Glycoconj J (2010) 27:61–68
DOI 10.1007/s10719-009-9255-8

Infection-associated *FUT2* (Fucosyltransferase 2) genetic variation and impact on functionality assessed by *in vivo* studies

Lara M. Silva · Ana S. Carvalho · Patrice Guillon · Susana Seixas · Maria Azevedo · Raquel Almeida · Nathalie Ruvoën-Clouet · Celso A. Reis · Jacques Le Pendu · Jorge Rocha · Leonor David

Received: 5 June 2009 / Revised: 10 August 2009 / Accepted: 11 August 2009 / Published online: 16 September 2009

Abstract The secretor (Se)/nonsecretor (se) histo-blood group variation depends on the action of the FUT2 enzyme and has major implications for human susceptibility to infections. To characterize the functionality of FUT2 variants, we assessed the correlation between saliva phenotypes and sequence variation at the *FUT2* gene in sixty seven individuals from northern Portugal. While most non-secretor haplotypes were found to carry the 428G > A nonsense mutation in association with a 739G > A missense substitution, we have also identified a recombinant haplotype carrying the 739*A allele together with the efficient 428*G variant in individuals with the Se phenotype. This finding suggested, in contrast to previous results, that the 739*A allele encodes an efficient Se allele. To test this hypothesis we evaluated the *in vivo* enzyme activity of full coding expression constructs in transient transfection of CHO-K1 cells using FACS (fluorescence-activated cell sorting) analysis and expression of type 2 and type 3 chain H structures as read out. We detected FUT2 activity for the 739*A expression construct, demonstrating that the 739G > A substitution is indeed not inactivating. In accordance with the hypothesis that *FUT2* is under long standing balancing selection, we estimated that the time depth of *FUT2* global genetic variation is as old as 3 million years. Age estimates of specific variants suggest that the 428G > A mutation occurred at least 1.87 million years ago while the 739G > A substitution is about 816,000 years old. The 385A > T missense mutation underlying the non-secretor phenotype in East Asians appears to be more recent and is likely to have occurred about 256,000 years ago.

Keywords α-1,2-fucosyltransferase (FUT2) · Secretor status · Sequence variation · Mutation age · Enzyme activity

L. M. Silva · A. S. Carvalho · S. Seixas · M. Azevedo · R. Almeida · C. A. Reis · J. Rocha · L. David (✉)
IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP),
Rua Dr. Roberto Frias s/n,
4200-465 Porto, Portugal
e-mail: ldavid@ipatimup.pt

P. Guillon · N. Ruvoën-Clouet · J. Le Pendu
Inserm, Université de Nantes,
U892 Nantes, France

R. Almeida · C. A. Reis · L. David
Medical Faculty of the University of Porto,
Porto, Portugal

N. Ruvoën-Clouet
Ecole Nationale Vétérinaire de Nantes,
Nantes, France

J. Rocha
Department of Zoology and Anthropology, Faculty of Sciences,
University of Porto,
Porto, Portugal

Introduction

Glycoconjugates are biosynthesized by the sequential action of glycosyltransferases and genetic variation on the genes that code for the glycosyltransferases generates inter-individual diversity on cell surface glycoconjugates. A disease-related variation occurs on the Lewis system, where polymorphisms of the fucosyltransferase 2 (FUT2) enzyme underlie the secretor (Se) or nonsecretor (se) phenotypes, according to the capability or incapability, respectively, to produce a biosynthetically active enzyme [1]. Several studies have shown that the Se/se variation has relevant

implications in human infections. For example, nonsecretors are known to be virtually resistant to infection by the prototype strain (Norwalk virus) of norovirus [2, 3] as well as by other strains [4, 5]. Inversely, women with the nonsecretor phenotype are more prone to *Escherichia coli*-mediated recurrent urinary tract infections than secretor women [6]. Recently, it has been suggested that BabA-positive *Helicobacter pylori* adhesion/infection is secretor-dependent both in humans and Rhesus monkeys [7, 8]. A further indication of the functional relevance of the Se/se variation is the increasing evidence that the *FUT2* gene has undergone non neutral evolution, suggestive of an important role in the host/pathogens arms race [9–11]. Proper characterization of efficiency of allelic variation at the *FUT2* locus is therefore of utmost importance.

In European and African populations the nonsecretor phenotype has similar frequencies of about 20% and is essentially due to homozigosity for the same null allele, caused by a 428G > A nonsense mutation at codon 143 (W143X) generating an inactive enzyme [12, 13]. In eastern Asians, nonsecretors have similar frequencies to Europeans and Africans, but are homozygous for a different, weak-activity allele resulting from a 385A > T missense mutation at codon 129 (I129F) [14, 15]. In addition, various other inactive or weak alleles with lower frequencies have been described and can be found at http://www.ncbi.nlm.nih.gov/gv/mhc/xslcgi.cgi?cmd=bgmut/hom [16]. We have previously reported that, in the Portuguese population, two FUT2 polymorphisms, 739G > A at codon 247 (G247S) and 839 T > C at codon 280 (F280S), are associated with decreased or absent FUT2 enzyme activity, respectively, in an *in vitro* assay [17].

