Rapid detection of phenotypes Bombay se_{del} and nonsecretor rs200157007 SNP (302C > T) by real-time PCR-based methods

Mikiko Soejima & Yoshiro Koda

The se_{del} allele is one of the nonsecretor alleles (se) of FUT2 generated by an Alu-mediated recombination event and was first found in Indian Bombay phenotype individuals who have anti-H, anti-A, and anti-B antibodies in their serum. As well as anti-A, and anti-B antibodies, anti-H is clinically significant because it causes severe hemolytic transfusion reactions. Like se_{del}, se302 having a missense single nucleotide polymorphism (SNP), 302C > T, is characteristic of South Asians with a frequency of 10–30%. We developed a real-time PCR melting curve analysis for detection of se_{del} using a 127-bp amplicon encompassing the breakpoint junction. In addition, by performing duplex PCR by amplifying a 65-bp amplicon of the FUT2 coding region at the same time, we could determine the zygosity of se_{del} in a single tube. We also developed an Eprobe-mediated PCR assay (Eprobe-PCR) for detection of 302C > T of FUT2. These methods were validated by analyzing 58 Tamils and 54 Sinhalese in Sri Lanka. Both the duplex PCR melting curve analysis for determination of se_{del} zygosity and the Eprobe-PCR assay for detection of 302C > T exactly determined three genotypes. In addition, the results of the present methods were in complete agreement with those obtained by previously established methods. The two present methods were reliable and seem to be advantageous for large-scale association studies of FUT2 polymorphisms in South Asian populations.

In humans, expression of the H antigen, a precursor of A and B antigens of the ABO blood group, is regulated by two α(1,2)fucosyltransferases. One is the H enzyme encoded by FUT1, which regulates expression of the H antigen, and thereafter ABO antigens, on the cell surface of erythrocytes and endothelial cells. The other is the Se enzyme, encoded by FUT2, which regulates expression of the H antigen on mucosal surfaces and in body fluids. Deficiency of the H enzyme (Bombay and para-Bombay phenotypes) is very rare, whereas deficiency of the Se enzyme (nonsecretor) is common. The Bombay phenotype individuals do not express H antigen and hence A antigen or B antigen either on red cells or in secretions, because they lack both H and Se enzyme activities. As a result, they have anti-H, anti-A, and anti-B antibodies in their serum and can receive only autologous blood or blood from another Bombay phenotype individual. Transfusing blood group O red cells to them can cause severe hemolytic transfusion reactions. Functional alleles of FUT1 (H) and FUT2 (Se) are dominant over nonfunctional alleles of FUT1 (h) and FUT2 (se). FUT1 and FUT2 locate on chromosome 19q3.3 beside a pseudogene (SEC1) having high sequence similarity particularly to FUT2.

Several single nucleotide polymorphisms (SNPs) of se alleles are distributed in a population-specific manner. The 428G > A nonsense SNP (rs601338, W143X) constituting se^{428} is common in Europeans, Africans, and West Asians with a frequency about 50% and in South Asians with a frequency of 10–30%. On the other hand, the 385A > T missense SNP (rs1047781, I129F) constituting a weak secretor allele (Se*) is common in East and Southeast Asians with a frequency about 50%. In addition, the 302C > T missense SNP (rs200157007, I101P) constituting se^{302} was first identified in a Thai population with a low frequency and was demonstrated to be exclusively encountered in South Asians with a frequency of 10–30%.

Five se alleles that resulted from copy number variations (CNVs) have also been identified. Four of them (se^{del}, se_{del2}, se_{del3}, se_{del4}) were complete deletions of the coding region, whereas one (se^{delet}) was generated by a homologous recombination between SEC1 and FUT2. Among them, the se^{del} allele was first identified in subjects of the Indian Bombay phenotype, and, like the se^{302} allele, is known to be characteristic to South Asians with a frequency of 10–30%. This allele was generated by a homologous recombination between two Alu elements that are 10-kb
away from each other (Fig. 1A). Alu elements are the most abundant repetitive elements, composing ~10% of the human genome. In order to detect sedel, we first designed a conventional PCR method to amplify a relatively long fragment (1.8-kb) and then developed a triplex hydrolysis probe (TaqMan) PCR assay to detect CNVs of FUT213,14. However, sedel, sedel2, sedel3, and sedel4 cannot be discriminated by the triplex TaqMan PCR assay.

