.

In contrast to our data, this nonsynonymous SNP was previously described to be responsible for the nonsecretor phenotype of two individuals of Chinese origin.[71] The individuals who were determined to be nonsecretor by histology of gastric epithelium, were only heterozygous for 428G>A nonsense mutation. Kinetic studies showed the 739G>A mutation to be almost inactivating, but not completely silencing. This discrepancy may be explained by the presence of an additional FUT2 inactivating mutation other than 739A somewhere else in the FUT2 genes (e.g. promoter).

As reported previously, there was a non-random association of FUT1 and FUT2 alleles such that the FUT1 alleles that are linked to alleles of FUT2 encode functional enzymes. Due to the close physical proximity of the two genes, genetic hitch hiking [78] of the two loci may be the reason for their linked appearance. In India the FUT1 mutation 725T>G is almost always present together with a deletion of FUT2, reflecting the Bombay phenotype. As in Reunion Island as well, the major inactivating mutation of FUT1 (349C>T) occurs together with the FUT2 inactivating mutation 428G>A.[14,25]

In the protein encoding region of FUT3, multiple gene variants, including 7 novel alleles and 4 novel SNPs, were found.

Although the novel 41G>A SNP results in amino acid exchange of (R14H) in the cytoplasmic domain of the protein, the mutation is not predicted to affect protein function. This conforms to the corresponding phenotype data of the individual who carried this allele (LE*17.01.01). He was also heterozygous for the non-functional LE*14.01.01 allele, but showed the Le(a-b+) phenotype, indicating LE*17.01.01 to encode a functional FUT3 enzyme.

An additional novel allele (LE*19.01.01) carried the 1060C>G variant. This mutation results in the substitution of arginine (polar, alkaline) to glycine (nonpolar, neutral) at position 354. Although located in the catalytic domain of the protein, the mutation is not predicted to affect the function of the translated protein. This variant is also present in a heterozygous state together with a functional allele, and the impact on the protein activity cannot be determined without further analysis. Notably, a mutation (I356K) situated close to R354G has been shown to inactivate Lewis enzyme activity.[33]

A further Lewis negative individual (ID 72) was haplotyped and identified to carry the non-functional LE*14.01.01 and the newly identified allele LE*20.01.01 (202C,

W68R). The single isolated mutation 202C of the new allele can be assumed to have an enzyme inactivating effect. The 202C and 314T variants have so far not been found separately, except for the rare non-functional le(59-202-1067) allele.[61] The results of invitro expression studies of Elmgren et al. are consistent with our data.[79] They demonstrated the W68R substitution to be the main amino acid change responsible for the Lewis negative phenotype in individuals who are homozygous for both the T202C and C314T mutations.

The lack of functional analysis has to be mentioned as a limitation of our study. We are aware that the prediction of the impact of allelic variants on the protein by computer based analysis has to be interpreted with caution. Predictions by PolyPhen cannot replace the performing of in vitro expression studies.

To conclude, in 100 individuals multiple sequence variations including 14 new alleles of fucosyltransferase 1, 2 and 3 loci were found. Four novel mutations resulted in amino acid substitution in the translated protein. Further, a new FUT2 nonsecretor allele was identified. As reported for other populations, our data also confirm the heterogeneity of fucosyltransferase genes for the Styrian population and may enhance the growing database of blood group antigen gene polymorphisms. Further investigations are needed to evaluate the functional characteristics of the newly found missense mutations.

5 Abstract

Background: Genes for fucosyltransferases 1 (FUT1:H), 2 (FUT2:Secretor) and 3 (FUT3:Lewis) encode enzymes crucial for ABH and Lewis blood group antigen synthesis. They are highly polymorphic and ethnically and geografically specific.

The aim of the present study was to investigate the genetic variations in Styrian blood donors.

Study design and methods: Genetic variations and allele frequencies of FUT1, FUT2 and FUT3 encoding regions and flanking sequences were analyzed in 100 Styrian blood donors by systematic sequencing. Haplotypes were verified with sequence-specific primers. To identify discrepancies, serologically determined ABO– and Lewis blood groups were correlated to respective genotypes.

Results: Two novel FUT1 alleles were defined by 9C>T (silent) and 991C>A (P331T) mutations, the latter located in the catalytic domain of the enzyme.

Three novel FUT2 alleles were detected, characterized by new base substitutions. One of which, 412G>A, also is located in the catalytic domain.

