Identification of a Partial Internal Deletion in the RH Locus Causing the Human Erythrocyte D-- Phenotype

By Cheng-Han Huang, Marion E. Reid, and Ying Chen

The D-- phenotype of the human erythrocyte is a genetic variant of the Rh blood group system associated with the expression of D but not C, c, E, and e (designated non-D) antigens. In this report, we characterize the structure and expression of Rh polypeptide genes in two D-- homozygotes of Italian origin. Southern blot analysis detected a gross deletion in their genomic DNA that correlated with the alteration of CcEe rather than the D polypeptide gene. With detailed exon mapping, the deletion was found to be partial and internal, encompassing exons 2 through 8 of the non-D gene. Analysis of Rh cDNAs showed that no functional mRNA was produced from the truncated non-D gene, whereas the D gene gave rise to one major and two minor mature transcripts. The full-length RhD cDNA sequence contained four nucleotide changes resulting in four amino acid substitutions on the polypeptide backbone. The shortened RhD cDNAs occurred as alternatively spliced isoforms lacking sequences corresponding to exons 7 and/or 8. The identification of a partial and internal deletion in the non-D gene shows that the molecular basis for the D-- phenotype is heterogeneous and that its alterations have occurred on different genetic backgrounds.

© 1995 by The American Society of Hematology.

MATERIALS AND METHODS

Blood samples and nucleic acid isolation. Peripheral whole blood samples were obtained from normal human blood donors and two D-- homozygotes. Their Rh blood group types were determined by standard hemagglutination tests. The two D-- individuals were members from the same family of Italian origin. Both D-- individuals were homozygotes.

Nucleic acid primers and probes. Figure 1 shows the size and location of the cDNA and exon-specific probes used. Exon-specific probes were generated by PCR from cDNA probes and purified by native 5% polyacrylamide gel electrophoresis (PAGE). Genomic DNA probes were prepared from purified lambda phage Rh clones by PCR amplification. Oligonucleotide primers (24 to 27 mers) were synthesized on a 380 A automated DNA synthesizer (Applied Biosystems, Foster City, CA) and purified by 15% PAGE with 7.0 mol/L urea. The nucleotide sequences of amplification primers were either selected from published cDNAs or obtained during the course of this study. All cDNA probes and synthetic oligonucleotides were numbered in accord with their distance from the first nucleotide (nt) position of initiation codon ATG.

Genomic DNA restriction mapping analysis. Genomic DNAs...
MOLECULAR ANALYSIS OF A DELETED Rh VARIANT

Fig 1. Schematic representation of Rh polypeptide gene probes. The Rh cDNA extending from the 5' to 3' untranslated (UT) region is diagrammed and divided into 10 exons. The location of initiation (ATG) and termination (TAA) codons is shown. Three cDNA probes covering the 5', the middle (m), and 3' portions, respectively, are drawn in straight lines. Exon (Ex)-specific probes except Ex 1 are designated according to their nucleotide (nt) positions relevant to the first residue of initiation codon ATG: Ex 2, nt 154-303; Ex 3, nt 337-480; Ex 4-5, nt 493-1150; Ex 6-7, nt 802-1073; Ex 7-8, nt 943-1150; Ex 8-9, nt 1075-1124; and Ex 10, nt 1226-1358. Ex 1 and Ex 10 contain a portion of 5' UT and 3' UT, respectively. All probes cross-hybridize with both D and non-D Rh polypeptide genes.

from D--- and control individuals were digested with restriction endonucleases (New England BioLab, Beverly, MA) and analyzed by Southern blot hybridization as described. Probes were labeled with [α-32P]dCTP (New England Nuclear, Boston, MA) by the random primer extension method.

Amplification and sequencing of genomic DNA. Amplification of Rh gene sequences by PCR was performed with AmpliTaq, a thermostable DNA polymerase (Perkin Elmer, Norwalk, CT). One microgram of template DNA together with the 5' and 3' primers was amplified for 30 cycles in a thermocycler (Ericornp, San Diego, CA). Template DNA was initially denatured at 94°C for 2.5 minutes; cycles at 55°C for 1.5 minutes, at 72°C for 1.5 to 2.5 minutes, and at 94°C for 1 minute were repeated 29 times. The final steps for primer annealing (55°C) and chain extension (72°C) were for 3 and 10 minutes, respectively. For direct DNA sequence analysis, the PCR products were all purified by native 5% PAGE and eluted in 1 X TE buffer (10 mmol/L Tris-Cl, 0.1 mmol/L EDTA, pH 8.0). An aliquot of eluted DNA was directly sequenced on an automated 373A DNA sequencer (Applied Biosystems) with PCR cycling and fluorescent tags as chain terminators.