Sri Lanka has a diverse ethnic composition and 74% are Sinhalese and 18% are Tamils. We have determined genetic diversity of the FUT2 previously for the same population in this study and found that sedel, se302, and se428 were common alleles as with other South Asian populations. The frequency of sedel was reported to be 28 and 13%, that of se302 was 9.5 and 27% and that of se428 was 9.5 and 22% for Tamils and Sinhalese, respectively12.

Recent studies suggested that FUT2 polymorphism (secretor status) is associated with susceptibility to various infectious diseases, such as norovirus, rotavirus, COVID-19, and several clinical conditions such as Crohn’s disease and low plasma vitamin B12 levels15–18. Large scale replication studies of various populations or independent samples are important for confirmation of these associations. Therefore, accurate and high-throughput genotyping should be performed. However, common nonsecretor alleles are not shared by different continental populations.

An Eprobe-mediated PCR method (Eprobe-PCR) was recently developed for detection of SNPs. Eprobe is a hybridization-dependent fluorescence probe based on the quenching of two dye moieties in the condition of a single-stranded oligonucleotide and can be applied to sequential quantitative PCR, followed by melting curve analysis in a single reaction tube with a real-time PCR instrument19.

The aim of present study was to develop high-throughput methods for detection of FUT2 polymorphisms applicable to South Asians and to examine the molecular basis of Indian Bombay phenotype in more detail. For this purpose, we developed a duplex real-time PCR melting curve analysis for detection of sedel using a short (127-bp) amplicon together with the FUT2 coding region using a 65-bp amplicon to determine sedel zygotosity in a single tube. We also developed an Eprobe-PCR method for detection of 302C>T of FUT2 using a 195-bp amplicon.
Results

The Lewis phenotype on red cells of each individual (58 Tamil and 54 Sinhalese) had been determined previously. Le(a−b+) was identified as a secretor and Le(a+b−) as a nonsecretor whereas discrimination of secretors from nonsecretors among 26 Le(a−b−) subjects by phenotyping is impossible. There was no discrepancy between phenotype and genotype determined by Sanger sequencing and denaturing high-performance liquid chromatography (dHPLC) in Lewis-positive subjects. As described previously, we could define that 16 of 26 Le(a−b−) subjects were secretors and ten of them were nonsecretors by genotyping of the FUT2.

PCR amplification of se

First we amplified a PCR product using a set of primers encompassing a 127-bp region of an se breakpoint (se-F and se-R) on a real-time PCR platform (Fig. 1A). Specific amplification of the deletion breakpoint of se was confirmed by an amplification signal only from individuals having se and direct DNA sequencing of the PCR products of four selected subjects (data not shown).

Duplex real-time PCR method for detection of se

We then designed a duplex real-time PCR to determine the zygosity of se in a single tube. In addition to primers for the 127-bp se specific amplicon, we added primers for detection of FUT2 that lacked the se allele in a single tube. We performed this on three selected individuals with genotypes of the wild type (+/+), heterozygote of se (+/−), and homozygote of se (−/−). Since these primers amplified a 65-bp region of the coding region of FUT2 (751–815 bp), the melting curve analysis of the duplex real-time PCR clearly distinguished the three genotypes from each other (data not shown). The melting temperature (Tm) value of the 127-bp amplicon of se was around 85 °C, while that of 65-bp FUT2 coding region was around 81 °C. The lower limits of discrimination were around 16, 64, and 8 pg of DNA for − (homozygote of se), +/− (heterozygote of se), and +/+ (wild type), respectively (data not shown). We then applied this method to analyze 58 Tamil and 54 Sinhalese samples and clearly discriminated genotypes from each other (data not shown). The results suggested the specific amplification of but not FUT2.

Genotyping of 302C>T of FUT2 using Eprobe-PCR.