In the present study, we characterized the sequence variation at the *FUT2* gene in a Portuguese sample and found that the 739G > A mutation is associated to the secretor phenotype in saliva, suggesting that the 739*A variant is not inactivating *in vivo*. In order to evaluate the possibility that the *FUT2* 739*A allele was affecting enzyme specificity to other acceptor substrates, we used CHO-K1 cells that offer the unique possibility of studying type 2 and type 3 substrate acceptors and confirmed the presence of enzyme activity for the 739*A expression construct. Together the data unequivocally demonstrate that the 739G > A substitution is not inactivating and that classical enzyme assays need to be confirmed in the *in vivo* situation.

Materials and methods

Population

We analyzed a random subset of 99 samples from a previous survey comprising 460 workers from a shipyard in northern Portugal, who donated blood samples and performed upper digestive endoscopies [7, 18, 19]. Saliva was collected from 67 individuals. All individuals gave informed consent and the study was approved by the ethical committee of "Hospital S. João".

PCR amplification and sequencing of the coding region of *FUT2*

Genomic DNA was extracted from blood cells. The entire coding region (exon 2) of *FUT2* gene, encompassing 1032 bp, was amplified by PCR and then sequenced in all 99 individuals. PCR was performed using forward primer: 5′-CCATCTCCCAGCTAACGTGTCC-3′ and reverse primer: 5′-GGGAGGCAGAGAAGGAGAAAAGG-3′. PCR was performed in 25 μL of reaction mixtures comprising 1 μL of DNA, 2.5 μL of 10 × *Taq* polymerase buffer, 25 pmol of each primer, 1 μL of dNTPs (10 mM) and 0.2 μL of *Taq* polymerase (Invitrogen). The conditions used in the reaction were as follows: 10 min of initial denaturation at 96°C, followed by 35 cycles of denaturation at 94°C for 20 s, annealing at 59°C for 30 s and extension at 72°C for 90 s, with a final extension at 72°C for 7 min. The resultant PCR products were verified by electrophoresis in 1% (*w*/*v*) agarose/TBE (90 mM Tris/85 mM boric acid/0.5 M Na_2EDTA, pH 8.0) gel using an operative voltage of 180 V. The products were then extracted using a GFX PCR DNA and Gel Band Purification Kit (GE Healthcare).

The PCR products were used as templates for the sequencing reactions. Each reaction was performed in a final volume of 10 μL, which comprised 3 μL of DNA, 3 pmol of primer (mentioned above) and 4 μL of TRR (terminator ready reaction) mix (ABI Prism® 3100 Sequencer; Applied Biosystems). The temperature profile used in sequencing PCR was 95°C for 2 min, followed by 30 cycles of denaturation at 95°C for 45 s, annealing at 50°C at 30 s and extension at 60°C for 4 min. The DNA sequence was then analyzed in the ABI Prism® sequencer.

Haplotype-based data analysis

Haplotypes were statistically inferred from the genotype data by using the program PHASE, version 2.1.1 [20, 21]. Haplotype networks were constructed with the NETWORK 4.5 software (http://www.fluxus-engineering.com/) using the median-joining algorithm [22]. To provide a temporal framework for the phylogenetic relationships among haplotypes and to estimate coalescent times and ages of relevant mutations, we used GENETREE version 9.0 [23], after removing rare recombinant haplotypes from the sequence data. The mutation rate per gene per generation was estimated from the average number of nucleotide substitu-

tions per site between human and chimpanzee reference sequences, calculated with DnaSP v.4.0 [24]. Time estimates in generations were converted into years using a 25 year generation time. Human/Chimpanzee divergence was assumed to have occurred 5 million years ago [25].

Expression vectors for wild-type and mutant *FUT2* variants

To assess the enzyme activity of wild-type (*wt*) and polymorphic variants of *FUT2* gene, expression vectors were assembled, as described previously [17]. Four vectors were obtained using pcDNA3.1: *FUT2wt (FUT2-428G-739G-839 T)*, *FUT2*-739G→A, *FUT2*-839 T→C and pcDNA3.1 empty vector. The sequence of all the constructs was confirmed by direct sequencing. *FUT2wt* and *FUT2*-839 T→C were used as positive and negative controls, respectively.

