Fiber-FISH Analysis Revealed Genomic Variations in the RHD as a Cause of Serologically Rh-negative Phenotype

Yumiko Suto¹,*, Momoki Hirai², Hironobu Hyodo¹, Hitoshi Okazaki¹, Masahiro Satake¹, Kenji Tadokoro¹, Makoto Uchikawa³, Takeo Juji¹, and Yoshihide Ishikawa¹

¹ Department of Research and Development, Japanese Red Cross Central Blood Institute, Tokyo, Japan
² International Research and Educational Institute for Integrated Medical Sciences, Tokyo Women's Medical University, Tokyo, Japan
³ Japanese Red Cross Tokyo Metropolitan Blood Center, Tokyo, Japan

Received July 1, 2009; accepted July 31 2009

Summary The human Rhesus (Rh) blood group is determined by 2 highly homologous RH genes, RHD and RHCE. These genes are located on chromosome 1p36.1, a genomic region representing copy number variation. In Rh-positive individuals, both RHD and RHCE genes are present, while in most Rh-negative individuals RHD genes are missing and are homozygous for a single RHCE gene. Some individuals with a serologically Rh-negative phenotype were diagnosed to be RHD-gene positive by a locus-specific DNA assay using polymerase chain reactions with sequence-specific primers (PCR-SSP). Therefore, the current PCR-based assay, examining the presence or absence of partial sequences of RHD genes, does not necessarily produce correct typing results. Alterations in the genomic structure of RHD that cause negative gene expressions are difficult to detect by the PCR-based assay because of highly homologous sequences of the 2 RH genes. In the present study, we examined, by extended chromatin fiber fluorescence in situ hybridization (fiber-FISH) with 2 DNA probes for introns 3 and 7 of the RH genes, the genomic structure of the RH gene region in 6 Rh-negative Japanese donors whose genotype was diagnosed as RHD-gene positive by the PCR-SSP assay. We demonstrated that varied sizes of deletion and/or insertion in the RH loci caused the Rh-negative phenotype.

Key words Copy number variation, Fiber-FISH, Gene rearrangement, Rhesus blood group system.

Rh blood group system is one of the most important erythrocyte antigen systems. Immunological and biochemical studies have demonstrated that the Rh antigens, D, C/c and E/e, are present on 3 homologous but distinct 32-kDa polypeptides (Le Van Kim et al. 1992a). Clinically speaking, incompatibility of blood type between the presence (Rh-positive type) and absence (Rh-negative type) of the D antigen can cause severe hemolytic reactions between a transfusion donor and a recipient, or between a mother and her fetus. It is also known that the frequency of each haplotype varies among different populations (Daniel 2002).

The gene organization of Rh-positive individuals consists of 2 RH genes, RHCE encoding C/c and E/e antigens, and RHD encoding D antigen. Alternative splicing of a primary transcript of the RHCE gene is the mechanism of encoding the C/c and E/e polypeptides with a single gene (Le Van Kim et al. 1992a, Mouro et al. 1994). Both RHCE and RHD genes are composed of 10 exons and their sequence identities in the mRNAs and genomic DNAs are about 96% and 92%, respectively (Colin et al. 1991, Le Van Kim et al. 1992a, Cherif-Zahar et al. 1994, Okuda et al. 1997). High-
sequence homology and the same gene location on chromosome 1p36.1 (Cherif-Zahar et al. 1994, Suto et al. 2000) suggest that RHD and RHCE originated from the duplication of a common ancestral gene. An RHCE gene and an RHD gene were found to face each other with 3’ of their ends within a genomic region of less than 200 kb (Suto et al. 2000, Wagner and Flegel 2000). The precise genomic structure of this RH gene cluster region was first unveiled by direct analysis using extended chromatin fiber fluorescence in situ hybridization (fiber-FISH) with 2 genomic DNA probes of the introns 3 and 7 of the RH gene (Suto et al. 2000). Most Rh-negative individuals lack RHD genes. In other words, they are homozygous for a single RHCE gene (Colin et al. 1991, Suto et al. 2000, Wagner and Flegel 2000). The complete deletion of the RHD gene in an Rh-negative allele could be explained as a result of an unequal crossover event.