Next, we performed melting curve genotyping using a 195-bp amplicon and a 25-bp Eprobe (see Fig. 1B). To amplify FUT2 specifically, we selected a reverse primer with a seven-base difference from SECI (which is identical to the reverse primer of an unlabeled probe based on high-resolution melt (HRM) analysis for detection of 385A>T of FUT2). Using this method, we clearly distinguished C/C (Tm: around 72 °C), C/T (Tm: around 63 °C and 72 °C), and T/T (Tm: around 63 °C) genotypes from each other (data not shown). We then applied it to 58 Tamils and 54 Sinhalese whose 302C>T genotypes had been determined by Sanger sequencing and dHPLC previously. As shown in Fig. 3, the three genotypes of 302C>T of FUT2 could be separated clearly, the results were fully in agreement with previous genotyping results, and the numbers of C/C, C/T, and T/T were 43, 27, and 3 for Tamils and 38, 16, and 0 for Sinhalese. The repeatability was confirmed by two independent assays.

Discussion

We recently developed several HRM-based real-time PCR methods for detection of se, se, and se alleles. Predominant se alleles in several South Asian populations are se, se, and se. In addition to these three se alleles, se has relatively high frequency in Bangladesh. Therefore, we should genotype se, se, and se in many South Asians, whereas se, se, and se should be genotyped in Bangladesh for association studies of FUT2.

Even when we do not consider the deletion allele (se), the estimation of secretor status is not affected. For example, the Se/− genotype is judged as Se/Se and the se− genotype as se/se. However, we need to be alert for se enrichment when targeting a population with high frequency of it. In this study, we did not detect any real-time PCR amplification signal of 302C>T in three Tamils with the −/− genotype (Fig. 3A). The results persuaded us to screen for se in South Asians; otherwise, we could not know whether the reason for the lack of an amplification signal of FUT2 was actually an absence of FUT2 or another problem, such as degradation of genomic DNA. In addition, it is likely that we overestimated homozygotes. In fact, without considering the results of se screening, we misjudged the 302C/− genotype of 19 Tamils and 13 Sinhalese as C/C and 302T/− genotype of 6 Tamils and 3 Sinhalese as T/T by the Eprobe-PCR assay.

Unfortunately, we overlooked three individuals with the +/− genotype by previous conventional PCR for detection of an 1.8-kb amplicon because of the relatively large size of the PCR product and by simplex PCR without an amplification control. On the other hand, the present duplex real-time PCR melting curve analysis included an amplification control (65-bp FUT2 coding sequence) that also allowed determination of se zygosity in a single tube. Therefore, the present duplex real-time PCR assay for detection of se is more reliable and faster than previous conventional PCR methods.

We previously developed a triplex TaqMan PCR assay to detect CNVs of FUT2. The advantage of this method is not only detection of known CNVs but the potential to detect novel CNVs. However, it depends largely...
on both the quality and quantity of DNA to work well and carries a cost in terms of three probes. On the other hand, the real-time PCR melting curve analysis we present here is dedicated to detection of the $s_{cdl}$ and unable to detect other CNVs. The present assay is cost-effective, easy to use, straightforward, and not very dependent on both quality and quantity of DNA.

Because a 164-bp sequence surrounding 302C > T is completely identical to that of SEC1 and this sequence contains another high frequency SNP (about 50% in global populations), 357C > T (rs281377, synonymous SNP), it was difficult to select appropriate primers for short amplicon HRM to detect 302C > T of $FUT2$. For this reason, in this study, we employed Eprobe-PCR instead of HRM analysis for detection of 302C > T of $FUT2$.

Compared with HRM analysis using a short amplicon, Eprobe-PCR needs a labeled probe, and is therefore more expensive. However, the HRM method is based on detection of subtle differences of the melting curve and melting temperature of PCR amplicons, whereas the Eprobe-PCR method is based on detection of relatively large differences of the melting curve and melting temperature of a short probe sequence. Therefore, a probe-based melting curve analysis seems to be one of the most specific and sensitive methods to detect SNPs. In fact, the Tm values of the wild-type (around 72 °C) and that of the mutant (around 63 °C) were quite different (around 9 °C). This significant difference made a clear distinction of the three genotypes of C/C, C/T, and T/T possible.

Thus we believe that the present Eprobe-PCR for detection of 302C > T of $FUT2$ is quite useful and reliable.

