

# Two $\alpha(1,2)$ fucosyltransferase genes on porcine Chromosome 6q11 are closely linked to the blood group inhibitor (S) and *Escherichia coli* F18 receptor (*ECF18R*) loci

E. Meijerink,<sup>1</sup> R. Fries,<sup>1,\*</sup> P. Vögeli,<sup>1</sup> J. Masabanda,<sup>1</sup> G. Wigger,<sup>1</sup> C. Stricker,<sup>1</sup> S. Neuenschwander,<sup>1</sup> H.U. Bertschinger,<sup>2</sup> G. Stranzinger<sup>1</sup>

<sup>1</sup>Institute of Animal Science, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zurich, Switzerland <sup>2</sup>Institute of Veterinary Bacteriology, University of Zurich, CH 8057 Zurich, Switzerland

Received: 17 February 1997 / Accepted: 30 May 1997

Abstract. The Escherichia coli F18 receptor locus (ECF18R) has been genetically mapped to the halothane linkage group on porcine Chromosome (Chr) 6. In an attempt to obtain candidate genes for this locus, we isolated 5 cosmids containing the  $\alpha(1,2)$  fucosyltransferase genes FUT1, FUT2, and the pseudogene FUT2P from a porcine genomic library. Mapping by fluorescence in situ hybridization placed all these clones in band q11 of porcine Chr 6 (SSC6q11). Sequence analysis of the cosmids resulted in the characterization of an open reading frame (ORF), 1098 bp in length, that is 82.3% identical to the human FUT1 sequence; a second ORF, 1023 bp in length, 85% identical to the human FUT2 sequence; and a third FUT-like sequence thought to be a pseudogene. The FUT1 and FUT2 loci therefore seem to be the porcine equivalents of the human blood group H and Secretor loci. Direct sequencing of the two ORFs in swine being either susceptible or resistant to adhesion and colonization by F18 fimbriated Escherichia coli (ECF18) revealed two polymorphisms at bp 307 (M307) and bp 857 (M857) of the FUT1 ORF. Analysis of these mutations in 34 Swiss Landrace families with 221 progeny showed close linkage with the locus controlling resistance and susceptibility to E. coli F18 adhesion and colonization in the small intestine (ECF18R), and with the locus of the blood group inhibitor S. A high linkage disequilibrium of M307-ECF18R in Large White pigs makes the M307 mutation a good marker for marker-assisted selection of E. coli F18 adhesion-resistant animals in this breed. Whether the FUT1 or possibly the FUT2 gene products are involved in the synthesis of carbohydrate structures responsible for bacterial adhesion remains to be determined.

# Introduction

Edema disease and post-weaning diarrhea in swine is associated with the colonization of the small intestine with toxigenic *Escherichia coli* strains of a limited number of serotypes. Colonization depends on adherence of the bacteria, which is mediated by bacterial fimbriae F18 or F4 (Sellwood et al. 1974; Bertschinger et al. 1990; Rippinger et al. 1995). Several candidate structures have been proposed as receptors for fimbriae F4 (Grange and Mouricout 1996), whereas receptors for F18 have not yet been characterized. Susceptibility to adhesion, meaning expression of receptors mediating the binding to the bacterial fimbriae F18 formerly designated fimbriae F107, has been shown to be genetically controlled by the host and is inherited as a dominant trait (Bertschinger et al. 1993) with B being the susceptibility allele and b the resistance allele. The genetic locus for this E. coli F18 receptor (ECF18R) has been mapped to porcine Chr 6 (SSC6), based on its close linkage to the S locus and other loci of the halothane (HAL) linkage group (Vögeli et al. 1996). The epistatic S locus suppresses the phenotypic expression of the A-0 blood group system when being S<sup>s</sup>S<sup>s</sup> (Vögeli et al. 1983) and allows expression when being  $S^{S}S^{-}$ . Likewise, the expression of the AB0 and Lewis (Le) blood group systems in humans is influenced by the blood group H and by the Secretor (Se) locus. Inactivating mutations in these genes prevent the expression of the precursor H determinant, and thus of A and B (Kelly et al. 1994) and Le<sup>b</sup> determinants. The expression of Lewis antigen in the small intestine of humans was shown to be under control of the Lewis gene ( $\alpha$ 1,3/4 fucosyltransferase) and the Secretor gene (Björk et al. 1987; Henry et al. 1994) with the type of Lewis antigen (Le<sup>a</sup> or Le<sup>b</sup>) determined by the Secretor locus (se or Se, respectively).

The blood group H and Se loci have been mapped genetically (Oriol et al. 1984; Reguigne-Arnould et al. 1995) and physically (Rouquier et al. 1995) to human Chr 19q13.3. This region is evolutionarily conserved, containing homologous genes from the HAL linkage group in pigs. Conservation was shown by ZooFISH analysis (Rettenberger et al. 1995) and earlier by the genetic and physical mapping of genes such as glucosephosphate isomerase (GPI; Davies et al. 1988), phosphogluconate dehydrogenase (PGD) and apolipoprotein E (APOE) in both chromosomal segments (Gedde-Dahl et al. 1984; Vögeli et al. 1988; Clamp et al. 1993; Rouquier et al. 1994). Larsen and associates (1990) characterized the blood group H encoding gene (FUT1). The structure and expression of the Se gene (FUT2) was reported by Kelly et al. (1995) and Rouquier et al. (1995). FUT1 determines H antigen expression in the erythroid cell lineage, whereas FUT2 regulates expression of the H antigen in secretory epithelia and saliva. Conservation of the FUT1 gene has been shown in lower mammals such as rat, rabbit, and pig, and mRNA expression has been found in rabbit brain tissue (Hitoshi et al. 1995) and rat colon (Piau et al. 1994). In all these species, two types of  $\alpha(1,2)$  fucosyltransferase genes have been reported that are structurally very similar to the human FUT1 and FUT2 genes, but in particular the FUT1 homologous genes show a species-specific expression pattern. In humans the FUT1 gene is responsible for synthesis of H antigens in the precursors of erythrocytes (Watkins 1980). However, in pigs erythrocytes passively adsorb H-like antigens from the serum (Oriol 1987), as is the case for the human Lewis antigens. In pigs, all H-like antigens detected by Oriol and colleagues (1993) were related to exocrine secretory tissues, and expression of the FUT2

Correspondence to: P. Vögeli

<sup>\*</sup> Present address: Chair of Animal Breeding, Technical University Munich, D-85350 Freising-Weihenstephan, Germany

The presented sequences are submitted to the NCBI- and GenEMBL databases under the accession numbers U70881, U70882, and U70883

(Secretor) gene is seen in secretory tissues of other animal species (Piau et al. 1994; Rouquier et al. 1995). Therefore, expression of the porcine A-0 blood group determinants, which cross-react with anti-human blood group H and A antibodies (King and Kelly, 1991), might be influenced by the FUT2 gene.

