RESEARCH COMMUNICATION

Correlation of a missense mutation in the human *Secretor* α 1,2-fucosyltransferase gene with the Lewis(a + b +) phenotype: a potential molecular basis for the weak *Secretor* allele (*Se^w*)

Lung-Chih YU,* Yun-Hsin YANG,* Richard E. BROADBERRY,* Yee-Hsiung CHEN,†§ Yung-Syu CHAN† and Marie LIN*†||

*Transfusion Medicine Laboratory, Department of Medical Research, †Department of Laboratory Medicine, Mackay Memorial Hospital, ‡Institute of Biochemical Science, College of Science, National Taiwan University, and §Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

A missense mutation (A^{385} to T), predicting an Ile¹²⁹ to Phe substitution, in the human *Secretor* α 1,2-fucosyltransferase gene was present in double dose in Lewis(a + b +) individuals, but not in Lewis(a – b +) individuals. Co-segregation of the Lewis(a + b +)

phenotype with homozygosity for the mutation was also verified. These results yield a potential molecular basis for the weak *Secretor* allele (*Se^w*) accounting for the Lewis(a + b +) phenotype.

INTRODUCTION

The human Lewis histo-blood-group system comprises two main oligosaccharide antigens, Le^a and Le^b, which were first discovered on red blood cells [1,2], but were later also identified in plasma and exocrine secretions, including saliva (for a review, see [3]). The Le^a [Gal β 1-3(Fuc α 1-4)GlcNAc β 1-R] and Le^b [Fuc α 1-2Gal β 1-3(Fuc α 1-4)GlcNAc β 1-R] epitopes are synthesized by the transfer of fucose to the O-4 position of GlcNAc on either type I precursor [Gal β 1-3GlcNAc β 1-R] or H type I [Fuc α 1-2Gal β 1-3GlcNAc β 1-R] respectively, through the action of α -1,3/1,4-fucosyltransferase, encoded by the *Lewis* (*Le* or FUT3) locus [4]. H type I may be synthesized from type I precursor by the α 1,2-fucosyltransferase encoded by the *Secretor* (*Se* or FUT2) locus [5–7], and thus the correlation of the Lewis phenotype with the ABH secretor status of an individual has long been observed [8]. The fact that Lewis phenotype is controlled by the *Se* and *Le* genotypes of an individual (for reviews, see [9,10]), and also the three major Lewis phenotypes identified in Caucasians [Le(a + b –), Le(a – b +) and Le(a – b –)] are a result of different combinations of the alleles of these two loci is well established [11,12].

However, a fourth Lewis phenotype, Le(a + b +), has been found among Taiwanese [13,14], Polynesians [15,16], Japanese [17], and Australian aborigines [18]. The Le(a + b +) phenotype is virtually absent in Caucasians, but has a relatively high frequency of 22–25% among Taiwanese [14,19]. Because this phenotype is associated with a reduced amount of salivary ABH substances, a weak *Secretor* allele (*Se^w*) has been proposed to be responsible for the formation of the Le(a + b +) phenotype [20]. It is thought that the *Secretor* enzyme in Le(a – b +) individuals is highly active, and converts most of type I precursor into H type I, which can then be modified into Le^b by the Lewis enzyme, whereas the weak *Secretor* enzyme expressed by the *Se^w* allele leaves a portion of the type I precursor unchanged which can then be modified into Le^a by the Lewis enzyme, thus resulting in the Le(a + b +) phenotype [21,22].

Although the *Se^w* allele has been proposed for more than 20 years, and also during these two decades more evidence for the correlation of the Le(a + b +) phenotype with the postulated *Se^w* allele has been proposed [14,16,21,22], only recently has it been possible to study the molecular basis for the *Se^w* allele as a result of cloning of the *Secretor* gene [6,7]. In the present study, homozygosity of a missense mutation (A^{385} to T) in the coding region of the *Se* gene in Le(a + b +) individuals, but not in Le(a – b +) individuals, was demonstrated. The mutation produces an amino acid substitution of Ile¹²⁹ to Phe in the corresponding *Secretor* α 1,2-fucosyltransferase. The segregation of the mutation in double dose with the Le(a + b +) phenotype was also verified. These results yield a potential molecular basis for the weak *Secretor* allele (*Se^w*) responsible for the expression of the Le(a + b +) phenotype.