Further, the uncommon arrangement of previously known single-nucleotide-polymorphisms produced a new nonsecretor allele.

Another newly identified FUT2 allele may result from an intragenic crossover event between the two most prevalent FUT2 alleles.

FUT3 analysis revealed seven novel alleles, partly based on the new mutations 41G>A (R14H), 1060C>G (R354G), 735G>C (silent) and 882C>T (silent). While 41G>A is placed in the cytoplasmic domain and functional, 1060C>G is placed in the catalytic domain.

Conclusion: Multiple common and sporadic sequence variations including 14 new alleles at FUT1, FUT2 and FUT3 loci were identified. Four novel mutations result in amino acid substitution in the translated protein. Three of them are predicted to have adverse effects on the enzyme activity. A novel nonsecretor allele containing 428G>A was found. As reported for other populations, our data confirm the heterogeneity of fucosyltransferase genes for the Styrian population and may enhance the growing database of blood group antigen gene polymorphisms.

6 Kurzfassung

Einleitung: Die Fucosylstransferase 1 (FUT1:H), 2 (FUT2:Secretor) und 3 (FUT3:Lewis) codierenden Gene spielen eine bedeutende Rolle in der ABH und Lewis Blutgruppensynthese. Die Gene sind sehr polymorph und spezifisch für unterschiedliche Bevölkerungspopulationen und geografischen Regionen.

Ziel dieser Arbeit war die Erfassung der vorhandenen genetischen Variationen in den Blutgruppen assozierten Genen in den steirischen Blutspendern.

Studien Design und Methoden: In 100 zufällig ausgewählten, unverwandten steirischen Blutspendern wurden genetische Variationen und Allelfrequenzen der FUT1, FUT2 und FUT3 codierenden Regionen mittels direkter Sequenzierung analysiert.

Die in den Blutspendern vorliegenden Haplotypen wurden einerseits aus den Frequenzen der Genotypen mittels der Software PyPop ermittelt und andererseits mittels allelspezifischer PCR bestätigt.

Die Ergebnisse der genetischen Analyse wurden mit den Ergebnissen der serologischen ABO und Lewis Blutgruppenbestimmung korreliert.

Resultate: Zwei neue FUT1 Allele, charakterisiert durch die Nukleotidsubstitutionen (SNPs) 9C>T (synonym) und 991CA (P331T) in der codierenden Region wurden identifiziert. 991C>A ist in der katalytischen Domaine des Fucosyltranserase Proteins lokalisiert.

Im FUT2 Gen wurden aufgrund der Anwesenheit dreier neuer Nukleotidsubstitutionen drei neue Allele identifiziert. Der SNP 412G>A liegt ebenso in der katalytischen Domäne des codierten Proteins.

Ein neues Non-Sekretor Allel, charakterisiert durch eine neue Kombination an schon bekannten Polymorphismen im FUT2 Gen wurde detektiert.

Ein weiteres neu identifiziertes FUT2 Allel ist vermutlich durch ein Crossover-Ereignis zwischen den beiden häufigsten FUT2 Allelen zustande gekommen.

Die Analyse des FUT3 Gens ergab sieben neue Allele, die teilweise auf die neuen SNPs 41G>A (R14H), 1060C>G (R354G), 735G>C (synonym) und 882C>T (synonym) beruhen. Die Variation 41G>A liegt in der zytoplasmatischen Domäine des codierten Proteins und 1060C>G ist in der katalytischen Domäne des Proteins lokalisiert.

Zusammenfassung: Mehrere häufige und sehr verbreitete Polymorphismen, aber auch sehr sporadisch auftretende Mutationen der FUT1, FUT2 und FUT3 Gene wurden identifiziert. Darunter waren 14 noch unbekannte Allele. Vier neue Sequenzvariationen führen zu einem Aminosäureaustausch im translatierten Protein. Drei davon liegen in der katalytischen Domäne des Proteins und könnten die Enzymaktivität verringern. Ein neues FUT2 Allel dass die 428G>A Stop Mutation enthält und somit ein Non-Sekretor Allele ist wurde entdeckt.

Die Ergebnisse dieser Studie bestätigen die Heterogenität der Fucosyltransferase codierenden Gene FUT1, FUT2 und FUT3 in der steirischen Bevölkerung.

Die Daten liefern ergänzende Informationen zum ABO und Lewis Blutgruppensystem und erweitern die wachsenden Datenbanken der Blutgruppenantigen Allel-Polymorphismen.

