Rh Gene Evolution in Primates: Study of Intron Sequences

Pol-André Apoil and Antoine Blancher
Laboratoire d’Immunogénétique Moléculaire, Université Paul Sabatier, Pavillon Charles Lefebvre, Hôpital Purpan, Toulouse, France

By amplification and sequencing of RH gene intron 4 of various primates we demonstrate that an Alu-Sx-like element has been inserted in the RH gene of the common ancestor of humans, apes, Old World monkeys, and New World monkeys. The study of mouse and lemur intron 4 sequences allowed us to precisely define the insertion point of the Alu-Sx element in intron 4 of the RH gene ancestor common to Anthropoidea. Like humans, chimpanzees and gorillas possess two types of RH intron 4, characterized by the presence (human RHCE and ape RHCLike-like genes) or absence (human RH and ape RHDLike-like genes) of the Alu-Sx element. This led us to conclude that in the RH common ancestor of humans, chimpanzees, and gorillas, a duplication of the common ancestor gene gave rise to two genes, one differing from the other by a 654-bp deletion encompassing an Alu-Sx element. Moreover, most of chimpanzees and some gorillas possess two types of RHDLike-like intron 4. The introns 4 of type 1 have a length similar to that of human RHDLike-like intron 4, whereas introns 4 of type 2 display an insertion of 12 bp. The latest insertion was not found in the human genome (72 individuals tested). The study of RH intron 3 length polymorphism confirmed that, like humans, chimpanzees and gorillas possess two types of intron 3, with the RHDLike-type intron 3 being 289 bases shorter than the RHCE intron 3. By amplification and sequencing of regions encompassing introns 3 and 4, we demonstrated that chimpanzee and gorilla RHLike-like genes displayed associations of introns 3 and 4 distinct to those found in man. Altogether, the results demonstrate that, as in humans, chimpanzee and gorilla RHLike genes experienced intergenic exchanges.

Introduction

The human RH blood group system, one of the most polymorphic in humans, is of utmost clinical importance because of the high immunogenicity of its antigens, particularly antigens D and c. Antibodies against Rh antigens are responsible for transfusion accidents and for most of fetomaternal incompatibilities (Mollison 1979). Ten years ago, it was demonstrated that humans possess two RH genes that are closely linked on chromosome 1: RHDLike, which encodes the D polypeptide and is present or absent depending on the RH haplotype (Colin et al. 1991), and RHCELike, which displays four common alleles responsible for the expression of the two allelic series of antigens C/c and E/e (Moouro et al. 1993). In a double dose, the complete deletion of RHDLike is responsible in Caucasians for the D-negative phenotype (Colin et al. 1991). In other populations (Africans, Japanese), D-negative individuals frequently possess in double-dose nonfunctional RHDLike genes which are wholly or partially deleted (Blunt, Daniels, and Carritt 1994; Hyland, Wolter, and Saul 1994; Daniels, Green, and Smart 1997; Okuda et al. 1997; Sun et al. 1998). RHCELike and RHDLike are highly homologous and most probably derived by duplication from a common ancestor gene (Colin et al. 1991; Le Van Kim et al. 1992). This homology between RHDLike and RHCELike probably promoted genetic exchanges between the two genes. This was demonstrated by the characterization at the genomic level of RHDLike genes which encode qualitative variants of the antigen D: most of these antigenic variants are encoded by low-frequency RHDLike alleles which display replacement of some exons by their RHCELike counterparts. These replacements can be due to various mechanisms of intergenic exchange (double crossing over or gene conversion) (for a general review, see Huang 1997a). Moreover, it was demonstrated that the four main alleles of the RHCELike gene derived through intergenic exchanges and interallelic recombinations from a few ancestor alleles (Carritt, Kemp, and Poulter 1997). Particularly, in humans, one gene exchange, consisting of the replacement of RHCELike exon 2 by its RHDLike counterpart, led to the appearance of RHCELike alleles coding for the C antigen (Carritt, Kemp, and Poulter 1997).

Nonhuman primates express counterparts of the human Rh antigens (Masouredis, Dupuy, and Elliot 1967; Moor-Jankowski and Wiener 1972). The use of monoclonal antibodies confirmed that chimpanzees and gorillas express polymorphic antigens, namely chimpanzee R c and gorilla D<sup>gor</sup>, which share epitopes with the human D antigen (Socha and Ruffié 1990; Blancher, Calvas, and Ruffié 1992). The expression of antigens R c and D<sup>gor</sup> was shown to depend on RHLike-like genes of chimpanzees and gorillas, respectively (Salvignol et al. 1993, 1994; Blancher and Socha 1997). Gibbons and orangutans also express D-like antigens, but the small number of animals studied did not allow the description of a polymorphism in these species (Blancher and Socha 1997). It has to be noted that if the expression of Rh-like antigens seems to be restricted to apes, the presence of Rh-like polypeptides at the surfaces of RBCs of Old World monkeys, New World monkeys, lemurs, and many other mammalian (cats, dogs, bovines, rats, mice) was evidenced by biochemical techniques (Saboroi, Denker, and Agre 1989) or by immunoblotting (Moouro et al. 1994; Salvignol et al. 1995; Apoil and Blancher 1999). Despite their conservation throughout evolution, the function of Rh polypeptides remains elusive.