EXPERIMENTAL

Materials

Mouse monoclonal anti-Le^a (LM112/161, batch 041/01/90) [23] and anti-Le^b (LM129/181 anti-Le^b^L, batch 010/01/90) [24] antibodies were a gift from Dr. R. H. Fraser (Glasgow and West of Scotland Blood Transfusion Centre, Law Hospital, Carlisle, U.K.). *Taq* polymerase, dNTP and pGEM-T vector were purchased from Promega (Madison, WI, U.S.A.). AmpliCycle Sequencing Kit was obtained from Perkin-Elmer (Foster City, CA, U.S.A.). [α -³²P]dATP and [γ -³²P]ATP were from Amersham Searle (Arlington Heights, IL, U.S.A.). Nylon membrane was purchased from Micron Separations Inc. (Westborough, MA, U.S.A.). *FokI* restriction endonuclease was obtained from Boehringer-Mannheim (Mannheim, Germany). All chemicals were of molecular-biology or reagent grade.

Sample preparation

The Lewis phenotypes of fresh blood samples from healthy Taiwanese were determined by a microplate method [19], using

Abbreviations used: Le^a, Lewis a, Gal β 1-3(Fuc α 1-4)GlcNAc; Le^b, Lewis b, Fuc α 1-2Gal β 1-3(Fuc α 1-4)GlcNAc; *Se*, *Secretor* allele; ASO, allele-specific oligonucleotide; RFLP, restriction-fragment-length polymorphism.

|| To whom correspondence should be addressed.

anti-Le^a and anti-Le^b monoclonal antibodies. Blood samples from 18 unrelated Le(a-b+) individuals, Seven unrelated Le(a+b+) individuals and two Le(a+b+) individuals' families were used in the present study. Genomic DNAs were prepared from leucocytes by a proteinase K/SDS method [25].

Molecular cloning of *Se* and *Le* alleles

PCR and a pair of specific primers for the Sec2 DNA fragment encoding the *Secretor* α 1,2-fucosyltransferase gene [7] were used to amplify the coding region of the *Se* genes of one Le(a+b+) and one Le(a-b+) individual, denoted P1 and P2 respectively. The sense primer locates at nucleotides -79 to -51 of the *Se* gene

CCTCCATCTCCCAGCTAACGTGTCCCGTT

5' to the first initiation codon, and the antisense primer

GCTTCTCATGCCCGGGCACTCATCTTGAG

is complementary of nucleotides 1042-1070 within the 3' untranslated region. A 200 ng portion of genomic DNA and 0.3 μ M of each primer were combined in 50 μ l of PCR buffer containing low concentrations of dNTP (15 μ M each) and *Taq* polymerase (0.4 unit) to minimize the PCR-mediated DNA sequence alteration. The PCR program consisted of 5 min at 94 °C, followed by 30 cycles of 1.5 min at 94 °C and 3 min at 72 °C. The 1149 bp PCR products were cloned into pGEM-T vectors.

The coding region of the *Le* gene [4] of the P1 and P2 propiti were also amplified by PCR and cloned. The sense primer

ATGGATCCCCTGGGTGCAGCCAAGCCACAAT

anneals to nucleotides 1-31 of the *Le* gene, and the antisense primer

CAGCGTGTGAGGTCCCAGGTAAGA

is complementary to nucleotides 1170-1193 within the 3' untranslated region. The PCR and cloning procedures for the *Le* gene were similar to those for the *Se* gene, except that 75 μ M of each dNTP and 0.5 unit of *Taq* polymerase were used in PCR amplification of the *Le* gene.

DNA sequences were determined by the dideoxy-chain-termination method [26], using the AmpliCycle Sequencing Kit. Multiple clones were sequenced to distinguish PCR errors from actual sequence polymorphisms.

Allele-specific oligonucleotide (ASO) hybridization and PCR/restriction-fragment-length polymorphism (RFLP) analyses

ASO hybridization and RFLP analyses were designed to detect the A³⁸⁵ to T sequence polymorphism in the *Se* allele. Nested PCR was used to sample the nucleotide position 385. The PCR products containing the whole coding region of the *Se* gene amplified as described above were used as templates for the second PCR. The pair of inner primers anneal to positions 338-367:

Sense:

GGCAGAACTACCACCTGAATGACTTGATGGAGG

and 527-555:

Antisense:

TACAAAGGTGCCCGGCCGCTCCCGTTCA

of the *Se* gene. After the first PCR, 4 μ l from a 100-fold dilution of the product was combined with 0.6 μ M of each inner primer in 50 μ l of PCR buffer containing 200 μ M of each dNTP and 1

unit of *Taq* polymerase, and then subjected to another 30 cycles of amplification (1.5 min at 94 °C, 1 min at 68 °C and 1 min at 72 °C). The sense primer was designed to alter artificially nucleotide position 362 of the *Se* gene from G to T (**bold**) to destroy a *Fok* I recognition site in the PCR product (explained in the Results section).