7 References

1. Koda Y, Soejima M, Kimura H. The polymorphisms of fucosyltransferases. Leg Med (Tokyo) 2001;**3**: 2-14.
2. Rouquier S, Lowe JB, Kelly RJ*, et al.* 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 1995;**270**: 4632-9.
3. Larsen RD, Ernst LK, Nair RP, Lowe JB. 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 U S A 1990;**87**: 6674-8.
4. Kelly RJ, Rouquier S, Giorgi D*, et al.* 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 non-secretor phenotype. J Biol Chem 1995;**270**: 4640-9.
5. Clausen H, Hakomori S. ABH and related histo-blood group antigens; immunochemical differences in carrier isotypes and their distribution. Vox Sang 1989;**56**: 1-20.
6. Watkins WM. Biochemistry and Genetics of the ABO, Lewis, and P blood group systems. Adv Hum Genet 1980;**10**: 1-136, 379-85.
7. Daniels G. Human blood groups. Oxford: Blackwell Science 1995: 8-120.
8. Oriol R, Danilovs J, Hawkins BR. A new genetic model proposing that the Se gene is a structural gene closely linked to the H gene. Am J Hum Genet 1981;**33**: 421-31.
9. Oriol R, Le Pendu J, Mollicone R. Genetics of ABO, H, Lewis, X and related antigens. Vox Sang 1986;**51**: 161-71.
10. Blumenfeld OO, Patnaik SK. Allelic genes of blood group antigens: a source of human mutations and cSNPs documented in the Blood Group Antigen Gene Mutation Database. Hum Mutat 2004;**23**: 8-16.
11. Bhende YM, Deshpande CK, Bhatia HM*, et al.* A "new" blood group character related to the ABO system. Lancet 1952;**1**: 903-4.
12. Kelly RJ, Ernst LK, Larsen RD*, et al.* Molecular basis for H blood group deficiency in Bombay (Oh) and para-Bombay individuals. Proc Natl Acad Sci U S A 1994;**91**: 5843-7.
13. Bhatia HM, Sanghvi LD. Rare blood groups and consanguinity: (Bombay) phenotype. Vox Sang 1962;**7**: 245-8.
14. Wagner FF, Flegel WA. Polymorphism of the h allele and the population frequency of sporadic nonfunctional alleles. Transfusion 1997;**37**: 284-90.
15. Sneath JS, Sneath PH. Transformation of the Lewis groups of human red cells. Nature 1955;**176**: 172.
16. Marcus DM, Cass LE. Glycosphingolipids with Lewis blood group activity: uptake by human erythrocytes. Science 1969;**164**: 553-5.
17. Kukowska-Latallo JF, Larsen RD, Nair RP, Lowe JB. A cloned human cDNA determines expression of a mouse stage-specific embryonic antigen and the Lewis blood group alpha(1,3/1,4)fucosyltransferase. Genes Dev 1990;**4**: 1288-303.
18. Kudo T, Iwasaki H, Nishihara S*, et al.* Molecular genetic analysis of the human Lewis histo-blood group system. II. Secretor gene inactivation by a novel single missense mutation A385T in Japanese nonsecretor individuals. J Biol Chem 1996;**271**: 9830-7.
19. Yu LC, Yang YH, Broadberry RE*, et al.* Correlation of a missense mutation in the human Secretor alpha 1,2-fucosyltransferase gene with the Lewis(a+b+) phenotype: a