Carbohydrate structures of blood group antigens have been shown to mediate the adhesion of some pathogenic microorganisms to host tissues; for example, Helicobacter pylori adheres to Lewis<sup>b</sup> blood group antigens (Borén et al. 1993), and E. coli strains causing urinary tract infections adhere to P blood group substance (Källenius et al. 1980; Svenson et al. 1983). Genes encoding glycosyltransferases that are responsible for the formation of the blood group-specific carbohydrate structures therefore represent candidate genes for the control of bacterial colonization by the host. It was on this basis that we attempted to clone porcine fucosyltransferase genes. These genes can be considered as positional candidates, since their localization was expected to be in the same chromosomal region as the locus responsible for adhesion/ non-adhesion of F18-positive E. coli in the small intestine. The detection of single nucleotide polymorphisms in one of the cloned fucosyltransferase (FUT1) genes allowed us to study linkage between this gene, the S-system, the ECF18R locus, and other loci of the HAL-linkage group.

The chromosomal localization of the fucosyltransferase genes in the *ECF18R* region was confirmed by fluorescence in situ hybridization and by linkage analysis.

#### Materials and methods

Primers (Table 1) derived from the human FUT1 gene were used for the amplification of its porcine counterpart from genomic DNA. From the resulting porcine sequences, specific primers were designed that were used in further amplification and sequencing reactions.

A porcine genomic library, constructed in SuperCos 1 (Stratagene, La Jolla, Calif., USA), was screened with an  $\alpha^{32}P$  dATP-labeled (Prime It II, Stratagene) FUT1 probe obtained from porcine genomic DNA with primers P7 and P10. After hybridization of replica filters at 42°C for 15 h (50% formamide, 6 × SSC, 5 × Denhardt's, 0.5% SDS, 0.1 mg/ml salmon sperm) and washing twice at 65°C for 30 min (1 × SSC, 0.1% SDS), positive colonies could be identified after exposure (15 h, -80°C) to X-ray film.

Cosmid clones ETHs1, ETHs2, ETHs3, ETHs4, and ETHs6 were subjected to fluorescence in situ hybridization (FISH) on porcine metaphases. Metaphase chromosomes were Q-banded and photographed before hybridization. The probes were labeled by random priming with biotin-16-dUTP. Signal detection and amplification were performed with avidin-FITC and biotinylated anti-avidin. The chromosomes were counterstained with 4,6diamidino-2-phenylindole, and the relative positions of the cosmids were determined as described by Solinas Toldo and coworkers (1993).

*Ksp*I-, *Eco*RI- and *Ksp*I/*Eco*RI digests of all cosmids were separated on a 0.8% agarose gel and transferred to a Hybond N nylon membrane. This blot was hybridized with  $\alpha^{32}P$  dATP-labeled porcine *FUT1* PCR products (primers P6–P11 and P7–P10). Based on the autoradiographic signals, ETHs1, -s2, and -s3 were subjected to further subcloning into pBluescript SK- (Stratagene), and FUT sequences were determined from subclones. The sequences of two FUT-like ORFs (*FUT1* and *FUT2*) (Fig. 3) obtained from cosmids ETHs2 and -s3 were compared in ECF18R-positive (BB/Bb) and -negative (bb) animals by direct sequencing of PCR products.

Using the Perkin Elmer Ready Reaction Dye Terminator kit (Perkin Elmer Cetus, Norwalk, CT, USA) and 10 pmol of primer, cycle sequencing was performed with a thermal program consisting of an initial denaturation of 5 min at 95°C, followed by 25 cycles of 30 s at 95°C, 15 s at 50°C, and 4 min at 60°C. Primers used for amplification and sequencing of the porcine  $\alpha(1,2)$  fucosyltransferase genes are listed in Table 1. Additional primers were designed on the sequences shown in Fig. 3, taking the possibility of cross-annealing of primers owing to the high similarity of *FUT1*, *FUT2*, and the *FUT2* pseudogene into account. Samples were analyzed on a 373A ABI sequencer (Applied Biosystems Inc.), and sequence analysis was performed with the GCG package (Devereux et al. 1984).

Single nucleotide polymorphisms were analyzed in 221 Swiss Landrace pigs produced from 4 boars and 16 sows, and in 29 Large White pigs produced from 9 unrelated matings. In order to produce a large number of informative offspring for the examination of linkage between porcine

Primer Name	Primer Sequence	Position
FUT1 P6 (R)	5'-CTTCAGCCAGGGCTCCTTTAAG-3'	+489
FUT1 P7 (F)	5'-TTACCTCCAGCAGGCTATGGAC-3'	+720
FUT1 P10 (R)	5'-TCCAGAGTGGAGACAAGTCTGC-3'	+1082
FUT1 P11 (F)	5'-CTGCCTGAACGTCTATCAAGATC-3'	+69
FUT1 P16 (F)	5'-AGAGTTTCCTCATGCCCACAGG-3'	-90
FUT1 P18 (R)	5'-CTGCTACAGGACCACCAGCATC-3'	+1203
FUT1 PBEST (R)	5'-ACCAGCAGCGCAAAGTCCCTGAC GGGCACGGCCTC-3'	+893
FUT2 P16 (R) FUT2 P17 (F)	5'-CTCCCTGTGCCTTGGAAGTGAT-3' 5'-AACTGCACTGCCAGCTTCATGC-3'	+1094 -83

genes encoding ECF18 receptors and selected polymorphic loci, only informative Swiss Landrace matings of the type  $B/b \times b/b$  were produced for this study. In a study of Bertschinger et al. (1993), the above-mentioned Swiss Landrace pigs were also tested for *ECF18* susceptibility in a colonization test. For this, pigs were inoculated shortly after weaning with bacteria of *E. coli* strain 124/76 of serotype O139:K12(B):H1:F(18). Fecal shedding of the bacteria was monitored daily. The extent of colonization was calculated as the mean of the two highest fecal scores. Pigs with a mean fecal score of 3.5, corresponding to 6.7 log colony forming units (CFU)/g or more, were considered susceptible to colonization. This limit was based on a lack of mortality below this value, and on scores obtained from completely resistant litters.