Synthesis, amplification, and sequencing of Rh cDNAs. Reverse transcription PCR (RT-PCR) of erythroid RNA was performed as previously described. Synthesis of the first strand of Rh cDNA was performed at 42°C for 60 to 90 minutes in 20 μL of reaction volume. The resultant cDNA products were directly added to 80 μL of premade PCR solution and then amplified for 35 cycles as

Fig 2. Genomic blot analysis of DNAs from D--- individuals. Genomic DNAs were digested with Sph I and HindIII, blotted onto filters, and hybridized with Rh cDNA 5', middle, and 3' probes (see Fig 1). Lanes 1 and 2 are controls from D-positive (DCCee) and D-negative (dccee) DNAs, respectively. Lanes 3 and 4 represent the two D--- homozygotes. The DNA size markers shown at the left margin are HindIII-cleaved λ DNA fragments.
described. Three sets of Rh gene-specific primers (a-b, c-d, and e-f) were used: a, 5'-ATGACCTTAAGTACCAGGCTGCTG-3' (nt 1-25), and b, 5'-TGGCCAGAAATCCACAAAGAG-3' (nt 640-663); c, 5'-ACGATACCCATGGTCTGCCGTG-3' (nt 607-630), and d, 5'-CCAGAAAACTTGGTCTACAAATATTT-3' (nt 1195-1221); and e, 5'-GCTAAATATTTTGATCGACCAAGTTTTC-3' (nt 1195-1221), and f, 5'-GTATTGCTACGTGCAATAAATGGGT-3' (nt 1432-1458). The sequences of primers a to e are specific for the D gene. The cDNA amplification products were directly sequenced as described above.

RESULTS

Identification of a gross deletion in the non-D gene. Restriction digestion of D-- genomic DNA with either HindIII or Sph I showed gross alterations in the Rh gene (Fig 2). It was evident that D-- genomic DNAs carried a relatively large deletion, as shown by hybridization with the middle and 3' probes. A comparison between the D negative and two D-- individuals showed an opposite pattern of DNA deletion, namely, some restriction fragments that occurred in the former were absent from the latter or vice versa. Because the D and non-D genes differ in the size of intron 4 and in the sequence of exon 7, amplification of the genomic region encompassing exons 4, 5, and 7 with gene-specific primers followed by diagnostic restriction digestions and DNA sequencing was performed. It was found that all the PCR products were from the D gene but not from the non-D gene (data not shown). Taken together, these results indicated that the deletion of the D-- complex is partial and resides in the non-D polypeptide gene.

Mapping of the limit of the deletion in the non-D gene. Having identified a gross deletion in the D-- complex, we then determined its limit by exon mapping. The Sph I and HindIII genomic blots were mapped in parallel using probes specific for exon 1 through exon 10 (Fig 1). Representative autoradiograms are shown in Fig 3. Among several restriction endonucleases tested, Sph I was found to be most telling because it only cleaves the noncoding region of Rh genes and yields more restriction fragments whose exonic content and gene origin can be readily defined (Figs 3 and 4).

Comparative deletion mapping of the D-negative and D-- variants showed that, whereas the entire D gene is missing in the former, the deletion of the non-D gene in the latter is indeed partial and internal. For both D-- individuals, the altered non-D gene apparently retained sequences for exons 1, 9, and 10 because these three exons each occurred in two separate bands as in D-positive controls (Fig 3, panels 1, 6, and 7). However, the altered non-D gene lacked the portion encompassing exons 2 through 7 because those bands from the ce gene of D-negative variants or the Ce gene of D-positive individuals were completely missing (Fig 3, panels 2 through 5). Regarding the exon 8 containing bands, D-- was similar to D-negative variants displaying a reduced in-
MOLECULAR ANALYSIS OF A DELETED RH VARIANT

Fig 4. A comprehensive Sph I map for the Rh polypeptide genes and their deletions. The distribution of exons in Rh D and Rh CeEe genes is shown for the DCe, dce, and D-- complexes. The D gene is presumed to reside upstream. Note that the organization of RhD is based on exon mappings and thus remains tentative. Below the gene, all exon-containing fragments flanked by Sph I (S) cleavage sites are denoted. The number of Sph I sites in some introns (introns 3, 7, and 9) is minimal and the size of the intergenic region not known. The deletions in dce and D-- are illustrated by open bars whose shaded ends denote the putative 5' and 3' breakpoint regions.

tensity as compared with the D-positive control (Fig 3, panels 5 and 6). These results indicated that the D-- complex harbors a deletion encompassing exons 2 through 8 of the non-D gene. Because the size of the exon 1 and exon 9 containing fragments found in the deleted gene was comparable to that of the normal Ce or ce gene, the 5' and 3' deletion endpoints could be localized to intron 1 and intron 8, respectively.