In vivo activity of α-1,2-Fucosyltransferase variants

For the α-1,2-fucosyltransferase activity assays *in vivo*, confluent CHO-K1 cells were transfected with the expression vectors described above. CHO-K1 cells were cultured in RPMI 1640, supplemented with 10% fetal calf serum, 2 mM L-glutamine, free nucleotides (10 µg/ml), 100 U/ml penicillin, and 100 mg/ml streptomycin (Gibco, Paisley, UK). They were cultured at confluence after dispersal with 0.025% trypsin and 0.02% EDTA. Cells were routinely checked for mycoplasma contamination using Hoechst 33258 (Sigma, St Louis, MO) labeling. Following amplification and purification with the Miniprep Qiagen kit, 8 µg of the recombinant pcDNA3.1 constructs were cotransfected with 4 µg of the pEGFP-N3 control vector (Invitrogen) into CHO-K1 cells using Lipofectamin 2000 (Invitrogen) according to the manufacturer's instructions. Forty-eight hours following transfection, cells were collected and the presence of α-1,2-linked fucose residues was tested by flow cytometry using the UEA-I lectin (*Ulex europaeus* agglutinin I, Vector Labs) and the MBr1 monoclonal antibody (Alexis Biochemicals) that detect the H type 2 (Fucα2Galβ4GlcNAcβ-R) and H type 3 (Fucα2Galβ3GalNAcα-R) motifs, respectively. To this aim, 2.5×10^5 viable transfected cells were incubated in the presence of either biotin-labeled lectin at 5 µg/ml (Vector Labs) or the MBr1 monoclonal antibody at 5 µg/ml for 20 min at 4°C. After 3 washings with PBS (phosphate buffered saline), cells were incubated under the same conditions in the presence of either phycoerythrin-conjugated streptavidin (BD Biosciences) at a 1/40 dilution or a Cy5-conjugated anti-mouse IgG (BD Biosciences) at a 1/500 dilution. After 3 more washings with PBS, cell fluorescence was measured on a FACScalibur flow cytometer (Becton-Dickinson) and analyzed using the CellQuest program (Becton-Dickinson). The transfected protein expression was quantified by the GFP fluorescence recorded on the FL1 channel. EFGP fluorescence was recorded on the FL1 channel, while the presence of the H type 2 and H type 3 motifs was detected on the FL2 and FL4 channels, respectively.

Lewis b and Ulex (secretor) phenotypes in saliva

Lewis b and Ulex levels in saliva were evaluated as previously described [7]. Briefly, saliva collected in sterile tubes was boiled and centrifuged and supernatant stored at −70°C and later used for coating NUNC immunoplates at 1/1000 in 100 mM/L carbonate-bicarbonate buffer by overnight incubation at 37°C. Anti-Lewis b (Le^b)-specific monoclonal antibody 2-25Le, a kind gift from Dr. J. Bara (CNRS, Villejuif, France) was used, followed by peroxidase anti-mouse immunoglobulin (Uptima: Interchim) in 5% milk/PBS at 1/2000 for 1 h at 37°C. Reactions were developed with TMB (3, 3′, 5, 5′-tetramethylbenzidine, BD Biosciences) and optical density read at 450 nm after addition of a stopping solution. For Ulex staining NUNC immunoplates were coated with samples serially diluted from 1/400. After blocking in PBS containing 5% deffated milk for 1 h at 37°C alkaline phosphatase-conjugated UEA-I lectin (Sigma) at 10 µg/ml was incubated for 2 h at 37°C. Reactions were developed using *p*-nitrophenyl phosphate (Sigma) and optical densities read at 405 nm.

Statistical analysis

Data were analyzed by Student's t test using the Statview program.

Results

Haplotype diversity

In order to assess the patterns of haplotype variation at the *FUT2* locus, we sequenced the entire coding region of the gene in the 99 sampled individuals. We identified 11 haplotypes with 10 single nucleotide polymorphisms (SNPs) (Table 1): 3 missense, 6 synonymous and 1 nonsense SNP. Most efficient lineages are represented by haplotype SQ2, while most deficient lineages belong to haplotype SQ8.

Reconstruction of the evolutionary relationships between *FUT2* haplotypes uncovered two major branches that reflect the divergence between efficient alleles and deficient alleles carrying the 428G > A nonsense mutation (Fig. 1). Haplotype SQ5 was the only deficient lineage that was not associated with the 428*A variant, and results from a 302C > T weakly efficient missense mutation at codon 101

Table 1 *FUT2* haplotypes based on 10 polymorphic sites. Polymorphic sites are classified as missense (M), synonymous (S) and nonsense (N). The 739G > A polymorphism is outlined in light gray and the 302C > T and 428G > A mutations in dark gray. Haplotypes are outlined accordingly

		1	2	3	3	4	4	7	8	9	
	4	7	1	0	5	2	8	3	5	6	
	0	1	6	2	7	8	0	9	5	0	
	M	S	S	M	S	N	S	M	S	S	
Anc*	A	G	C	C	C	G	C	G	A	G	Frequencies
SQ1	*	A	*	*	*	*	*	*	*	A	10
SQ2	*	A	*	*	T	*	*	*	*	A	69
SQ3	*	A	*	*	T	*	*	A	*	*	10
SQ4	*	A	*	*	T	*	T	*	*	A	10
SQ5	*	A	*	T	*	*	*	*	*	A	1
SQ6	*	*	T	*	*	A	*	*	*	*	2
SQ7	*	*	T	*	*	A	*	A	*	A	1
SQ8	*	*	T	*	*	A	*	A	*	*	86
SQ9	*	*	T	*	*	A	*	A	C	*	2
SQ10	*	*	T	*	T	*	T	*	*	A	6
SQ11	G	A	*	*	*	*	*	*	*	A	1

*Ancestral haplotype from chimpanzee

(L101F) previously described in Asian populations [26–29]. Except for haplotype SQ3, all haplotypes carrying the 739G > A mutation were found to be associated with the 428*A inactivating variant. The decoupling of the association between 739*A and 428*A in haplotype SQ3 is likely to have resulted from an historical recombination event between haplotypes SQ2 and SQ8 in the region flanked by positions 428 and 739 (Table 1; Fig. 1). The recombinant