The results of the single nucleotide polymorphisms were compared with typing data for ECF18R, which were identified in an in vitro adhesion assay described by Vögeli and associates (1996), and with typing data for the *GP1-*, *PGD-*,  $\alpha$ -1-B-glycoprotein- (*A1BG*), ryanodine receptor (*RYR1*), *EAH-* and *S-* loci as published by Vögeli and colleagues (1996). Pairwise linkage analysis and calculation of recombination fractions was performed with the CRI-MAP version 2.4 program (Green et al. 1990). Multipoint linkage analysis was performed by sequential insertion of above loci into the map. Haplotype frequencies were calculated from the parental animals in the Swiss Landrace families and from the 8 parental Large White animals, which were haplotyped for *ECF18R* from progeny information. Tetrachoric correlations of ECF18R and mutations in *FUT1* (FUT1/M307) were calculated on all Swiss Landrace and Large White progeny.

# Results

Cosmids ETHs1, -s2, -s3, -s4 and -s6 were identified after screening of the cosmid library with a FUT1 probe obtained from porcine genomic DNA with primers P7 and P10 and were mapped by FISH to Chr 6 in band q11 (Fig. 1).

When hybridizing with radiolabeled porcine *FUT1* P6–P11 and P7–P10 PCR fragments, Southern blot analysis of *KspI*, *Eco*RI, and *KspI/Eco*RI cosmid digests revealed identical autoradiography signals for ETHs2, -s4 and -s6, whereas different signals were obtained from cosmids ETHs1 and -s3 (Fig. 2A,B). From ETHs2, we isolated *KspI* subclones 940 bp and 6.2 kb in length, corresponding to the estimated length of hybridizing *KspI* fragments on the Southern blot. The sequence results of both subclones were combined to yield a 1501-bp sequence, which was in agreement with results of direct sequencing of genomic PCR products. The 1501-bp sequence contains an ORF of 1098 bp corresponding to the human *FUT1* ORF, with 82.3% nucleotide and 80.8% amino acid identity.

ETHs1 shows one fragment (2.7 kb) (Fig. 2B) hybridizing to FUT1 sequences, whereas ETHs3 shows two (2.7 kb and 8.2 kb). Subcloning and partial sequencing of the 2.7-kb *Eco*R1 fragments of ETHs1 and -s3 confirmed that these two fragments are identical. The sequence is highly similar to the human FUT2, but shows several changes in the NH<sub>2</sub>- and -COOH terminal regions. These lead to frame shifts that are not compatible with a conserved ORF; therefore, we assume the sequence obtained from the 2.7-kb fragment to represent a pseudogene (*FUT2P*). After subcloning of



**Fig. 1.** (**A**) Q-banding pattern of porcine metaphase chromosomes and (**B**) fluorescence in situ hybridization (FISH) with biotin-labeled cosmid DNA, assigned the highly specific signals to the chromosomal region 6q11.



**Fig. 2.** Southern blot analysis of cosmids ETHs1 (1–3), ETHs2 (4–6), and ETHs3 (7–9) after digestion with enzymes *KspI* (1, 4, 7), *Eco*RI (2, 5, 8), and *KspI/Eco*RI (3, 6, 9) and separation on a 0.8% agarose gel. In Fig 2A is seen that hybridization with an  $\alpha^{32}P$  dATP labeled 5' *FUT1* fragment (primers P6–P11) results in the same hybridizing 940-bp band in both the *KspI* digest (lane 4) and the *KspI/Eco*RI digest (lane 6). In Fig. 2B, however, hybridization with a 3' *FUT1* fragment (primers P7–P10) shows a 6.2-kb *KspI* band in lane 4 and a 1.1-kb *KspI/Eco*RI band in lane 6. Both the 5' and 3' *FUT1* fragments hybridize to the same 4.6 kb *Eco*RI fragment in lanes 5 (Fig. 2A, 2B). This indicates a *KspI* site in the *FUT1* gene contained in cosmid ETHs2. In Fig. 2B, cross-hybridization of the 3' *FUT1* fragment detects 2.7-kb (lanes 2, 3, 8, and 9) and 8.2-kb (lanes 8 and 9) bands, containing the *FUT2* pseudogene and the *FUT2* gene sequences, respectively.

ETHs3 *Bam*HI digests, the hybridizing sequences contained in the 8.2 kb *Eco*RI fragment could be identified. The sequence of the obtained subclones represents a 1023-bp ORF and is 85% identical at the nucleotide, and 83% identical at the amino acid level to the human FUT2 sequence. Many differences in the NH<sub>2</sub>- and -COOH terminal regions were observed between the porcine *FUT2* sequence and the *FUT2P* sequence derived from the 2.7-kb fragment. The predicted amino acid sequence corresponds to the partially determined amino acid sequence of the porcine *Secretor* enzyme (Thurin and Blaszczyk-Thurin 1995) (Fig. 3). The obtained porcine *FUT1*, *FUT2*, and *FUTP* sequences are submitted to GenBank and have accession numbers U70883, U70881, and U70882, respectively.