potential molecular basis for the weak Secretor allele (Sew). Biochem J 1995;**312 (Pt 2)**: 329-32.
20. Rouquier S, Giorgi D, Bergmann A*, et al.* The fucosyltransferase locus FUT1 maps distal to apolipoprotein loci E and C2 on human chromosome 19. Cytogenet Cell Genet 1994;**66**: 70-1.
21. Koda Y, Soejima M, Wang B, Kimura H. Structure and expression of the gene encoding secretor-type galactoside 2-alpha-L-fucosyltransferase (FUT2). Eur J Biochem 1997;**246**: 750-5.
22. McCurley RS, Recinos A, 3rd, Olsen AS*, et al.* Physical maps of human alpha (1,3)fucosyltransferase genes FUT3-FUT6 on chromosomes 19p13.3 and 11q21. Genomics 1995;**26**: 142-6.
23. The universal protein resource (UniProt). Nucleic Acids Res 2008;**36**: D190-5.
24. Soejima M, Fujihara J, Takeshita H, Koda Y. Sec1-FUT2-Sec1 hybrid allele generated by interlocus gene conversion. Transfusion 2008;**48**: 488-92.
25. Koda Y, Soejima M, Liu Y, Kimura H. 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 1996;**59**: 343-50.
26. Koda Y, Soejima M, Johnson PH*, et al.* An Alu-mediated large deletion of the FUT2 gene in individuals with the ABO-Bombay phenotype. Hum Genet 2000;**106**: 80-5.
27. Pang H, Fujitani N, Soejima M*, et al.* Two distinct Alu-mediated deletions of the human ABO-secretor (FUT2) locus in Samoan and Bangladeshi populations. Hum Mutat 2000;**16**: 274.
28. Saunier K, Barreaud JP, Eggen A*, et al.* Organization of the bovine alpha 2-fucosyltransferase gene cluster suggests that the Sec1 gene might have been shaped through a nonautonomous L1-retrotransposition event within the same locus. Mol Biol Evol 2001;**18**: 2083-91.
29. Oriol R, Candelier JJ, Mollicone R. Molecular genetics of H. Vox Sang 2000;**78 Suppl 2**: 105-8.
30. Abrantes J, Posada D, Guillon P*, et al.* Widespread gene conversion of alpha-2-fucosyltransferase genes in mammals. J Mol Evol 2009;**69**: 22-31.
31. Wagner T, Vadon M, Staudacher E*, et al.* A new h allele detected in Europe has a missense mutationin alpha(1,2)-fucosyltransferase motif II. Transfusion 2001;**41**: 31-8.
32. Liu Y, Koda Y, Soejima M*, et al.* Extensive polymorphism of the FUT2 gene in an African (Xhosa) population of South Africa. Hum Genet 1998;**103**: 204-10.
33. Cooling L, Gu Y. Identification of two new single-nucleotide polymorphisms in FUT3 associated with the Lewis-null phenotype. Transfusion 2003;**43**: 1760-1.
34. Liu TC, Chang JG, Lin SF*, et al.* Lewis (FUT3) genotypes in Taiwanese, Thai, and Filipino populations. Ann Hematol 2000;**79**: 599-603.
35. Pang H, Liu Y, Koda Y*, et al.* Five novel missense mutations of the Lewis gene (FUT3) in African (Xhosa) and Caucasian populations in South Africa. Hum Genet 1998;**102**: 675-80.
36. Salomaa V, Pankow J, Heiss G*, et al.* Genetic background of Lewis negative blood group phenotype and its association with atherosclerotic disease in the NHLBI family heart study. J Intern Med 2000;**247**: 689-98.
37. Fujitani N, Liu Y, Toda S*, et al.* Expression of H type 1 antigen of ABO histo-blood group in normal colon and aberrant expressions of H type 2 and H type 3/4 antigens in colon cancer. Glycoconj J 2000;**17**: 331-8.
38. Hakomori S. Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens. Adv Cancer Res 1989;**52**: 257-331.