In an earlier stage of DNA typing studies of RH, Southern blot analysis (Colin et al. 1991) was conducted. Recently, polymerase chain reaction with sequence-specific primers (PCR-SSP) for RH genes has increasingly been utilized for the prenatal diagnosis and the analysis of cell kinetics after transfusion as an alternative assay system for serological typing, particularly among Caucasians (Prager et al. 2007). Confusingly, some serologically Rh-negative individuals were found to be RHD-gene positive by the PCR-SSP assay. Alterations in the genomic structure of the RHD gene, that result in the negative expression of the gene, were not detected by the PCR-based assay, possibly because of the duplicated genomic structure of this region. This may be one of causes of the discordant typing results between serological and DNA assay systems.

In Caucasians, the D-antigen negative phenotype with RHD-gene positive alleles is rare (Hyland et al. 1994, Wagner et al. 2001). On the other hand, East Asians and South Africans are different (Grootkerk-Tax et al. 2005), and that makes it difficult to apply the PCR-SSP assay using well-established primer sets suitable for Caucasians (Wagner et al. 2001) compared to other groups. Okuda et al. (1997) reported that in the Japanese, as high as 27.7% of Rh-negative donors showed the presence of the RHD gene. Nonsense mutations in the RHD gene or formation of hybrid alleles are thought to cause the negative expression of the RHD gene. The RHD-gene positive but D-antigen negative alleles are also found in the Chinese (Sun et al. 1998, Lan et al. 2000, Peng et al. 2003, Ye et al. 2007). Singleton et al. (2000) reported that, of 82 Rh-negative native South Africans, 54 (66%) had an RHD pseudogene that carries 37-bp insertion at the intron 3/exon 4 boundary, resulting in the gain of a stop codon. On the other hand, 13 (15%) of them had an RHD-CE-D hybrid gene, and only 15 (18%) of them completely lacked the RHD gene. Therefore, the presence of some other causal mechanisms such as deletions and insertions has been suspected.

In the previous fiber-FISH study (Suto et al. 2000), we found that 3 RHD-gene positive but Rh-negative Japanese (D−C+c+E+e+, D−C+c+E−e, and D−C+c−E−e+) showed complex genomic rearrangements including: partial deletion, duplication and recombination in the RHD gene region, despite the presence of RHD detected by the PCR-SSP assay. These complicated genomic structures associated with the Rh-negative phenotype could not have been detected by PCR-based assays alone. In the present study, we collected blood samples from serologically Rh-negative Japanese donors who were diagnosed by PCR-SSP as RHD-gene positive to precisely determine the correlation between the Rh phenotype and genotype.

Materials and methods

Blood samples

Peripheral blood samples were obtained from 6 unrelated healthy Japanese donors who were diagnosed serologically to be Rh-negative by standard techniques and reagents, and were found to be RHD-gene positive by PCR-SSP for the intron 4, exon 7 and exon 10 of the RHD gene, as previously described (Suto et al. 2000). We confirmed that no samples showed 37-bp insertion by the DNA sequencing of the exon 4 of the RHD gene (Singleton et al. 2000). One common Rh-positive
and one common Rh-negative blood samples were also prepared as controls.

**Fiber-FISH**

A plasmid clone containing a 1.55-kb cDNA of the *RHD* gene (GenBank accession no. L08429) was prepared for chromosome mapping. Two genomic DNA probes containing the intron 3 and intron 7 of the *RH* genes, respectively, were prepared by PCR for the fiber-FISH assay (Suto et al. 2000). DNA probes for the intron 3 (12 kb) and intron 7 (11 kb) were labeled with biotin (Invitrogen) and digoxigenin (Roche), respectively, by the nick-translation. The total size of these 2 probes (23 kb) was about 40% of the genomic region of 1 *RH* gene (60 kb). The probe for intron 3 was hybridized with intron 3 regions of both the homologous *RH* genes (*RHCE* and *RHD*) because of their high sequence identity. The same holds for the probe of intron 7. The hybridization efficiency of these probes was almost 100% (Suto et al. 2000).