Direct sequencing of porcine FUT1 and FUT2 sequences and

their flanking regions in animals of different ECF18R genotype (*Bb*, *bb*) resulted in the identification of two G  $\rightarrow$  A transitions at positions 307 and 857 (termed *M307* and *M857*, respectively) of the *FUT1* ORF. The *M307* transition eliminates a restriction site for *CfoI*. Amplification with primers P6 and P11 (3 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 56°C, and 30 s at 72°C, followed by a 7-min final extension at 72°C) followed by *CfoI* digestion and separation on a 3% agarose gel results in a restriction fragment length polymorphism (RFLP; Fig. 4). Homozygous *M307*<sup>A/A</sup> animals show 93- and 328-bp fragments where homozygous *M307*<sup>G/G</sup> animals show 93-, 241- and 87-bp fragments. Heterozygous animals show all four fragments.

The *M857* transition eliminates an *Aci*I site. Primer PBEST was designed to mismatch two additional *Aci*I sites at positions 866 and 872. PCR with primers P7 and PBEST (3 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 56°C, and 30 s at 72°C, followed by a 7-min final extension at 72°C) followed by *Aci*I digestion enables PCR-RFLP analysis on a 3% agarose gel. Homozygous  $M857^{A/A}$  animals show a 174-bp fragment, while amplification products of  $M857^{G/G}$  animals show 136- and 38-bp fragments (Fig. 4).

In Swiss Landrace families, recombination events between M307 and the loci of the HAL linkage group (*S*, *ECF18R*, *RYR1*, *GPI*, *PGD*) revealed recombination fractions  $\theta < 0.04$  (Table 2). The lodscores Z for the overall recombination fractions were between 24.5 and 50.6, showing strong evidence for linkage between these loci. These data allow genetic mapping of the *FUT1* gene to the HAL linkage group in close proximity of *S* and *ECF18R*. Owing to low recombination rates, however, no locus order could be determined as being significantly better than others.

The haplotype frequencies are presented in Table 3. In Swiss Landrace (SL) and Large White (LW) parental pigs,  $ECF18R^b$  (the edema and post-weaning diarrhea resistance allele) is 100% associated with  $M307^A$ , and  $ECF18R^B$  (the edema and post-weaning diarrhea susceptibility allele) 100% with  $M307^G$ . It is worth while mentioning that in SL pigs the  $RYR1^T$  (the malignant hyperthermia susceptibility allele) accounted for 93% (28/30) of all  $ECF18R^b$  haplotypes. In SL pigs 88% (30/34) of  $S^s$  accounted for all  $ECF18R^b$  and  $M307^A$  haplotypes, respectively. The corresponding values for both the  $S^s$ - $ECF18R^b$  and  $S^s$ - $M307^A$  haplotypes were 82% (9/11) in Large White pigs. In our experimental SL families the occurrence of the  $M857^A$  allele at FUT1 was low and even absent in LW pigs. Therefore, a significant gametic association could not be observed between the alleles of M857 and the alleles of the flanking genes.

Table 4 shows that the distribution of *FUT1* genotypes at nucleotide 307 among *ECF18R* types was significantly different from the expected ratio. Of the 119 edema and post-weaning diarrhea-resistant *ECF18R<sup>b/b</sup>* animals, 118 were determined as  $M307^{A/A}$  in the DNA-based test. One resistant animal had the genotype  $M307^{A/G}$ . Of the 131 susceptible pigs, 130 were  $M307^{A/G}$  or  $M307^{G/G}$ . One animal, susceptible to *E. coli* adhesion, was shown to be homozygous  $M307^{A/A}$  by the DNA test.

# Discussion

The genetic and physical mapping of the identified *FUT1* ORF is in agreement with our expectations showing that the porcine *FUT1* homolog is located on SSC6q11. Similarities in length, nucleotideand amino acid sequence with the human FUT1 suggest that the established ORF in our study represents the porcine *FUT1*. Furthermore, except for nucleotides 229 and 714, our sequence is identical to the recently published sequence of Cohney and coworkers (1996), who showed that the *FUT1* gene product exhibits  $\alpha(1,2)$ fucosyltransferase activity. We found two *FUT2*-like sequences, one representing the porcine *FUT2* and the other one being a pseudogene (*sFUT2P*). As in humans, the  $\alpha(1,2)$ fucosyl3A

sfuti sfuti hfuti	M W V P S R R H L C L T F L L V C V L A A I F F L N V Q D L F Y S G L D L L A ANGIGGETGEGEGEGEGEGEGEGEGEGEGEGETGETGETGETAGETA	120
sfuti sfuti hfuti	L C P D H N V V S S P V A I F C L A G T P V H P N A S D S C P K H P A S L S G T CTGTGTCCAAACCATAACGTGGTATCATCTCCCCGGGGGCAACGCCCGGTACCACCCCCAACGCCTCCGATTCCCGGCCCCAACGCCTCCCAATCCCCCAACGCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCAACGCCCAACGCCCAACGCCCAACGCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCAACGCCCCCC	240
sfuti sfuti hfuti	W T I Y P D G R F G N Q H G Q Y A T L L A L A Q L N G R Q A P I Q P A H H A V L Togactattraccogategoccogtteggalccaategalacaategalacaategalacategalacaacegoccacegoccacegoccaetegalacegoccategoccategalacegoccategococategococategococategococategococategococategococategococategocc	360
SFUT1 SFUT1 hFUT1	A P V F R I T L P V L A P E V D R H A P W R E L E L H D W H S E D Y A H L K E F GCCCCCGATTACGCATCACGCATGCCTGCCGGCGAGGTAGACAGGAGGCACGCTGGGGGGAGGTGGAGCTTCACGACGGAGGATTATGCCCACGAGGAC G-A	480
sfut1 sfut1 hfut1	WLKLTCFPCSWTFFHHLREQIRSET LHDHLRQEAQGC AGAGGGAGTGAAGCCCAGGGAAGCACCACCTCGGCAAAAAGCCCAGGGGAGTACTAAG TGGCTGAAGCTCACCGGCTTCCCCCTGCCGGAACCCCCTGCGGAACCACCTTCGGCAAAAGCCCAGGGGAGTACTAAG -TC	600
sFUT1 sFUT1 hFUT1	C F R L P R T G D R P S T F V G V H V R R G D Y L R V M P K R W K G V V G D G A CAGTICCTCTACCCCCGACAGCGCCCCAGCACCTCCTGGGGGCCCAGCGCCCAGGGGCGACCATCTGCCGTGATGCCCCAAGCGCCTGGAGGGGGTGGCTGATGCCCC CCGGG	720
PFUT1 PFUT1 HFUT1	YLQQAMDWFRARYEAPVFVVTSNCMEWCRKNIDTSRGDVI TACTCCAGCAGGCTATGGACGGGCCCGGATAGGACGGCCCGGATGGACGGCCGGAAGAACGACACCTCCCGGGGGGACGTATG 	840
sFUT1 sFUT1 hFUT1	F A G D G R E A A P A R D F A L L V Q C N H T I M T I G T F G F W A A Y L A G. G TTTGCTGGCGAGGCGGGGGCGCGCCGCCGCCGCGGGCTTTGGCGTGCGGTGCAGTGGAGCACCACCATCATTGGCATTGGCAGCCTTCGGCGCCCCTACCTGGCTGG	960
sfut1 sfut1 hfut1	D T I Y L A N F T L P T S S F L K I F K P E A A F L P E W V G I N A D L S P L C GATACCATCTACTAGCTAACTTCACCCTGCCGACTTCCTGALGATCTTTAAACCCGGGGGGGGCGGCCGGGGGGGGGG	1080
sfuti sfuti hfuti	м L A G P * Атоттосстосстора -сааа	1098