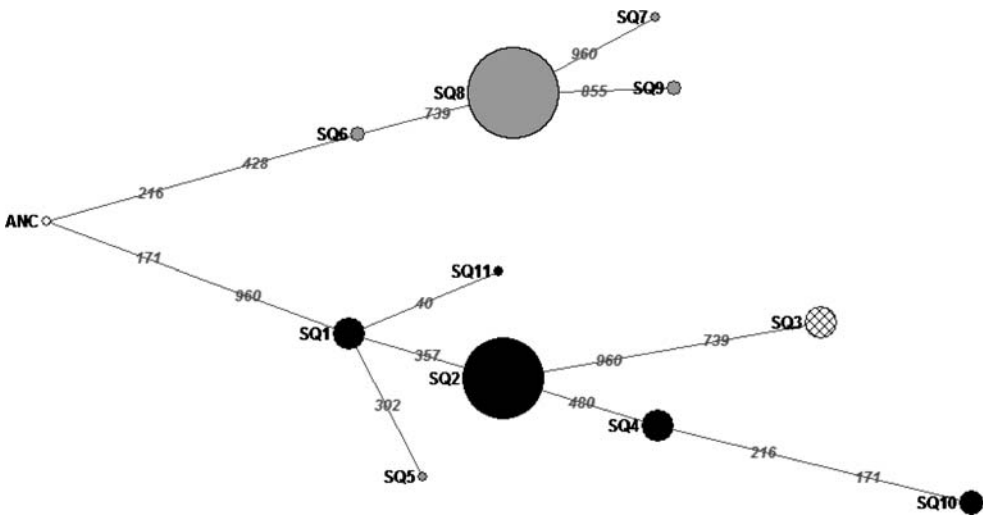

Fig. 1 Median-joining network showing relationships between *FUT2* haplotypes. Efficient haplotypes are represented in black and deficient in grey. The haplotype, carrying the 739G > A mutation without the 428G > A mutation is represented in dashed lines. Each circle is proportional to the haplotype frequency and nucleotide differences between haplotypes are indicated on network branches. Mutations are displayed in the branches connecting the different haplotypes. ANC stands for the ancestral haplotype corresponding to the chimpanzee sequence

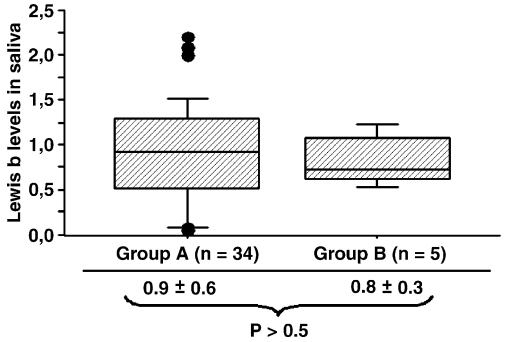

Fig. 2 Measurement of Lewis b levels in saliva. Box plot showing Lewis b levels in individuals with a combination of a deficient haplotype together with any efficient haplotype (group A) or SQ3 haplotype (group B)

SQ3 haplotype was used to evaluate the effect of the 739G > A mutation on FUT2 activity without the confounding effect of the 428*A variant (see below). Haplotype SQ10 is also likely to have resulted from a recombination event, involving haplotypes SQ4 and SQ8, between positions 216 and 357 (Table 1; Fig. 1).

FUT2 739*A allele encodes an efficient Se enzyme in saliva

To identify the *in vivo* relevance of the SQ3 haplotype carrying the 739G > A mutation without the inactivating 428*A variant, we analyzed the levels of Lewis b and Ulex expression in saliva in two groups of individuals. The first group included heterozygous individuals for one deficient and one efficient haplotype (group A). The second group included heterozygous individuals for one deficient haplotype and the SQ3 haplotype (group B) (Fig. 2). All individuals from group A were secretors, as evaluated by UEA-I reactivity (three times above background). Surprisingly, and in contrast to our expectations, expression of Lewis b (and UEA-I—data not shown) was not affected by the SQ3 haplotype. This observation suggests that, unlike our previous *in vitro* characterization [17], the 739G > A substitution is not an inactivating mutation *in vivo*.

FUT2 739*A allele is active in CHO-K1 cells

The finding that the *FUT2* 739*A variant is associated with the Se phenotype in saliva from individuals producing Lewis b, implies that the enzyme acts towards type 1 chain acceptor substrates. We therefore wanted to test the enzyme variant with other potential acceptor substrates known to function with the wild type Se enzyme such as type 2 and type 3 (O-linked) substrates. CHO-K1 cells only produce type 2 lactosamine chains on N-linked glycoproteins and core 1 (type 3) chains on O-linked glycoproteins, offering the possibility to evaluate activity with these two substrates. The α-1,2-fucosyltransferase activity of *FUT2* variants was analyzed *in vivo* by transfection into CHO-K1 cells, which are devoid of such endogenous enzyme activity. Following