Figure 4 shows a comprehensive Sph I map for the DCe, dce, and D-- complexes. The assignment of 10 exons to the various restriction fragments of individual Rh genes was further confirmed by selected oligonucleotide mapping (data not shown). In the D-- complex, whereas the non-D gene is partially deleted, the D gene remains intact and apparently bears a similar, if not identical structural organization with the RHCE gene.26

Composition and splicing pattern of Rh transcripts in D-- erythroid cells. To investigate the expression of Rh genes in D-- erythroid cells, the composition and splicing pattern of Rh transcripts were analyzed by RT-PCR. Three overlapping cDNA segments that cover the entire coding sequence of Rh mRNAs were synthesized. Figure 5 shows the cDNA amplification products resolved by agarose gel electrophoresis. For segment a-b, one major band of expected size was apparent for D-- and D-positive individuals, whereas D-negative individuals produced two cDNA species of which the larger one carried an insertion sequence causing premature chain termination.25 For segment e-f that was primed with a D-specific sequence, D-- was also identical with the D-positive control showing a single cDNA band, whereas the D-negative variant gave no product because of lack of the sequence for primer f. The main difference between D-- and controls was noted in the segment c-d encompassing exons 4 through 9. The banding pattern in D-- was apparently less complex than that in controls.

To delineate the structure of Rh transcripts in D-- cells, all cDNA segments were eluted from gels and their nucleotide sequences were determined. The sequencing data showed that the three overlapping cDNA segments of expected size comprised a unique sequence from the D gene. The shortened c-d segments were also derived from the D gene, but they involved the exclusion of sequences corresponding to exons 7 and/or 8. Although the deleted gene still retained three exons and, probably, an intact promoter region (Fig 4), we failed to detect any mature forms of its transcript using several pairs of primers located in exons 1 and 10. Apparently the pre-mRNA of this gene, if any were transcribed, would be very unstable and thus rapidly degraded.

Predicted amino acid sequence of RhD polypeptide. Figure 6 shows the nucleotide sequence and the deduced amino acid sequence of the RhD polypeptide. Compared with the
reported D sequence, the RhD mRNA identified in D−− cells differ by four nucleotides that result in four amino acid changes at the following positions of the polypeptide backbone: 16 W−C, 174 V−M, 218 I−M, and 398 E−V. Of these alterations, the change of amino acid residues at positions 174 and 218 may be attributed to spontaneous nucleotide substitutions, whereas the replacement of the two other residues (16 W−C and 398 E−V) is also present at identical positions of the RhCe polypeptide. Further studies are needed to determine whether the latter nucleotide changes occur as simple DNA polymorphisms or result from templated transfers by the mechanism of gene conversion analogous to that occurring in the human MNSSs glycoporphins.

A hydropathy plot analysis of the RhD polypeptide from D−− individuals showed an arrangement of transmembrane segments identical with the previously predicted topologic model (figure not shown). This finding suggested that the introduced amino acid changes neither significantly alter the membrane disposition of the D polypeptide nor elicit variant antigenicities. Consistent with this notion, serologic typing showed an identical banding pattern; for brevity, only one lane is played an identical banding pattern. For brevity, only one lane is shown. The amplified cDNA products are designated by the primer used (see Materials and Methods for their sequences and function). The RT-PCR analysis of Rh transcripts expressed in D−− erythrocytes. Total RNA isolated from DCCee, dccee, and D−− erythrocytes (lanes 1 through 3) were analyzed by RT-PCR. (Top) Electrophoretic pattern of the resulting cDNA products on ethidium bromide staining. The two D−− individuals display an identical banding pattern; for brevity, only one lane is shown. The amplified cDNA products are designated by the two primers used (see Materials and Methods for their sequences and function). M is the Hae III fragments of ΦX174 RF DNA shown in basepairs. (Bottom) Schematic representation of the strategy for synthesis and amplification of Rh cDNAs by RT-PCR. The initiation and termination codons are indicated and the three overlapping cDNA segments illustrated. Note that all but primer f are common but primer f are common D and non-D. Primer f is a D-specific sequence located in 3′UT.