# 3**B**

hFUT2 sFUT2 sFUT2 sFUTP	AT	GAG	ACC	TAC	ccg	 ccc.	AGCO	CAT	ccc	rcc	TAA	GCG			. сто Г	CCC	CTG C	AC. T-T	A GT-	GCC	M ATG -CC	- L CTC	V S AGC	V M ATG T-C	Q CAGO	M A GCA	P S TCT A-C	F TTC TG-	CCT	S F TTC	- F FTCC	- P 2007	M T CGG	A G GTC TCT	H P CCC1	- F (TC)	- I NTCC	- L TC1	F FTTG G-CA	V TC	57
hFUT2 sFUT2 sFUT2 sFUTP	F TTC	T ACG -T-	V A GCT -TG	5 TCC	T ACC.	I ATA	F TTTC	H	V L CTTO FGC	Q CAGO	- Q CAGI	R AGGZ	L M TG	A V STG	K AAGI	- I ATT G-G	- CAA -CC	A P CCC AA-	M T ACG C-C	W TGG	E GAG -CC	L TTA -CC	P Q CAG	V M ATG GGC	Q V GTG	I T ACG TTG	P Q CAGO GTCO	v STG	<b>L</b>	<b>A</b>	s 	T ACCJ	S T CAG	K E GAGA	A 5 AGCO	L P CCC	G 5 ICG2 A-C1	P S GCC	s P CCCC	Q AG C-	168
hFUT2 sFUT2 sFUT2 sFUTP	- CTG	R K AAG	- G G G G G G C C	MATG	w TGG.	T ACG.	I	N NAATO	A BCC	- I ATCO	G GGC(	R	L	g GGG	- HAAC	Q CAG	N ATG	- 6 6 6 6 6 6 6 6 6	GAG	TAC	GCC	- T ACC	L CTG	TAC	A GCG	L CTG	A GCC/	K R AGG.	M	N AAC	- G 3660	R	P	A BCC1	F F	- I ATC	P	A P CCC	Q E BAGA	- M TG	288
hFUT2 sFUT2 sFUT2 sFUTP	H CAC	S AGC	т АСО	L CTG	A	- P ccc.	I	F	R	I	T	L	P CG	- V STC	- L CTG	H	S A GCC	A S AGC	T ACG	A GCC	S R CGC	R AGG	I ATC	- ccc	W TGG	Q CAG	N AAC	Y TAC	H CAC	L CTG	- M	D BACT	W GGJ	- M	- B BAGG	E BAG	E R CGG1	Y NACO	R	н АС	408
hFUT2 sFUT2 sFUT2 sFUTP	I	- P CCG	- 660	E GAG	Y TAC	V GTG	R	F L CTC	T	G G GCC	Y TAC	P	c GC	s rcc	W TGGJ	TACC	F	Y TAC	H	H	L	R CGC	Q T ACC	E GAG	I	L CTC	Q R CGGG	EGAG	F TTC	TACC	L	H H	D N ACC	H	- V STG(	R	EGAGO	E	A SCCC	Q AG	528
hFUT2 sFUT2 sFUT2 sFUTP	K D GAT	F	L CTG	R	GGT	L	Q R CGGC	V TG	N	G G GCG	S AGC	R	P CG	G S AGT	T	F Y TAC	V GTG	- 	V GTG	H	V	R	R	GGG	D GAC	Y TAC	V GTG	H CAC	V GTG	M	P P CCC/	K N ACC	- V TG1	w rGG2	K AAGO	G	V STGC	v TGG	A BCCG	D AC	648
hFUT2 sFUT2 sFUT2 sFUTP	R	R	Y TAC	L CTG	Q E GAG	CAG	A	L	D GAC	W TGG	F	R	A	R	Y	S R CGC	s TCC	L P CCC	I V GTC	F	v v	v GTC	T S TCC	S AGC	N	GGC	MATG	A GCC	W	C TGT	R	- E BAAJ	N ACJ	I TCJ	D N ATC	T A SCC	s	H R CGCC	G G GGCG	- D AT	768
hFUT2 sFUT2 sFUT2 sFUTP	GTO	V GTG	F	A GCC	- G GGC.	D N AAT	G G G G G G G C I	I	E	G 30C	s rcco	P	À CC.	К АЛА	D GAC	F	A GCG	L CTG	L CTC	T	Q CAG	C TGT	N AAC	H CAC	T	U V GTC	M ATG	T ACC.	I	GGC	T	F	G G GGGJ	I	W	A	A	Y	L	T A CT	888
hFUT2 sFUT2 sFUT2 sFUTP	GGT	GGA	D E GAG	T ACC	I ATC	Y TAC	L	A BCC	N AAT	Y FAC	- T ACGO	- L TCC	P CG	- D GAC'	- S TCT( C	- P CCC	F	- L CTC	- к ала -тс	I L CTC G-G	- F TTT	K AAG	-	- B GAG	A GCA	A GCC	TTC	L CTG	- P CCC	В GAG <sup>4</sup> G	W TGGJ	T I ATTO G-G-	- 6 1667	I TCC	A E SAGO	- - 	D GACC	L	- 5 TCCC -GG-	P CA AG	1008
hFUT2 sFUT2 sFUT2 sFUTP	L CTC GC-	L CTT	K K AAG G-C	- H CAC	TGA		 TAG																																		1023