Peripheral blood lymphocytes were stimulated with phytohemagglutinin, cultured for 3 d and harvested according to the standard procedures. FISH was performed on metaphase chromosomes using the cDNA probe to confirm the location of Rh genes. Chromatin DNA fiber preparations and fiber-FISH with the 2 intron probes were performed as previously described (Suto et al. 2000). Human Cot-1 DNA (Invitrogen) was used as a competitor. Hybridization signals were amplified using the combinations of rabbit anti-biotin antibody/FITC-goat anti-rabbit IgG (Enzo), and sheep anti-digoxigenin (Roche)/Cy3-donkey anti-sheep IgG (Jacksons). Finally, slides were counterstained with DAPI II (Abbott).

**Image capture and analysis of fiber-FISH signals**

Hybridization signals were detected with a fluorescence microscope (BX60, Olympus) equipped with a CCD camera (SPOT, Diagnostic Instruments Inc.). At least 20 images per sample were captured and processed with IPLab Spectrum v. 3.5.1 software (Scanalytics Inc.) for analysis. The lengths of signal arrays and intervals between the signals were measured on digitized images.

**Results**

**Chromosomal localization of the RH locus**

FISH signals of the labeled probe of a cDNA clone for the *RH* gene were observed on both homologs of chromosome 1 at band p36.1 in all samples (data not shown). This indicated that the *RH* gene regions were not involved in any chromosomal rearrangements such as inversions or translocations. Since every hybridization signal was large enough, the chromosomal region of *RH* gene clusters did not contain complete or considerable sizes of deletion, which, if any, might have caused signal size reduction.

**Fiber-FISH**

A control sample with the typical Rh-positive phenotype showed 2 sets of green (intron 3) and red (intron 7) signal arrays arranged in the opposite direction, denoting the following gene arrangement: *RH* (5′→3′)–*RH* (3′→5′) (Fig. 1). This was concordant with results described earlier (Suto et al. 2000, Wagner and Flegel 2000). A control sample with typical Rh-negative phenotype showed 1 set of the array (Fig. 1) As for the samples tested in the present work, we found the following 3 different signal patterns: type (I), 1 set of the signal array; type (II), 2 sets of the signal array with varied sizes either in one of the red signals or in the interval between the red and green signals; and, type (III), miscellaneous patterns that were not classified into I or II. Irrespective of varied signal sizes in type (II), the gene orientations of *RH* (5′→3′)–*RH* (3′→5′) were the same as those seen in typical Rh-positive alleles (Suto et al. 2000, Wagner and Flegel 2000). According to the genomic database of this gene region (UCSC Human Genome Browser, hg18, http://genome.ucsc.edu/), the
Fig. 1. Schematic representation of the arrangement of the human RH locus based on the database of UCSC Human Genome Browser (hg18, http://genome.ucsc.edu/), modified (a). Fiber-FISH signal arrays of intron 3 probe (green) and intron 7 probe (red) of RH genes on stretched chromatin fibers from cells (b). Bar indicates \(-10\mu m (30\, kb)\) when a degree of stretching was 90–100%. The size estimation of signal arrays and intervals was made by calculating the relative lengths of those to the longer signal arrays (i.e. RHCE). Magnification of some images was slightly changed so that normal RH (i.e. RHCE) signal arrays showed almost the same size. Note that 1 allele of JR-04 showed \(-30\,-30\) kb deletion between the 2 signal arrays of intron 7.
distance between the 3′-end of intron 3 and the 5′-end of intron 7 in RHCE (40 kb) is slightly longer than that of RHD (38 kb). However, these 2 signal arrays from 2 different genes were not discernible under the conditions employed in this fiber-FISH study. Since all subjects tested in the present work showed normal C/E serotypes, the signal arrays showing aberrant sizes were regarded as those of aberrant RHD genes, as evidenced in our previous work (Suto et al. 2000). The chromatin fiber preparations used for the present analysis exhibited a degree of stretching of 80–100%, which was estimated from the ratio of the observed lengths of signal arrays to the expected lengths of probe DNA sizes based on an assumption that the physical length of one base-pair is 0.34 nm. The estimation of short (deleted) signal arrays and the insertions between signals was made by calculating the relative lengths of those to the longer signal arrays. Distribution of these patterns is shown in Fig. 1 and Table 1.