In conclusion, the present two protocols seem to be a reliable and high throughput method for detection of $s_{cdl}$ and $s_{c02}$ in South Asian subjects and for examination of genetic basis of Indian Bombay phenotype in more detail.

**Figure 2.** Melting peak profiles of duplex PCR for detection of $s_{cdl}$ zygosity. Melting peak profiles obtained for 58 randomly selected Tamils (A) and 54 randomly selected Sinhalese (B). The individuals with genotypes of ++/− (wild type, blue), +/− (heterozygote of $s_{cdl}$, red), and −/− (homozygote of $s_{cdl}$, green) were completely separated by melt curve genotyping.
Materials and methods

Statements and DNA samples. All methods were carried out in accordance with relevant guidelines and regulations. DNA samples from 58 Tamil and 54 Sinhalese in Sri Lanka whose FUT2 genotypes had been already determined were used. Genomic DNA was extracted from frozen whole blood using Gentra Puregene Blood Kit (Qiagen, Tokyo, Japan). The oral informed consent was obtained and the DNA samples were taken from participants in 2002. The statement for oral informed consent approved by ethical committee of Kurume University in 2002. However, present study protocol was approved by the ethical committee of Kurume University School of Medicine in 2017 using existing and already anonymized DNA samples (Bioethics approval No. 342).

Duplex real-time PCR melting curve analysis for detection of se
del. Since both the 5′ and 3′ deletion breakpoints of se
del are located within Alu-repetitive elements, the primers were carefully designed to amplify only a recombination allele (se
del) but not other Alu-elements using Primer 3 (http://bioinfo.ut.ee/ primer3-0.4.0/) and a BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Fig. 1A). For amplification of the 127-bp amplicon of se
del encompassing a homologous sequence of 25 bp at each breakpoint, we selected a se
del-F primer (5′-TCTCAGTAGAGACAGCTGG-3′, 35–11 bp upstream from the 5′ end of the 25-bp homologous sequence) and a se
del-R primer (5′-GACAGGTTTCACCATGTCGAC-3′, 46–67 bp downstream from the 3′ end of the 25-bp homologous sequence) from several candidates. For duplex PCR to determine the zygosity of se
del, we added primers which were recently designed for HRM analysis (FUT2-778-F: 5′-TTTGCTGCGATGGAGATT-3′ and FUT2-778-R: 5′-TGTTACCTGTGATTGAGCA-3′) to detect 778C > del (P260Lfs*16, rs1799761) as FUT2-specific primers. These primers amplified the 751–815 bp coding region of FUT2 which lacked the se
del allele (Fig. 1A). The primers were synthesized by Eurofins Genomics (Tokyo, Japan). The 10 µL PCR reaction contained 1–10 ng genomic DNA, 5 µL of LightCycler 480 HRM Master mix (Roche Diagnostics, Mannheim, Germany), 2.5 mM MgCl₂, 125 nM of se
del-F and se
del-R primers, and 187.5 nM of FUT2-778-F and FUT2-778-R primers. Touchdown PCR was performed on a LightCycler 480 instrument II.
Sanger sequencing of PCR products of se\textsuperscript{del}. PCR amplification was performed using sedel-F and sedel-R primers and LightCycler 480 HRM Master mix as described above. Direct sequencing of the PCR amplicons of se\textsuperscript{del} was performed using each PCR primer as a sequence primer.  