**Fig. 3.** (A) Alignment of the porcine (sFUT1) and human (hFUT1) FUT1 coding regions. Amino acid translation is given for the porcine FUT1 (upper row). At nucleotides 307 (M307) and 857 (M857) of the porcine FUT1 G $\rightarrow$ A transitions were observed (bold printed). (B) Alignment of the porcine *FUT2* gene (sFUT2) and *FUT2* pseudogene (sFUTP) by the Pileup function of the GCG sequence analysis program. Basepair count is

given for the porcine *FUT2*, starting from the ATG startcodon. In the upper two rows, amino acid translations of human- (hFUT2) and porcine-(sFUT2) *FUT2* coding regions are given. Similarities in nucleotide- and amino acid sequences are indicated by dashes. Bold printed amino acid sequences are identical to the partial amino acid sequence of porcine Secretor type  $\alpha(1,2)$ fucosyltransferase (Thurin and Blaszczyk-Thurin 1995).

transferase genes in pigs seem to form a cluster of functional genes and pseudogenes. However, in human the pseudogene seems to be derived from the FUT1 gene, whereas in pigs it is most similar to the *FUT2* gene. Linkage was demonstrated between inheritance of the alleles of the *ECF18R* gene and polymorphisms in the *FUT1* (*M307/M857*), *S*, *RYR1* (the halothane gene), *GPI*, and *PGD* genes. The mutation of *M307* at *FUT1* is linked to *ECF18R* with a recombination fraction of 0.01 in 221 informative SL meiosis. In both recombinants where lack of association between M307 and ECF18R was found (Table 4), adhesion status, although determined by the microscopic adhesion test, may have been falsely diagnosed. Indeed, the fecal score in the colonization test was 5.5 for the  $M307^{A/G}$  animal classified as resistant ( $ECF18R^{B/b}$ ) and 2.3 for the  $M307^{A/A}$  animal classified as susceptible ( $ECF18R^{B/b}$ ). Animals with a fecal score of 3.5 or greater are considered to be susceptible (see Materials and methods). Therefore, if the resis-



**Fig. 4.** Detection of **(A)** the M307 G to A and **(B)** the M857 G to A mutation in the porcine *FUT1* gene. Digestion of amplified *FUT1* fragments with *CfoI* **(A)** and *AciI* **(B)** results in a restriction fragment length polymorphism. First lane: 100-bp marker. Fragment lengths are indicated in basepairs. **(A)** The  $M307^{A/A}$  genotype (lane 2) generates 328 and 93 bp restriction fragments, while the  $M307^{G/G}$  genotype (lane 4) generates 93, 241, and 87 bp fragments and heterozygous  $M307^{A/A}$  genotype (lane 3) show all four fragments. **(B)** Digestion of the  $M857^{A/A}$  genotype (lane 2) generates 174-bp fragments while it generates 136- and 38-bp fragments in the  $M857^{G/G}$  genotype (lane 4), and in  $M857^{A/G}$  genotypes (lane 3) all three fragments are generated.

**Table 2.** Overall recombination fractions ( $\theta$ ), lodscores (Z), and number, of informative animals (N) for *M307* and loci of the *HAL* linkage group in the Swiss Landrace experimental population.

Locus Pair	Ν	θ	Z
S-ECF18R	183	0.01	50.6
M307–S	183	0.01	50.6
M307–ECF18R	216	0.01	57.1
M307–RYR1	198	0.02	47.2
M307–GPI	147	0.03	34.2
M307–PGD	147	0.04	24.5

**Table 3.** Haplotype frequencies (%) at the four loci (*S-FUT1* (*M307,M857*)-*ECF18R-RYR1*) in the Swiss Landrace (SL) experimental population and randomly selected Large White (LW) pigs.

Breed	Haplotype at S, FUT1(M307, M857), ECF18R, RYR1 <sup>a</sup>	Frequency <sup>b</sup> (Number)
SL	sAGbT	70 (28)
	sAGbC	5 (2)
	SGGBC	15 (6)
	sGABC	10 (4)
LW	sAGbC	56 (9)
	SGGBC	31 (5)
	sGGBC	13 (2)

<sup>a</sup> S: Suppressor locus for A and 0 blood types (*S* and *s*). *FUT1(M307*); alteration of adenine (A) to guanine (G) at nucleotide 307 of the  $\alpha(1,2)$ fucosyltransferase (*FUT1*) gene. *FUT1(M857)*; alteration of adenine (A) to guanine (G) at nucleotide 857 of the *FUT1* gene. *ECF18R:E. coli* F18 receptor. The dominant susceptible allele is indicated by *B* and the resistant allele by *b. RYR1:* skeletal muscle ryanodine receptor. *C* (cytosine) is the dominant resistant and *T* (thymine) the susceptible allele for malignant hyperthermia.

<sup>b</sup> Haplotype frequencies in % and absolute number of haplotypes between brackets.

tance/susceptibility status was determined on the basis of the colonization test, there would be no recombinants. Conflicting results may be due to heterogeneity in the genetic control of the F18 receptor phenotypes.