The PCR products of 218 bp were fixed to nylon membranes and probed with ³²P-labelled [27] ASOs:

Wild-type probe: CGCCACATCCCGG

A³⁸⁵ to T mutant probe: CGCCACTTCCCGG

Membranes were hybridized at 37 °C in 5 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), 5 $\times$ Denhardt's solution [28], 0.5% SDS and 0.1 mg/ml of sheared salmon sperm DNA. The membranes were washed once at room temperature, and three times at 47 °C in 2 $\times$ SSC and 0.1% SDS for 20 min at each interval, and subjected to autoradiography.

The PCR products were also digested with *Fok*I and analysed by 3%-agarose-gel electrophoresis.

RESULTS

A missense mutation found in the *Se* allele

The coding region of the *Se* genes of the Le(a+b+) and Le(a-b+) propiti were cloned. The DNA sequences of 1063 bp, from nucleotides -50 to 1013 of the Sec2 DNA segment, including the whole coding region of the longer isoform of the *Se* gene [7], were analysed. The inserts of eight recombinant plasmids from the Le(a+b+) propiti (P1) were sequenced, and all were found to differ from the wild-type sequence of the Sec2 DNA at four positions, denoted SM1, SM2, SM3 and SM4, as shown in Figure 1. SM1 is a T for C substitution at nucleotide position 357. This sequence alteration does not change the protein sequence predicted at this position (Asn¹¹⁹). SM2 is a missense mutation corresponding to a T for A substitution at position 385, predicting an amino acid substitution of Ile¹²⁹ to Phe. SM3 (G¹⁰⁰⁹ to A) and SM4 (C¹⁰¹¹ to T) are within the 3' untranslated region. The propiti P1 with the Le(a+b+) phenotype is most probably homozygous for the mutated *Se* allele, designated as *Sw*, containing a missense mutation SM2 in its coding region. In all, 11 clones from the Le(a-b+) propiti (P2) were sequenced, and another variant of the *Se* allele distinct from *Sw* was found and designated as *Se1* (Figure 1). The *Se1* allele comprises the SM1, SM3 and SM4 mutations, but not the missense mutation SM2. Of the 11 clones, nine presented with all of the four mutations (SM1, SM2, SM3 and SM4), and the remaining two contained the three mutations (SM1, SM3 and SM4) only, suggesting that the propiti P2 with the Le(a-b+) phenotype is a heterozygote with the *Se1*/*Sw* genotype. Since the SM1 is a silent mutation and the SM3 and SM4 mutations are within the 3' untranslated region, a wild-type *Secretor* enzyme expressed by the *Se1* allele of the Le(a-b+) propiti can be expected. Homozygosity and heterozygosity for the SM2 mutation of the P1 and P2 propiti respectively were further ascertained by ASO hybridization and PCR/RFLP analyses.

Correlation of homozygosity for the missense mutation SM2 in the *Se* allele with the Le(a+b+) phenotype

ASO hybridization analysis for detection of the sequence polymorphisms at position 385 of the *Se* allele was used to further explore whether the Le(a+b+) phenotype is accompanied by the missense mutation SM2 in the *Se* allele. As shown in Figure 2, each of the seven unrelated Le(a+b+) individuals analysed

Figure 1 Location of mutations within the *Se* allele

Sec2 indicates the coding region of the *Secretor* α 1,2-fucosyltransferase gene reported by Kelly et al. [7]. The two ATG and the TAA codons correspond to the enzyme's initiation and termination codons respectively. PCR primers used to amplify the coding region are indicated by arrows. *Se1* and *Sw* represent the variants of the *Se* allele. Point mutations found in these two variants are located.

Figure 2 ASO hybridization analysis of the A³⁸⁵ to T sequence polymorphism in Le(a + b +) and Le(a - b +) individuals

The nested PCR products of 218 bp encompassing the polymorphic site were generated from genomic DNA obtained from each of seven unrelated Le(a + b +) and 18 unrelated Le(a - b +) individuals and probed with ³²P-labelled wild-type and mutant oligonucleotides, as described in the Experimental section. Results obtained with control samples derived from cloned wild-type and mutant alleles are displayed in the boxed area. Results obtained from the Le(a - b +) proband P2 are indicated by arrows.

was homozygous for the SM2 mutation, and at least one wild-type *Se* allele at position 385 was obtained in each of the 18 unrelated Le(a - b +) individuals analysed, including the P2 proband. These results distinctly showed that homozygosity for the SM2 mutation in the *Se* allele correlates with the Le(a + b +) phenotype. PCR/RFLP analysis also corresponded with the genotypes of these 25 individuals determined by ASO hybridization analysis (results not shown).