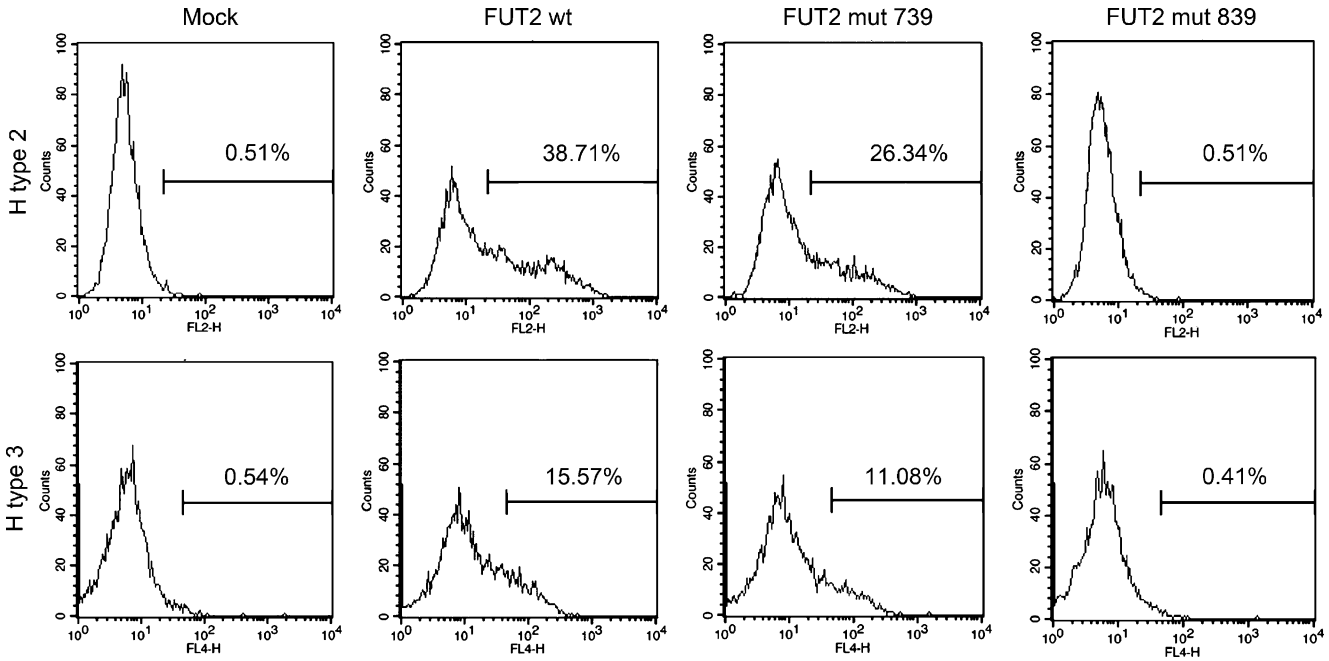

Fig. 3 Expression of H type 2 and H type 3 structures in CHO-K1 cells transfected with wild-type and mutant *FUT2* variants subcloned into pcDNA3.1 plasmids. The presence of α 1-2-linked fucose residues was analyzed by flow cytometry using the UEA-I lectin and the MBr1 monoclonal antibody for H type 2 and H type 3, respectively. Transfection efficiency was controlled by cotransfection with an EFGP expression vector and diagrams are representative of three independent experiments