The discovery of the close linkage of M307 at FUT1 and ECF18R has made it possible to develop a diagnostic test for *E. coli* F18 adhesion resistant, heterozygous (carrier) and homozygous susceptible animals in SL and LW. In Duroc, Hampshire, and Pietrain pigs, we observed the same polymorphism at M307 as in the SL and LW (data not shown). Extended studies yield preliminary evidence that no more than 5–10% of Swiss Landrace, Large White, and Duroc pigs are resistant to infection with highly toxigenic. *E. coli* F18. The availability of a diagnostic test for the M307 mutation provides breeders with the opportunity to eliminate the  $ECF18R^B$  allele from their herds, thereby eliminating a prerequisite for *E. coli* F18 bacterial adhesion causing edema and post-weaning diarrhea. In such a selection program, identification

**Table 4.** Distribution of the genotypes, tetrachoric correlation (r) and significance of the association ( $\chi^2$  and w  $\times \chi^2$ ) of the associated polymorphic *FUT1(M307)* and *ECF18R* loci in Swiss Landrace (SL) experimental population and randomly selected Large White (LW) pigs.

		FUT1/M30	)7				
Breed	Locus	genotype	A/G	A/A	r	$\chi^2$	$\chi^2 \times w^a$
SL	<sup>b</sup> ECF18R	b/b B/b	1 106	113 1	0.98	213.1	42.6***
		genotype	A/G,G/G	A/A	-		
LW	<sup>b</sup> ECF18R	b/b B/b,B/B	0 24	5 0	1.00	29.0	11.6***

<sup>a</sup> A weight factor of w = 0.2 (SL) and 0.4 (LW) was applied to correct the lack of precision resulting from inclusion of related animals in the data according to Cotterman (1947).

\*\*\* p < 0.001.

<sup>b</sup> Animals of genotype b/b at the *ECF18R* locus are resistant and those of genotype B/b and B/B are susceptible to adhesion of F18ab *E. coli* bacteria.

of the 1843C to 1843T mutation in the *RYR1* gene has to be applied for positive diagnosis of malignant hyperthermia susceptibility in order to prevent a rise of the *RYR1* susceptibility allele in the population.

The functional relation of the *FUT1* or the *FUT2* gene to the *ECF18* receptor needs to be elucidated by studying the expression of these genes and the implications for bacterial adhesion. If *FUT1* is expressed in pigs, the observed mutations might reflect on enzyme activity and change the phenotypic expression, as seen in the Bombay and para-Bombay phenotypes in humans (Kelly et al. 1994). The  $G \rightarrow A$  changes at bp +307 and bp +857 that we observed in our material result in a predicted amino acid threonine (neutral-polar) instead of an arginine (basic), respectively, and may indeed have functional consequences.

Although we did not find mutations in the FUT2 sequence, FUT2 cannot be ruled out to be involved in the expression of the adhesion and resistance phenotypes, respectively. We are presently screening the FUT2 promoter region for potentially causal mutations.

Acknowledgments. These studies were supported in part by grants of the ETH Zurich, the Swiss Performance Testing Station Sempach (Drs. A. Rebsamen and Schwörer), the Swiss National Research Foundation, and other public and private pig breeding organizations in Switzerland. We thank Mrs. E. Wenk for excellent technical assistance and Dr. J. Rapacz for critically reviewing the manuscript. The PCR-RFLP tests to detect genetic variation within the porcine FUT1 gene and the use of these tests in breeding programs are subject to patent protection.

# References

- Bertschinger HU, Bachmann M, Mettler C, Pospischil A, Schraner EM, Stamm M, Sydler T, Wild P (1990) Adhesive fimbriae produced in vivo by *Escherichia coli* O139:K12(B):H1 associated with enterotoxaemia in pigs. Vet Microbiol 25, 265–281
- Bertschinger HU, Stamm M, Vögeli P (1993) Inheritance of resistance to oedema disease in the pig: experiments with an *Escherichia coli* strain expressing fimbriae 107. Vet Microbiol 35, 79–89
- Björk S, Breimer ME, Hansson GC, Karlsson KA, Leffler H (1987) Structure of blood group glycosphingolipids of human small intestine: a relation between the expression of fucolipids of epithelial cells and the AB0, Le and Se phenotype of the donor. J Biol Chem 257, 755–760
- Borén T, Falk P, Roth KA, Larson G, Normark S (1993) Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science 262, 1892–1895
- Clamp PA, Feltes R, Shalhevet D, Beever JE, Atac E, Shook LB (1993) Linkage relationships between ALPL, ENO1, GPI, PGD, and TGFB1 on porcine chromosome 6. Genomics 17, 324–329
- Cohney S, Mouhtouris E, McKenzie IFC, Sandrin MS (1996) Molecular

cloning of the gene coding for pig  $\alpha(1,2)$ fucosyltransferase. Immunogenetics 44, 76–79