Pedigrees of the Le(a + b +) Individuals' families

The pedigree analysis of two Le(a + b +) individuals' families, including the proband P1's family, are illustrated in Figure 3.

The *Se* genotypes of the individuals in the pedigrees were determined by PCR/RFLP (Figure 3b) and ASO hybridization analyses (Figure 3c). Since the SM2 mutation destroys a *FokI* recognition site (CATCC) locating at position 384–388 of the *Se* allele, PCR/RFLP analysis for *FokI* was used to examine their genotypes. In the amplified 218 bp product, another *FokI* recognition site at position 362–366 (GGATG) was artificially destroyed by the designed sense primer (see the Experimental section). This modification insured that the difference of length between the resistant and the digested products at the *FokI* site encompassing position 385 would be large enough to be distinguished by agarose-gel-electrophoretic analysis. Two DNA fragments, of 35 and 183 bp, were generated from the 218 bp PCR product amplified from a wild-type *Se* allele by *FokI* enzyme digestion, whereas the 218 bp product amplified from a mutated *Sw* allele was resistant to digestion. As shown in Figures 3(b) and 3(c), the *Se* genotypes determined by the PCR/RFLP analysis agreed with the result of ASO analysis. The pedigree analyses demonstrated the inheritance of the SM2 mutation of the *Sw* allele and the segregation of the SM2 mutation in double dose with the Le(a + b +) phenotype, which was further confirmed in the family 1 in which the Le(a + b +) sibling inherited the *Sw* alleles from each of his heterozygous parents.

DISCUSSION

The results reported here show a correlation of a missense mutation (A³⁸⁵ to T) in the *Se* allele with the Le(a + b +) phenotype. Therefore it would appear that the mutation in the *Se* allele is responsible for the weak *Secretor* allele (*Se*^w) resulting in the Le(a + b +) phenotype.

Although all previous evidence [14,16,20–22] has indicated that the expression of the Le(a + b +) phenotype is influenced by the *Se* locus only, we also analysed the coding sequence of the *Le* genes of the two probands (see the Experimental section), and found that the wild-type *Le* allele having a coding sequence identical with that reported by Kukowska-Latallo et al. [4] was present in both of the probands. Thus the formation of the Le(a + b +) phenotype among Lewis-positive individuals actually does not appear to be influenced by the *Le* locus. Furthermore, taking the frequencies of the three major *Le* phenotypes in Taiwanese {Le(a - b +), 67–70%; Le(a + b +), 22–25%; and Le(a - b -), 8% [14,19]} into consideration, the frequency of the *Se*^w gene among Taiwanese should be about 50%. The high

Figure 3 Segregation of *Se* alleles in two families

(a) Lewis phenotypes of the families. Open and solid symbols for male (square) and female (circle) denote an individual with the $Le(a-b+)$ and $Le(a+b+)$ phenotypes respectively. *Se* locus genotypes under each symbol are inferred from each individual's results of PCR/RFLP and ASO hybridization analyses. The $Le(a+b+)$ propositus P1 is in family 2 and indicated by an arrow. (b) PCR/RFLP analysis. The SM2 mutation in the *Se* allele destroys a *FokI* recognition site. The PCR fragments of 218 bp encompassing the polymorphic site were generated for each family member, as described in the Experimental section. The PCR products were digested with *FokI*, followed by 3%-agarose-gel-electrophoretic analysis. *FokI* cleavage of the segment from a wild-type allele yields two fragments of 35 and 183 bp, whereas the 218 bp PCR product from a mutant allele is resistant to digestion. Dc and Bc are control samples derived from cloned wild-type and mutant alleles respectively. Molecular-size standards (bp) are in lane M. (c) ASO hybridization analysis. The PCR fragments encompassing the polymorphic site were generated for each individuals and control templates (Dc and Bc) and probed with ASO probes, as described in Figure 2 and the Experimental section.

occurrence of the SM2 missense mutation in $Le(a-b+)$ individuals (Figure 2) corresponds with this prediction. These results provide further evidence that the single missense mutation SM2 in the *Se* allele accounts for the formation of the $Le(a+b+)$ phenotype.