**Eprobe-PCR for detection of 302C>G of FUT2.** Nucleotide positions of the SECI and FUT2 genes are numbered as described previously. For PCR amplification surrounding 302C>G, Primer3 (https://www.bioinfo.ut.ee/primer3-0.4.0/) was also used to design primers for specific amplification of FUT2. In addition, Edesign software ver.2.00 (http://www.dnaform.com/edesign2/) provided by K.K. DNAFORM (Yokohama, Japan) was used to design an Eprobe for detection of 302C>G of FUT2. For Eprobe-PCR, we used a FUT2-302-F primer (5′-CCCTGGCAAAGATGAAG-3′, 218–236 bp of FUT2) and identical with the corresponding SECI sequence, Fig. 1B underlined) a FUT2-302-R primer (5′-CCGGTAGAAGCGACGTACT-3′, 395–412 bp of FUT2, a 7-base difference from SECI, Fig. 1B underlined) and Eprobe (5′-TCTTCCAGAAUCACCTGCGGTCT-3′-AmC3; U indicates the position of the modified T by thiazole orange, 284–308 bp of FUT2). The Eprobe was blocked on the 3′ end (3′-amino-modifier C3) to prevent extension during PCR. The primers were synthesized by Eurofins Genomics, and the Eprobe was synthesized by K.K. DNAFORM. We performed real-time PCR and melting curve analysis using a LightCycler 480 Instrument II. Asymmetric PCR amplification was performed in 10 μL reaction mixture including 1–10 ng of genomic DNA, 5 μL of E-Taq 2× PCR Mix (K.K. DNAFORM), 50 nM of FUT2-302-F primer, 250 nM of FUT2-302-R primer, and 250 nM of the Eprobe. The thermal profile was as follows: one cycle at 95 °C for 30 s, followed by 50 cycles with denaturation at 95 °C for 15 s, annealing at 58 °C for 30 s, and extension at 72 °C for 15 s. The fluorescence data for monitoring real-time PCR amplification were collected at the end of the annealing step of each cycle using a filter (533–580 nm). The products were heated to 95 °C for 1 min, rapidly cooled to 45 °C for 1 min, and fluorescence data for melting curve analysis were collected over the range from 50 to 80 °C. The melting curve genotypes were automatically clustered into separate groups by LightCycler 480 Gene Scanning Software (Roche Diagnostics).
17. McGovern, D. P. et al. Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn’s disease. *Hum. Mol. Genet.* **19**, 3468–3476. https://doi.org/10.1093/hmg/ddq248 (2010).
18. Hazra, A. et al. Common variants of FUT2 are associated with plasma vitamin B12 levels. *Nat. Genet.* **40**, 1160–1162. https://doi.org/10.1038/ng.210 (2008).
19. Hanami, T. et al. Eprobe mediated real-time PCR monitoring and melting curve analysis. *PLoS One* **8**, e70942. https://doi.org/10.1371/journal.pone.0070942 (2013).
20. Nordgren, J. et al. Both Lewis and secretor status mediate susceptibility to rotavirus infections in a rotavirus genotype-dependent manner. *Clin. Infect. Dis.* **59**, 1567–1573. https://doi.org/10.1093/cid/ciu633 (2014).
21. Soejima, M. & Koda, Y. Detection of the weak-secretor rs1047781 (385A>T) single nucleotide polymorphism using an unlabeled probe high-resolution melting-based method. *Electrophoresis* **42**, 1362–1365. https://doi.org/10.1002/elps.202000386 (2021).
22. Soejima, M. & Koda, Y. High-resolution melting analysis for detection of fusion allele of FUT2. *Electrophoresis* **42**, 315–318. https://doi.org/10.1002/elps.202000241 (2021).
23. Soejima, M. & Koda, Y. Estimation of secretor status of ABO antigens by high-resolution melting analysis of rs601338 (428G>A). *Clin. Chim. Acta* **517**, 86–91. https://doi.org/10.1016/j.cca.2021.02.019 (2021).
24. Parker, E. P. K. et al. FUT2 secretor status is not associated with oral poliovirus vaccine immunogenicity in south indian infants. *J. Infect. Dis.* **219**, 578–581. https://doi.org/10.1093/infdis/jiy553 (2019).
25. Erali, M., Voelkerding, K. V. & Wittwer, C. T. High resolution melting applications for clinical laboratory medicine. *Exp. Mol. Pathol.* **85**, 50–58. https://doi.org/10.1016/j.xom.2008.03.012 (2008).
26. Untergasser, A. et al. Primer3—New capabilities and interfaces. *Nucleic Acids Res.* **40**, e115. https://doi.org/10.1093/nar/gks596 (2012).
27. 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 alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J. Biol. Chem.* **270**, 4640–4649 (1995).

Acknowledgements
We thank Ms. Katherine Ono for the English editing of this manuscript.

Author contributions
M.S. performed experiments and contributed to the data analysis and drafting of the manuscript. Y.K. planned the experiment, contributed to the data analysis and interpretation, and drafted the paper.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Y.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021