Frequencies of type (III) in the present 6 donors were found to be 10% or less. The majority of signal arrays observed in each of these samples were determined as a 2-RH gene type. Therefore, we identified type-III signals as artifacts derived from type (II) in which fibers were physically damaged through experimental procedures or due to insufficient hybridization and/or signal detection.

From the distribution of signal patterns in types (I) and (II), and the estimates of signal length and distance, the genomic structure of the 6 Japanese donors was deduced as shown below and listed in Table 2. Rough estimates of structural alterations are indicated in parentheses.

### Table 1. Distribution of fiber-FISH signal patterns of RH genes in donors who were identified as Rh-negative by serological typing and RHD-positive by PCR typing

<table>
<thead>
<tr>
<th>Donor</th>
<th>Serological phenotype</th>
<th>Distribution of signal patterns§</th>
<th>No. of fibers observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR-01</td>
<td>D–C+c–E–e+c</td>
<td>0.17 0.83 0.05 24</td>
<td></td>
</tr>
<tr>
<td>JR-02</td>
<td>D–C+c–E–e+c</td>
<td>0.50 0.45 0.05 20</td>
<td></td>
</tr>
<tr>
<td>JR-03</td>
<td>D–C+c–E–e–</td>
<td>0.45 0.50 0.05 20</td>
<td></td>
</tr>
<tr>
<td>JR-04</td>
<td>D–C+c–E–e–</td>
<td>0.25 0.65 0.10 20</td>
<td></td>
</tr>
<tr>
<td>JR-05</td>
<td>D–C+c–E–e–</td>
<td>0.49 0.42 0.09 35</td>
<td></td>
</tr>
<tr>
<td>JR-06</td>
<td>D–C+c–E–e–</td>
<td>0.20 0.70 0.10 20</td>
<td></td>
</tr>
</tbody>
</table>

§ Hybridization signal patterns of fibers are classified as follows: I: 1-RH gene type. II: 2-RH gene type arranged in the order of RH (5′→3′)–RH (3′→5′) with partial deletion in 1 RH gene. III: Other patterns that were not classified in any of I and II. All of them seen in this study had 2 RH genes either arranged in the order of RH (intron 3→intron 7)–RH (intron 3) or RH (intron 3→intron 7)–RH (intron 7). Therefore, they were likely to be broken artifacts through experimental procedures derived from type II in this table.

### Table 2. Fiber-FISH typing for Rh-negative and RHD-positive donors

<table>
<thead>
<tr>
<th>Donor</th>
<th>Fiber-FISH typing</th>
<th>allele 1</th>
<th>allele 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>JR-01</td>
<td>CE+D*</td>
<td>CE+D*</td>
<td></td>
</tr>
<tr>
<td>JR-02</td>
<td>CE</td>
<td>CE+D*</td>
<td></td>
</tr>
<tr>
<td>JR-03</td>
<td>CE</td>
<td>CE+D*</td>
<td></td>
</tr>
<tr>
<td>JR-04</td>
<td>CE+D* (del: ~30 kb)</td>
<td>CE+D* (del: ~11 kb, ins: ~11 kb)</td>
<td></td>
</tr>
<tr>
<td>JR-05</td>
<td>CE</td>
<td>CE+D*</td>
<td></td>
</tr>
<tr>
<td>JR-06</td>
<td>CE+D* (del: &lt;10 kb, ins: ~15 kb)</td>
<td>CE+D* (del: &lt;10 kb, ins: ~15 kb)</td>
<td></td>
</tr>
</tbody>
</table>

*: Partial deletion, ‡: Partial insertion, del: deletion, ins: insertion. Sizes of deletion or insertion without notice were less than 10 kb.
JR-01: homozygous for an allele with a normal RHCE and an RHD with partial deletion including part of intron 7 (<10 kb) [RHCE-RHD*/RHCE-RHD*; D* denotes deleted D].