39. Szulman AE. The ABH blood groups and development. Curr Top Dev Biol 1980;**14**: 127-45.
40. Mejias-Luque R, Lopez-Ferrer A, Garrido M*, et al.* Changes in the invasive and metastatic capacities of HT-29/M3 cells induced by the expression of fucosyltransferase 1. Cancer Sci 2007;**98**: 1000-5.
41. Kindberg E, Hejdeman B, Bratt G*, et al.* A nonsense mutation (428G-->A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection. AIDS 2006;**20**: 685-9.
42. Oba-Shinjo SM, Uno M, Ito LS*, et al.* Association of Lewis and Secretor gene polymorphisms and Helicobacter pylori seropositivity among Japanese-Brazilians. J Gastroenterol 2004;**39**: 717-23.
43. Lindesmith L, Moe C, Marionneau S*, et al.* Human susceptibility and resistance to Norwalk virus infection. Nat Med 2003;**9**: 548-53.
44. Carlsson B, Kindberg E, Buesa J*, et al.* The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 Norovirus infection. PLoS One 2009;**4**: e5593.
45. Bucardo F, Nordgren J, Carlsson B*, et al.* Pediatric norovirus diarrhea in Nicaragua. J Clin Microbiol 2008;**46**: 2573-80.
46. Ellison RC, Zhang Y, Myers RH*, et al.* Lewis blood group phenotype as an independent risk factor for coronary heart disease (the NHLBI Family Heart Study). Am J Cardiol 1999;**83**: 345-8.
47. Kauffmann F, Frette C, Pham QT*, et al.* Associations of blood group-related antigens to FEV1, wheezing, and asthma. Am J Respir Crit Care Med 1996;**153**: 76-82.
48. Djousse L, Karamohamed S, Herbert AG*, et al.* Fucosyltransferase 3 polymorphism and atherothrombotic disease in the Framingham Offspring Study. Am Heart J 2007;**153**: 636-9.
49. Barthel SR, Wiese GK, Cho J*, et al.* Alpha 1,3 fucosyltransferases are master regulators of prostate cancer cell trafficking. Proc Natl Acad Sci U S A 2009;**106**: 19491-6.
50. Barthel SR, Gavino JD, Wiese GK*, et al.* Analysis of glycosyltransferase expression in metastatic prostate cancer cells capable of rolling activity on microvascular endothelial (E)-selectin. Glycobiology 2008;**18**: 806-17.
51. Koda Y, Ishida T, Tachida H*, et al.* DNA sequence variation of the human ABO-secretor locus (FUT2) in New Guinean populations: possible early human migration from Africa. Hum Genet 2003;**113**: 534-41.
52. Soejima M, Nakajima T, Fujihara J*, et al.* Genetic variation of FUT2 in Ovambos, Turks, and Mongolians. Transfusion 2008;**48**: 1423-31.
53. Soejima M, Koda Y. Molecular mechanisms of Lewis antigen expression. Leg Med (Tokyo) 2005;**7**: 266-9.
54. Lancaster AK, Single RM, Solberg OD*, et al.* PyPop update--a software pipeline for large-scale multilocus population genomics. Tissue Antigens 2007;**69 Suppl 1**: 192-7.
55. Nomenclature Committee of the International Union of Biochemistry (NC-IUB). Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984. Biochem J 1985;**229**: 281-6.
56. Lancaster A, Nelson MP, Meyer D*, et al.* PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. Pac Symp Biocomput 2003: 514-25.
57. Terwilliger J OJ. Handbook for human genetic linkage: Johns Hopkins University Press, Baltimore, 1994.
58. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002;**30**: 3894-900.

59. Marsh SG, Albert ED, Bodmer WF*, et al.* Nomenclature for factors of the HLA system, 2004. Tissue Antigens 2005;**65**: 301-69.
60. Xu XG, Hong XZ, Wu JJ*, et al.* [Study on the frequency of alpha-1,2-fucosyltransferase gene h4 allele (C35T) in Chinese population]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 2005;**22**: 657-60.
61. Elmgren A, Rydberg L, Larson G. Genotypic heterogeneity among Lewis negative individuals. Biochem Biophys Res Commun 1993;**196**: 515-20.
62. Liu Y, Koda Y, Soejima M*, et al.* PCR analysis of Lewis-negative gene mutations and the distribution of Lewis alleles in a Japanese population. J Forensic Sci 1996;**41**: 1018-21.
63. Soejima M, Kimura H, Koda Y. Two novel FUT3 alleles responsible for Lewis-null phenotypes in Sri Lanka. Transfusion 2004;**44**: 1534-5.
64. Chang JG, Yang TY, Liu TC*, et al.* Molecular analysis of secretor type alpha(1,2)-fucosyltransferase gene mutations in the Chinese and Thai populations. Transfusion 1999;**39**: 1013-7.
65. Henry S, Mollicone R, Fernandez P*, et al.* Homozygous expression of a missense mutation at nucleotide 385 in the FUT2 gene associates with the Le(a+b+) partial-secretor phenotype in an Indonesian family. Biochem Biophys Res Commun 1996;**219**: 675-8.
66. Henry S, Mollicone R, Lowe JB*, et al.* A second nonsecretor allele of the blood group alpha(1,2)fucosyl-transferase gene (FUT2). Vox Sang 1996;**70**: 21-5.
67. Liu YH, Koda Y, Soejima M*, et al.* The fusion gene at the ABO-secretor locus (FUT2): absence in Chinese populations. J Hum Genet 1999;**44**: 181-4.
68. Mitani T, Akane A. Detection of single nucleotide polymorphism in the FUT2 gene by temperature gradient gel electrophoresis. Leg Med (Tokyo) 2005;**7**: 217-21.
69. Park KU, Song J, Han KS, Kim JQ. The fusion allele of the FUT2 (secretor type alpha(1,2)-fucosyltransferase) gene at a high frequency and a new se385 allele in a Korean population. Ann Hematol 2005;**84**: 656-60.
70. Peng CT, Tsai CH, Lin TP*, et al.* Molecular characterization of secretor type alpha(1, 2)-fucosyltransferase gene deficiency in the Philippine population. Ann Hematol 1999;**78**: 463-7.
71. Serpa J, Mendes N, Reis CA*, et al.* Two new FUT2 (fucosyltransferase 2 gene) missense polymorphisms, 739G-->A and 839T-->C, are partly responsible for non-secretor status in a Caucasian population from Northern Portugal. Biochem J 2004;**383**: 469-74.
72. Soejima M, Koda Y. Denaturing high-performance liquid chromatography-based genotyping and genetic variation of FUT2 in Sri Lanka. Transfusion 2005;**45**: 1934-9.
73. Yu LC, Broadberry RE, Yang YH*, et al.* Heterogeneity of the human Secretor alpha(1,2)fucosyltransferase gene among Lewis(a+b-) non-secretors. Biochem Biophys Res Commun 1996;**222**: 390-4.
74. Yu LC, Chu CC, Chan YS*, et al.* Polymorphism and distribution of the Secretor alpha(1,2)-fucosyltransferase gene in various Taiwanese populations. Transfusion 2001;**41**: 1279-84.
75. Yu LC, Lee HL, Chu CC*, et al.* A newly identified nonsecretor allele of the human histo-blood group alpha(1,2)fucosyltransferase gene (FUT2). Vox Sang 1999;**76**: 115-9.
76. Pang H, Koda Y, Soejima M*, et al.* Polymorphism of the human ABO-Secretor locus (FUT2) in four populations in Asia: indication of distinct Asian subpopulations. Ann Hum Genet 2001;**65**: 429-37.