Although, in our previous studies, the Le^b antigen strength in the $Le(a+b+)$ phenotype was found to be relatively weak compared with that in the $Le(a-b+)$ phenotype [14,19], and it is believed to be the result of a weak Secretor enzyme, nevertheless, the influence of the protein sequence alteration (Ile¹²⁹ to Phe), predicted from the missense mutation, on the enzyme activity or enzyme character of the corresponding Secretor $\alpha 1,2$ -fucosyltransferase awaits further investigation. Such research will not only directly examine whether the mutated *Se* allele, *Sw*, expresses a weak Secretor enzyme, but also provide evidence to prove the proposed mechanism of the formation of $Le(a+b+)$ phenotype. Functional analysis of the *Sw* allele is in progress.

This work was partially supported by Grant DOH-84-HR-223 from the National Institute of Health and Grant NSC-84-2331-B-195-001 from the National Science Council, Taiwan. We thank many of our colleagues of the Department of Laboratory Medicine, Mackay Memorial Hospital, for assistance in collecting the blood samples.

REFERENCES

- Mourant, A. E. (1946) *Nature (London)* **158**, 237–238
- Anderson, P. H. (1948) *APMIS* **25**, 728–731
- Watkins, W. M. (1980) *Adv. Hum. Genet.* **10**, 1–136
- Kukowska-Latallo, J. F., Larsen, R. D., Nair, R. P. and Lowe, J. B. (1990) *Genes Dev.* **4**, 1288–1303
- Oriol, R., Danilovs, J. and Hawkins, B. R. (1981) *Am. J. Hum. Genet.* **33**, 421–431
- Rouquier, S., Lowe, J. B., Kelly, R. J., Fertitta, A. L., Lennon, G. G. and Giorgi, D. (1995) *J. Biol. Chem.* **270**, 4632–4639
- Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G. and Lowe, J. B. (1995) *J. Biol. Chem.* **270**, 4640–4649
- Grubb, R. (1948) *Nature (London)* **162**, 933
- Oriol, R., Pendu, J. L. and Mollicone, R. (1986) *Vox Sang.* **51**, 161–171
- Hakomori, S.-I. (1991) *Bailliere's Clin. Haematol.* **4**, 957–974
- Mollicone, R., Reguigne, I., Kelly, R. J., Fletcher, A., Watt, J., Chatfield, S., Aziz, A., Cameron, H. S., Weston, B. W., Lowe, J. B. and Oriol, R. (1994) *J. Biol. Chem.* **269**, 20987–20994
- Nishihara, S., Narimatsu, H., Iwasaki, H., Yazawa, S., Akamatsu, S., Ando, T., Seno, T. and Narimatsu, I. (1994) *J. Biol. Chem.* **269**, 29271–29278
- Lin-Chu, M., Broadberry, R. E. and Chang, F. J. (1988) *Transfusion* **28**, 350–352
- Broadberry, R. E. and Lin-Chu, M. (1991) *Hum. Hered.* **41**, 290–294
- Henry, S. M., Simpson, L. A., and Woodfield, D. G. (1988) *Hum. Hered.* **38**, 111–116
- Henry, S. M., Benny, A. G. and Woodfield, A. G. (1990) *Vox Sang.* **58**, 61–66
- Hirano, K., Kawa, S., Oguchi, H., Kobayashi, T., Yonekura, H., Ogata, H. and Homma, T. (1987) *J. Natl. Cancer Inst.* **79**, 1262–1268
- Vos, G. H., and Comley, P. (1967) *Acta Genet.* **17**, 495–510
- Lin, M. and Shieh, S.-H. (1994) *Vox Sang.* **66**, 137–140
- Sturgeon, P. and Arcilla, M. B. (1970) *Vox Sang.* **18**, 301–322
- Henry, S. M., Woodfield, D. G., Samuelsson, B. E. and Oriol, R. (1993) *Vox Sang.* **65**, 62–69
- Henry, S. M., Oriol, R. and Samuelsson, B. E. (1994) *Glycoconjugate J.* **11**, 593–599
- Fraser, R. H., Allan, E. K., Inglis, G., Munro, A. C., Mackie, A. and Mitchell, R. (1984) *Exp. Clin. Immunogenet.* **1**, 145–151
- Fraser, R. H., Allan, E. K., Murphy, M. T., Inglis, G. and Mitchell, R. (1990) *Proceedings of the Second International Workshop and Symposium on Monoclonal Antibodies against Human Red Blood Cells and Related Antigens*, p. 73 (abstr.) Lund, Sweden
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., vol. 2, pp. 9.16–9.19, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., vol. 1, pp. 5.68–5.69, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun.* **23**, 641–646