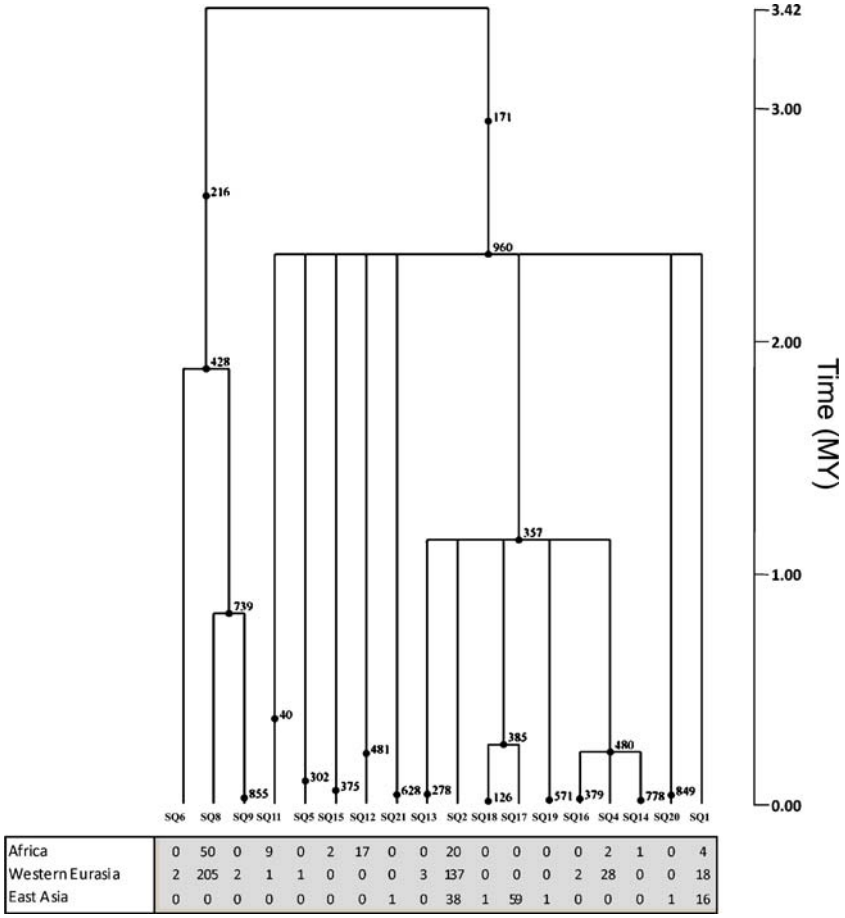

Fig. 4 *FUT2* scaled gene tree for a pooled data set including samples from Africa, western Eurasia and East Asia. Africa includes samples from South Africa (Xhosa) studied by Koda *et al*. [9] and African-Americans from the SeattleSNPs database (http://pga.gs.washington.edu/). The western Eurasia sample includes the Portuguese individuals studied in the present work, European–Americans from the SeattleSNPs database, and Europeans from South Africa and Iranians studied by Koda *et al*. [9]. East Asia includes Chinese and Japanese samples from Koda *et al*. [9]. Time is scaled in million years (MY). Nucleotide differences between haplotypes are indicated on the tree branches. Absolute frequencies of *FUT2* lineages are shown in the bottom of the tree

transient transfection, the presence of α-1,2-fucosylated structures at the cell surface was assessed using reagents against either H type 2 or H type 3. Relative fluorescence intensity values were normalised based on EGFP (enhanced green fluorescent protein) fluorescence of the *FUT2 wt* allele. The ratio of EGFP values for the two other *FUT2* variants transfected were 0.97 and 0.95, indicating that transfection efficiencies were reproducible.

The FUT2 *wt* variant exhibited slightly higher levels of activity compared to the *FUT2* 739G > A mutant both for H type 2 and H type 3 structures (Fig. 3). Although the difference was observed in three independent experiments, it did not reach statistical significance. By contrast, the *FUT2* 839 T→*C* mutant did not show any detectable activity since mean fluorescence values for either H type 2 or H type 3 were at background level, identical with those obtained after mock transfection.

Discussion

We have uncovered several polymorphic sites at the *FUT2* gene defining 11 haplotypes coding for active (SQ1, SQ2, SQ4, SQ10 and SQ11) or inactive (SQ5, SQ6, SQ7, SQ8 and SQ9) forms of the FUT2 enzyme. Except for SQ5, carrying the 302C > T missense substitution, all deficient haplotypes in our sample share the 428G > A nonsense mutation. Although the 428G > A mutation is commonly associated with the 739G > A mutation [9], we found a recombinant haplotype (SQ3) carrying the 739*A allele together with the active 428*G variant. Our previous results based on *in vitro* assays using a detergent solubilized enzyme and small acceptor substrates, suggested that the 739G > A mutation on haplotype SQ3 was associated with a nonsecretor phenotype [17]. In the present study, we observed that the levels of Lewis b in saliva were not affected by the presence of SQ3 in combination with deficient haplotypes, suggesting that the 739*A variant codes for an active Se allele *in vivo*. Efficiency of the SQ3 haplotype was further confirmed by transfection in CHO-K1 cells, where at best some reduction towards H type 2 and H type 3 precursors, but by no means inactivation, of the enzyme activity was demonstrated. Previous reports have described discrepancies between *in vitro* and *in vivo* assays of glycosyltransferases activities. For example, a mutation in the transmembrane domain of FUT3 that did not affect the *in vitro* enzyme catalytic properties was shown to be associated with a negative Lewis phenotype,

most likely because of an altered Golgi membrane anchoring [30]. Another example was published by our group showing that α-2,6-sialyltransferase ST6GalNAc-II is capable of biosynthesizing the sialyl-Tn antigen *in vitro* on different substrates but leads only to very limited biosynthesis of sialyl-Tn in transfected gastric carcinoma cell lines [31]. Our present results add another example of the lack of a strict relationship between *in vivo* and *in vitro* enzymatic assays. However, in the present case, the enzyme proved active *in vivo* but only very weakly *in vitro,* suggesting that the protein conformation may be unstable upon solubilisation. Regardless of the underlying mechanism, this observation stresses the need for *in vivo* data when describing new, putatively inactivating polymorphisms of glycosyltransferases.

Koda *et al*. [9, 32] have previously hypothesized that *FUT2* was subjected to balancing selection based on unusual patterns of DNA sequence variation and an estimated time to the most recent common ancestor (TMRCA) of *FUT2* genetic variation as high as 3 million years obtained with a phylogenetic approach. We have combined our haplotype information with an extended dataset including available sequences from Europe, Asia and Africa in order to generate a scaled tree of *FUT2* variation, using GENETREE (Fig. 4). Our analysis yielded a TMRCA estimate of 3.42±0.88 million years, in close agreement to that obtained by Koda *et al*. [32] with a different method, confirming that *FUT2* is among the human genes with highest coalescent times [33, 34]. The age of the 428G > A causing the nonsecretor phenotype in Africans and western Eurasians was calculated at 1.87± 0.85 million years. Since the 428*A is always associated with the 216*T variant (Fig. 4), the relative position of the 428G > A and 216C > T mutations along the branch containing deficient lineages is arbitrary and this age should be considered a minimum estimate based on the assumption that the 428G > A mutation is younger than 216 C > T. The age of the 739G > A was estimated at 816±324 thousand years. The 385A > T causing the nonsecretor phenotype in Asians had a younger age estimate of 256±155 thousand years. Taken together, these results are difficult to reconcile with the standard neutral model, in which panmixy, absence of selection and constant population size would have favoured a more rapid lineage turnover. Although high TMRCAs can result from ancient population structure [34], the fact that in both Asian and non-Asian populations the nonsecretor phenotype is caused by different mutations, reaching approximately the same frequency, further suggests that the persistence of efficient and deficient variants is a convergent feature of human populations that is favoured by selection. Consistent with the selective hypothesis, *FUT2* was found to be one of the six genes displaying evidence for non-neutral evolution in a recent survey of 168 genes related to immune function [10]. Moreover, it has been recently shown that *FUT2* allele frequencies are correlated with pathogen richness, suggesting that the secretor and nonsecretor status may have an important role in shaping susceptibility and resistance to different types of pathogens [11]. The detailed characterization of the function of *FUT2* alleles is therefore critical to understand the involvement of this gene in host-pathogen evolution.