Fig 6. The nucleotide and deduced amino acid sequence of RhD from one D−− individual. The nucleotide sequence for the full-length RhD cDNA was derived from direct sequencing of the three overlapping cDNA segments (Fig 5) that are 663, 618, and 264 bp in size, respectively. Note that the RhD cDNA sequence described here differs from the reported D sequence by four nucleotides. These nucleotide replacements are predicted to result in four amino acid changes at the following positions of the polypeptide backbones: 16 W−C, 174 V−M, 218 I−M, and 398 E−V. show the occurrence of rare Rh antigens other than the D Rh antigens.

DISCUSSION

In this report we have characterized the structure and expression of a D−− complex in two D−− homoygotes of Italian origin. By Southern blot analysis and exon mapping, we have delineated a genomic alteration in the non-D gene that is a partial and internal deletion encompassing exons 2 through 8. With a detailed analysis of transcript expression, we detected mature transcripts expressed from the remaining D gene but not from the altered non-D gene. Based on these results, we conclude that the truncated non-D gene does not give rise to functional Rh polypeptides bearing C, c, E, and e antigens. Thus, the identification of the gene deletion explains the absence of CcEe antigens from the erythrocyte membrane.

Similar to other variant Rh phenotypes lacking the expression of non-D antigens, such as DCWE−, Dc−, and D−*, the occurrence of homoygous D−− is frequently associated
with consanguinity of which the genetic basis is still not completely understood. At the molecular level, two types of D— complex have been recognized. For the nondeletion type of D—, the gross structure of the RH locus and the sequence of D and non-D transcripts were apparently normal. Thus, the molecular basis for this type remains to be elucidated and subtle changes in the non-D gene, including splicing mutation, gene recombination, or promoter alteration may be responsible. Detailed studies of such nondeletion D— complexes should provide further insights into the regulation of gene expression and structure-phenotype relationships of the Rh polypeptides.

In the case of deletion type of D—, the absence of a large portion of the non-D gene has been described in an Iceland family. Southern blot analysis suggested that the deletion covered the entire gene except for exon 1 and the upstream sequence, but the 3' deletion limit and the structure of the D transcript were not determined. The deletion described here is apparently partial and internal and thus represents a new genetic form of D—. The identification of such distinct forms of Rh gene complexes shows that the molecular basis for the D— phenotype is heterogeneous and that the alterations have occurred on different genetic backgrounds. Elucidation of the mechanisms underlying these DNA deletions requires detailed molecular dissection of the corresponding breakpoint regions.

By taking advantage of the resolving power of SpH 1, we have mapped not only the deletion but also the remaining D gene in the D— complex of the two homozygotes. Comparative deletion mapping of the D negative and D— DNAs showed that the D gene bears a similar, if not identical, organization with the non-D gene encoding CeEe antigens. Using RT-PCR analysis, we showed that the remaining D gene is functionally expressed, giving rise to one major and two minor mature transcripts. The full-length RhD mRNA represents alternatively spliced D mRNA isoforms lacking sequences corresponding to exons 7 and/or 8. These D-specific isoforms appear to be authentic as they also have been found in human erythroleukemic K562 cells. Nevertheless, their functional significance remains elusive because no corresponding protein products have been shown to be expressed in the red blood cell membranes.

We show here, by parallel exon mapping, that the dce and D— complexes are associated with deletion in the D and non-D genes, respectively. This finding provides further evidence that the D and non-D (CeEe) antigens are encoded by two separate structural genes. Erythrocytes with these and other variant phenotypes (i.e., DC^W-, Dc-, and D-•-) do not exhibit morphologic or functional abnormality. This dispensability suggests a redundancy of Rh genes, regardless of whether the role of their products is structural or functional. Nevertheless, the Rh protein appears to be an essential membrane component, as indicated by the rare occurrence of Rh null erythrocytes. Because of the absence of all Rh antigen or protein products, these red blood cells manifest stomatocytosis and multiple membrane abnormalities. It is also noteworthy that the absence of the D antigen (Rh negative) is much more frequently encountered than the absence of non-D antigens in human populations. Whether this genetic diversity reflects a selective advantage driven at the population level remains an open question.

ACKNOWLEDGMENT

We thank Giorgio Reali (Milan, Italy) for supplying D— blood samples. We are grateful to Dr Olga Blumenfeld for critical reading of the manuscript. Thanks are also due to Bob Ratner for manuscript preparation and Tellervo Huima-Byron and Howard Wynder for photographs.

REFERENCES

Identification of a partial internal deletion in the RH locus causing the human erythrocyte D--phenotype [see comments]

CH Huang, ME Reid and Y Chen