77. Yip SP, Lai SK, Wong ML. Systematic sequence analysis of the human fucosyltransferase 2 (FUT2) gene identifies novel sequence variations and alleles. Transfusion 2007;**47**: 1369-80.
78. Barton NH. Genetic hitchhiking. Philos Trans R Soc Lond B Biol Sci 2000;**355**: 1553-62.
79. Elmgren A, Mollicone R, Costache M*, et al.* Significance of individual point mutations, T202C and C314T, in the human Lewis (FUT3) gene for expression of Lewis antigens by the human alpha(1,3/1,4)-fucosyltransferase, Fuc-TIII. J Biol Chem 1997;**272**: 21994-8.

8 Appendix

Table 7: Accession numbers of fucosyltransferase 1, 2 and 3 alleles

Gene	Allele name	Accession numbers	Comment
FUT1	H*01.02.01	FM180557	Novel
FUT1	H*39.01.01	FM162557	Novel
FUT2	SE*33.02.01	FM164934	Novel
FUT2	SE*25.01.02	FM162558	Novel
FUT2	SE*01.03.01	FM180564	Extension of previously published allele
FUT2	SE*01.05.01	FM180563	Extension of previously published allele
FUT2	SE*01.04.01	FM180562	Extension of previously published allele
FUT2	SE*25.01.03	FM180561	Novel
FUT2	SE*01.04.02	FM180560	Novel
FUT2	SE*25.01.01	FM180559	Extension of previously published allele
FUT2	SE*34.01.01	FM180558	Novel
FUT3	LE*19.01.01	FM164937	Novel
FUT3	LE*01.03.01	FM164936	Novel
FUT3	LE*01.02.01	FM164935	Novel
FUT3	LE*13.01.01	FM162559	Extension of previously published allele
FUT3	LE*01.04.01	FM180567	Novel
FUT3	LE*17.01.01	FM180566	Novel
FUT3	LE*14.02.01	FM210025	Novel
FUT3	LE*20.01.01	FM210024	Novel

New alleles of fucosyltransferase 1 (FUT1), fucosyltransferase 2 (FUT2) and fucosyltransferase 3 (FUT3) were submitted to EMBL Nucleotide sequence Database (http://www.ebi.ac.uk/embl/) and specific accession numbers were produced.

Additionally, if our sequences contained non-redundant information enhancing previously described alleles, the data were uploaded to EMBL. The allele designation dbRBC/NCBI was used to name the newly identified alleles.[10]