Acknowledgements We gratefully acknowledge Nuno Ferrand for discussions and help in the use of the GENETREE software. We thank Nuno Mendes for expert technical assistance. This study was supported by Fundação para a Ciência e a Tecnologia (FCT) and Programa Operacional Ciência e Inovação 2010 do Quadro Comunitário de Apoio III (PTDC/SAU-MII64153/2006), by FCT project (PPCDT/BIA-BDE/56654/2004) and by the CIMATH project from the Région des Pays de la Loire. AS Carvalho thanks Fundação para a Ciência e a Tecnologia for post-doctoral fellowship (SFRH/BPD/36912/2007).

References

1. Marionneau, S., Cailleau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B., Ruvoën-Clouet, N., Clément, M., Le Pendu, J.: ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. Biochimie **83**, 565–573 (2001)
2. Lindesmith, L., Moe, C., Marionneau, S., Ruvoen, N., Jiang, X., Lindblad, L., Stewart, P., Le Pendu, J., Baric, R.: Human susceptibility and resistance to Norwalk virus infection. Nat. Med. **9**, 548–553 (2003)
3. Hutson, A.M., Airaud, F., Le Pendu, J., Estes, M.K., Atmar, R.L.: Norwalk virus infection associates with secretor status genotyped from sera. J. Med. Virol. **77**, 116–120 (2005)
4. Thorven, M., Grahn, A., Hedlund, K.O., Johansson, H., Wahlfrid, C., Larson, G., Svensson, L.: A homozygous nonsense mutation (428G > A) in the human FUT2 gene provides resistance to symptomatic norovirus (GGII) infections. J. Virol. **79**, 15351–15355 (2005)
5. Kindberg, E., Akerlind, B., Johnsen, C., Knudsen, J.D., Heltberg, O., Larson, G., Böttinger, B., Svensson, L.: Host genetic resistance to symptomatic norovirus (GGII.4) infections in Denmark. J. Clin. Microbiol. **45**, 2720–2722 (2007)
6. Stapleton, A., Hooton, T.M., Fennell, C., Roberts, P.L., Stamm, W.E.: Effect of secretor status on vaginal and rectal colonization with fimbriated Escherichia coli in women with and without recurrent urinary tract infection. J. Infect. Dis. **171**, 717–720 (1995)
7. Azevedo, M., Eriksson, S., Mendes, N., Serpa, J., Figueiredo, C., Resende, L.P., Ruvoën-Clouet, N., Haas, R., Borén, T., Le Pendu, J., David, L.: Infection by Helicobacter pylori expressing the BabA adhesin is influenced by the secretor phenotype. J. Pathol. **215**, 308–316 (2008)
8. Lindén, S., Mahdavi, J., Semino-Mora, C., Olsen, C., Carlstedt, I., Borén, T., Dubois, A.: Role of ABO secretor status in mucosal innate immunity and H. pylori infection. PLoS Pathog. **4**, e2 (2008)
9. Koda, Y., Tachida, H., Pang, H., Liu, Y., Soejima, M., Ghaderi, A. A., Takenaka, O., Kimura, V.: Contrasting patterns of polymorphisms at the ABO-secretor gene (FUT2) and plasma alpha(1, 3)