JR-02: an allele with a single normal RHCE and another allele with an RHCE and an RHD with deletion including part of intron 7 (<10 kb) [RHCE/RHCE-RHD*].

JR-03: an allele with a single RHCE and another allele with an RHCE and an RHD with partial deletion in intron 7 (<10 kb), accompanied with inserted sequences between intron 7 of RHCE and intron 3 of RHD (<10 kb) [RHCE/RHCE-ins-RHD*].

JR-04: an allele with an RHCE and an RHD, accompanied with partial deletion in interval of two RH genes (~30 kb), and another allele with an RHCE and an RHD with deletion including part of intron 7 (~11 kb) and additional sequences in the region between the RH genes (~11 kb) [RHCE RHD/RHCE-ins-RHD*].

JR-05: an allele with a single RHCE gene and an allele with a normal RHCE gene and an RHD with deletion including part of intron 7 (<10 kb). [RHCE/RHCE-RHD*].

JR-06: homozygous for an allele with an RHCE and an RHD with deletion including part of intron 7 (<10 kb) and insertion in the region between RHCE intron 7 and RHD intron 3 (~15 kb). [RHCE-ins-RHD*/RHCE-ins-RHD*].

Discussion

Since the complete sequencing of the human genome was accomplished (Human Genome Consortium 2003), interests have shifted to focus on the elucidation of individual genomic variation. Knowledge of human genomic diversity has accumulated drastically by detecting single nucleotide polymorphisms (SNPs), copy number variations (CNVs), and large-scale genomic structural variations (SVs) [Iafrate et al. 2004, see The Genome Structural Variation Consortium (2009): Database of Genomic Variants (http://projects.tcag.ca/)]. Among others, CNVs attract attention with respect to their involvement in the phenotypic consequences such as diseases. The major mechanism underlying the occurrence of CNVs is thought to be genomic duplications. Because of highly homologous sequences between different copies of genes in the CNV-regions, especially in the diploid genomic state, the precise analysis of the genomic structure of these regions is often difficult by sequence-dependent assay systems. As for the newly devised high-density microarray that aims to precisely quantify copy numbers by comparative genomic hybridization (CGH), the probe density of CNV-regions is low because unique site-specific oligo-nucleotide probes are difficult to design. The RH locus represents such CNV-regions, and therefore encounters discrepancies between the genotype determined by the PCR assay and the phenotype determined by the serological assay. For the detection of such genomic variations, the fiber-FISH is a powerful tool that directly visualizes the genomic alterations.

We performed fiber-FISH analysis using DNA probes for introns 3 and 7 of the RH genes for a total of 6 Japanese donors who were Rh-negative serologically but RHD-gene positive by PCR-SSP typing. It should be noted that every donor had alleles with different types of genomic alterations, such as deletion (2.5–11 kb, varying in size among different individuals) in intron 7, 30-kb deletion or 5–15-kb insertions within the region between RHCE intron 7 and RHD intron 3. The estimated sizes of the deletions or insertions described above were not accurate because the estimates are based on the relative lengths of signal arrays or signal intervals on the DNA fibers. We used only 2 intronic probes for the fiber analysis. Although further analyses by DNA re-sequencing of the genomic regions of RH genes are needed to precisely determine the nature of genomic alterations, our present results suggest the presence of several hot spots of genomic rearrangements in the RH gene region. Present sequencing methods using conventional Sanger sequencers or next-generation sequencers are insufficient at establishing the precise assembly of haplotype sequences for those duplicated regions of the human genome; therefore, development of DNA sequencing methods evading multiple
PCR steps awaits complete detection of genomic alterations of the RH locus. Therefore, direct physical analyses by fiber-FISH could be crucial for detecting genomic alterations, especially at regions containing copy number variations that are widely present in the human genome.

Acknowledgements

This work was partly supported by grants from the Japan Society for the Promotion of Science (JSPS), the Sumitomo Foundation, the Ministry of Health, Labour and Welfare of Japan, and the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References