- Cottermann LW (1947) A weighting system for the estimation of gene frequencies from family records. Contrib Lab Vertebr Biol Univ Mich 33, 1–21
- Davies W, Harbitz I, Fries R, Stranzinger G, Hauge JG (1988) Porcine malignant hyperthermia carrier detection and chromosomal assignment using a linked probe. Anim Genet 19, 203–212
- Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 1, 387–395
- Gedde-Dahl T Jr, Olaisen B, Teisberg P, Constant Wilhelmy M, Mevåg B, Helland R (1984) The locus for apolipoprotein E (apoE) is close to Lutheran (Lu) blood group locus on human chromosome 19. Hum Genet 67, 178–182
- Grange PA, Mouricout MA (1996) Transferrin associated with the porcine intestinal mucosa is a receptor specific for K88ab fimbriae of Escherichia coli. Infect immunol 64, 606–610
- Green P, Falls K, Crooks S (1990) Documentation for CRI-MAP. version 2.4. (St. Louis: Washington University School of Medicine)
- Henry SM, Samuelsson BOE, Oriol R (1994) Immunochemical and immunohistochemical expression of Lewis histo-blood group antigens in small intestine including individuals of the Le(a+b+) and Le(a-b-) nonsecretor phenotypes. Glycoconj J 11, 600–607
- Hitochi S, Kusunoki S, Kanazawa I, Tsuji S (1995) Molecular cloning and expression of two types of rabbit  $\beta$ -galactoside  $\alpha$ 1,2-fucosyltransferase. J Biol Chem 270, 8844–8850
- Källenius G, Möllby R, Svenson SB, Winberg J, Svensson S, Cedergren B (1980) The P(k) antigen as a receptor for the haemaglutinin of pyelonephritic *Escherichia coli*. FEMS Microbiol Lett 7, 297–302
- Kelly RJ, Ernst LK, Larsen RD, Bryant JG, Robinson JS, Lowe JB (1994) Molecular basis for H blood group deficiency in Bombay (O<sub>h</sub>) and para-Bombay individuals. Proc Natl Acad Sci USA 91, 5843–5847
- Kelly RJ, Rouquier S, Giorgi D, Lennon GG, Lowe JB (1995) Sequence and expression of a candidate for the human *Secretor* blood group  $\alpha(1,2)$ fucosyltransferase gene (*FUT2*). J Biol Chem 270, 4640–4649
- King TP, Kelly D (1991) Ontogenic expression of histo-blood group antigens in the intestines of suckling pigs: lectin histochemical and immunohistochemical analysis. Histochem J 23, 43–54
- Larsen RD, Ernst LK, Nair RP, Lowe JB (1990) Molecular cloning, sequence and expression of a human GDP-L-fucose:β-D-galactoside 2-α-L-fucosyltransferase cDNA that can form the H blood group antigen. Proc Natl Acad Sci USA 87, 6674–6678
- Oriol R (1987) Tissular expression of ABH and Lewis antigens in humans and animals: expected value of different animal models in the study of AB0-incompatible organ transplants. Transplant Proc 19, 4416–4420
- Oriol R, Le Pendu J, Bernez L, Lanbert F, Dalix AM, Hawkins BR (1984) An international co-operative study confirming close linkage between H and SE. Seventh International Workshop on Human Gene Mapping. Cytogenet Cell Genet 37, 564
- Oriol R, Ye Y, Koren E, Cooper DKC (1993) Carbohydrate antigens of pig tissues reacting with human natural antibodies as potential targets for hyperacute vascular rejection in pig-to-man organ xenotransplantation. Transplantation 56, 1433–1442

- Piau J-P, Labarriere N, Dabouis G, Denis MG (1994) Evidence for two distinct  $\alpha(1,2)$  fucosyltransferase genes differentially expressed throughout the rat colon. Biochem J 300, 623–626
- Reguigne-Arnould I, Couillin P, Mollicone R, Fauré S, Fletcher A, Kelley RJ, Lowe JB, Oriol R (1995) Relative positions of two clusters of human α-L-fucosyltransferases in 19q (FUT1-FUT2) and 19p (FUT6-FUT3-FUT5) within the microsatellite genetic map of chromosome 19. Cytogenet Cell Genet 71, 158–162
- Rettenberger G, Klett C, Zechner U, Kunz J, Vogel W, Hameister H (1995) Visualization of the conservation of synteny between humans and pigs by heterologous chromosomal painting. Genomics 26, 372–378
- Rippinger P, Bertschinger HU, Imberechts H, Nagy B, Sorg I, Stamm M, Wild P, Wittig W (1995) Comparison of recently described adhesive fimbriae of *Escherichia coli* isolated from postweaning diarrhea and from oedema disease: proposed designations F18ab and F18ac for the antigenic variants. Vet Microbiol 45, 281–295
- Rouquier S, Giorgi D, Brandriff B, Lennon G (1994) The fucosyltransferase locus FUT1 maps distal to apolipoprotein loci E and C2 on human chromosome 19. Cytogenet Cell Genet 66, 70–71
- Rouquier S, Lowe JB, Kelly RJ, Fertitta AL, Lennon GG, Giorgi D (1995) Molecular cloning of a human genomic region containing the *H* blood group  $\alpha(1,2)$  fucosyltransferase gene and two *H* locus-related DNA restriction fragments. J Biol Chem 270, 4632–4639
- Sellwood R, Gibbons RA, Jones GW, Rutter JM (1974) A possible basis for the breeding of pigs relatively resistant to neonatal diarrhea. Vet Rec 95, 574–575
- Solinas Toldo S, Fries R, Steffen P, Neijbergs HL, Barendse W, Womack JE, Hetzel DJS, Stranzinger G (1993) Physically mapped, cosmidderived microsatellite markers as anchor loci on bovine chromosomes. Mamm Genome 4, 720–727
- Svenson SB, Hultberg H, Källenius G, Korhonen TK, Möllby R, Winberg J (1983) P-fimbriae of pyelonephritogenic *Escherichia coli:* identification and chemical characterization of receptors. Infection 11, 73–79
- Thurin J, Blaszczyk-Thurin M (1995) Porcine submaxillary gland GDP-L-fucose:β-D-galactoside α-L-fucosyltransferase is likely a counterpart of the human Secretor gene-encoded blood group transferase. J Biol Chem 270 (44), 26577–26580
- Vögeli P, Gerwig C, Schneebeli H (1983) The A-0 and H blood group systems, some enzyme systems and halothane sensitivity of two divergent lines of Landrace pigs using index selection procedures. Livest Prod Sci 10, 159–169
- Vögeli P, Kühne R, Gerwig C, Kaufmann A, Wysshaar M, Stranzinger G (1988) Bestimmung des Halothan Genotyps (HAL) mit Hilfe der S, PHI, HAL, H, PO2, PGD Haplotypen von Eltern und Nachkommen beim Schweizerischen Veredelten Landschwein. Züchtungskunde 60, 24–37
- Vögeli P, Bertschinger HU, Stamm M, Stricker C, Hagger C, Fries R, Rapacz J, Stranzinger G (1996) Evidence for linkage of genes of the halothane (HAL) linkage group to the genes specifying receptors for F18 fimbriated *Escherichia coli* strains causing oedema disease and postweaning diarrhoea in Swiss Landrace pigs. Anim Genet 27, 321–328
- Watkins WM (1980) Biochemistry and genetics of the AB0, Lewis and P blood group systems. Adv Hum Genet 10, 1–116