fucosyltransferase gene (FUT6) in human populations. Genetics **158**, 747–756 (2001)
10. Walsh, E.C., Sabeti, P., Hutcheson, H.B., Fry, B., Schaffner, S.F., de Bakker, P.I.W., Varilly, P., Palma, A.A., Roy, J., Cooper, R., Winkler, C., Zeng, Y., de The, G., Lander, E.S., O'Brien, S., Altshuler, D.: Searching for signals of evolutionary selection in 168 genes related to immune function. Hum. Genet. **119**, 92–102 (2006)
11. Fumagalli, M., Cagliani, R., Pozzoli, U., Riva, S., Comi, G., Menozzi, G., Bresolin, N., Sironi, M.: Widespread balancing selection and pathogen-driven selection at blood group antigen genes. Genome Res. **19**, 199–212 (2009)
12. Kelly, R.J., Rouquier, S., Giorgi, D., Lennon, G.G., Lowe, J.B.: Sequence and expression of a candidate for the human secretor blood group α(1, 2)fucosyltransferase gene (FUT2). J. Biol. Chem. **270**, 4640–4649 (1995)
13. Liu, Y., Koda, Y., Soejima, M., Pang, H., Schlaphoff, T., du Toit, E.D., Kimura, H.: Extensive polymorphism of the FUT2 gene in an African (Xhosa) population of South Africa. Hum. Genet. **103**, 204–210 (1998)
14. Koda, Y., Soejima, M., Kimura, H.: The polymorphisms of fucosyltransferases. Leg. Med. **3**, 2–14 (2001)
15. Soejima, M., Nakajima, T., Fujihara, J., Takeshita, H., Koda, Y.: Genetic variation of FUT2 in Ovambos, Turks, and Mongolians. Transfusion **48**, 1423–1431 (2008)
16. Blumenfeld, O.O., Patnaik, S.K.: Allelic genes of blood group antigens: a source of human mutations and cSNPs documented in the blood group antigen gene mutation database. Hum. Mutat. **23**, 8–16 (2004)
17. Serpa, J., Mendes, N., Reis, C.A., Santos Silva, L.F., Almeida, R., Le Pendu, J., David, L.: Two new FUT2 (fucosyltransferase 2 gene) missense polymorphisms, 739G→A and 839 T→C, are partly responsible for non-secretor status in a Caucasian population from Northern Portugal. Biochem. J. **383**, 469–474 (2004)
18. Nogueira, C., Figueiredo, C., Carneiro, F., Gomes, A.T., Barreira, R., Figueira, P., Salgado, C., Belo, L., Peixoto, A., Bravo, J.C., Bravo, L. E., Realpe, J.L., Plaisier, A.P., Quint, W.G., Ruiz, B., Correa, P., van Doorn, L.J.: Helicobacter pylori genotypes may determine gastric histopathology. Am. J. Pathol. **158**, 647–654 (2001)
19. Peleteiro, B., Lunet, N., Figueiredo, C., Carneiro, F., David, L., Barros, H.: Smoking, Helicobacter pylori virulence, and type of intestinal metaplasia in Portuguese males. Cancer Epidemiol. Biomarkers Prev. **16**, 322–326 (2007)
20. Stephens, M., Smith, N.J., Donnely, P.: A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. **68**, 978–989 (2001)
21. Stephens, M., Donnelly, P.: A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. **73**, 1162–1169 (2003)
22. Bandelt, H.J., Forster, P., Röhl, A.: Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. **16**, 37–48 (1999)
23. Griffiths, R.C., Tavaré, S.: Sampling theory for neutral alleles in a varying environment. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **344**, 403–410 (1994)
24. Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X., Rozas, R.: DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. **19**, 2496–2497 (2003)
25. Patterson, N., Richter, D.J., Gnerre, S., Lander, E.S., Reich, D.: Genetic evidence for complex speciation of humans and chimpanzees. Nature **441**, 1103–1108 (2006)
26. Chang, J.G., Yang, T.Y., Liu, T.C., Lin, T.P., Hu, C.J., Kao, M.C., Wang, N.M., Tsai, F.J., Peng, C.T., Tsai, C.H.: Molecular analysis of secretor type alpha(1, 2)-fucosyltransferase gene mutations in the Chinese and Thai populations. Transfusion **39**, 1013–1017 (1999)
27. Pang, H., Fujitani, N., Soejima, M., Koda, Y., Islam, M.N., Islam, A.K., Kimura, H.: Two distinct Alu-mediated deletions of the human ABO-Secretor (FUT2) locus in Samoan and Bangladeshi populations. Hum. Mutat. **16**, 274 (2000)
28. Soejima, M., Koda, Y.: Molecular mechanisms of Lewis antigen expression. Leg. Med. **7**, 266–269 (2005)
29. Soejima, M., Koda, Y.: Denaturing high-performance liquid chromatography-based genotyping and genetic variation of FUT2 in Sri Lanka. Transfusion **45**, 1934–1939 (2005)
30. Mollicone, R., Reguigne, I., Kelly, R.J., Fletcher, A., Watt, J., Chatfield, S., Aziz, A., Cameron, H.S., Weston, B.W., Lowe, J.B.: Molecular basis for Lewis alpha(1, 3/1, 4)-fucosyltransferase gene deficiency (FUT3) found in Lewis-negative Indonesian pedigrees. J. Biol. Chem. **269**, 20987–20994 (1994)
31. Marcos, N.T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., Almeida, R., Silva, F., Morais, V., Costa, J., Kihlberg, J., Clausen, H., Reis, C.A.: Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. Cancer Res. **64**, 7050–7057 (2004)
32. Koda, Y., Tachida, H., Soejima, M., Takenaka, O., Kimura, H.: Ancient origin of the null allele se(428) of the human ABO-secretor locus (FUT2). J. Mol. Evol. **50**, 243–248 (2000)
33. Tishkoff, S.A., Verrelli, B.C.: Patterns of human genetic diversity: implications for human evolutionary history and disease. Annu. Rev. Genomics Hum. Genet. **4**, 293–340 (2003)
34. Garrigan, D., Hammer, M.F.: Reconstructing human origins in the genomic era. Nat. Rev. Genet. **7**, 669–680 (2006)