Southern blot studies demonstrated that only humans and African apes (chimpanzees and gorillas) possess more than one gene per haploid genome (Salvignol
et al. 1993; Westhoff and Wylie 1994; Blancher and Socha 1997). Chimpanzees and gorillas possess at least three and two RH-like genes, respectively (Blancher and Socha 1997). Thus, it was deduced that the duplication event which led to the appearance of the RHCE and RHD genes arose most probably in the common ancestor species of humans, chimpanzees, and gorillas.

The two human RH genes differ in their coding sequences, but also in their noncoding elements, including length polymorphisms of intron 3 (Matassi et al. 1997) and intron 4 (Arce et al. 1993). Intron 4 of the RHCE gene differs from intron 4 of the RHD gene in the presence of a 654-bp fragment which encompasses an Alu-Sx element (Westhoff and Wylie 1996; Huang 1997b; Okuda et al. 1997). Like humans, gorilla possesses introns 3 and 4 length polymorphisms which are counterparts of the RHD/RHCE intronic polymorphisms observed in humans (Westhoff and Wylie 1996; Apoil, Roubinet, and Blancher 1999). For example, some gorillas RH-like genes exhibit RHCE-like introns 4 which harbor an Alu-Sx element orthologous to the Alu in human RHCE intron 4 (Apoi, Roubinet, and Blancher 1999). It was not possible from previous results to determine whether the Alu insertion in the RHCE gene arose after the duplication of a common ancestor gene or whether the RHD ancestor gene lost a 654-bp fragment. Results presented here are in good agreement with the deletion hypothesis.

Although chimpanzees and gorillas possess more than one gene per haploid genome (Blancher, Calvas, and Ruffé 1992), comparative studies of chimpanzee and gorilla RH-like cDNAs did not bring a definitive demonstration that genetic exchanges have occurred between RH-like genes in these species (Salvignol et al. 1995; Apoil and Blancher 1999). With the aim of demonstrating such intergenic exchanges, we decided to study noncoding parts of chimpanzee and gorilla RH-like genes.

Materials and Methods

Genomic DNA Samples

Human genomic DNA samples of 72 D-positive individuals (24 Caucasians, 24 African Blacks, 24 Americans) were provided by Jean Michel Dugooujon (Centre National de la Recherche Scientifique Unité propre de Recherche 8291, Toulouse, France), Chimpanzee (Pan troglodytes), gorilla (Gorilla gorilla), baboon (Papio papio), rhesus monkey (Macaca mulatta), and marmoset (Callithrix jacchus) blood samples were obtained from animals maintained at the Laboratory for Experimental Medicine and Surgery in Primates (LEMSIP, New York Medical Center, New York University). Squirrel monkey (Saimiri sciureus) and brown lemur (Eulemur fulvus) blood samples were obtained from the Centre de Primatologie de Strasbourg, Nierderhausergen, France. A blood sample from a Swiss mouse (Mus musculus) was drawn after euthanasia.

Amplification and Sequencing of Introns 4 of RH-like Genes

Primers complementary to exons 4 or 5 of RH-like genes were deduced from the Rh cDNA nucleotide sequences previously characterized (Mouro et al. 1994; Salvignol et al. 1995; Apoil and Blancher 1999). The sequences of primers were Ex.4-dir1 (CGATACCCAG-TTTGCTCTGACCAG), Ex.5-rev (TTGGGGTGACC-AAGGATGCAC(C/A)/C), Ex.4-dir2 (AGCTATTTGG-GCTGACTG), Ex.4.cons-dir (GCCTATTTTGGGCTGACTG), and Ex.5.cons-rev (GGCCGAATATCCC-ACAAAGAG).

The primer pairs Ex.4-dir1/Ex.5-rev, Ex.4-dir2/Ex.5-rev, and Ex.4.cons-dir/Ex.5.cons-rev were used for the amplification of RH-like introns 4 in humans and African apes, Old World monkeys and New World monkeys, and mice and lemurs, respectively. Thirty amplification cycles were carried out using an enzyme mix of Taq and Pwo DNA polymerases (Expand High Fidelity, Boehringer-Mannheim, Indianapolis, Ind.). Purified PCR products were sequenced by using the fluorescent dye terminator cycle sequencing method (PE Applied Biosystems, Foster City, Calif.). When required, amplified fragments were cloned into pCR 2.1.TOPO plasmid vector (TOPO TA Cloning kit, Invitrogen, Leek, the Netherlands) and sequenced.

Length Polymorphism of Primate RHD-like Introns 3 and 4

Primers Int.3-dir ((A/G)GGATTACAAGCAAGC-ATCACC) and Int.3-rev (CACGCAC(C/T)TCACT-GATTCTACTCTTC) were deduced from the intron 3 sequences of human RH genes (Matassi et al. 1997). This pair of primers led to the amplification of 580-bp (RHCE intron 3) or 290-bp (RHD intron 3) fragments with human genomic DNA. This set of primers was used to test for the presence of length polymorphism in the 3′ part of RH-like gene introns 3 in chimpanzees and gorillas. The pair of primers Int.4.cons-dir (CTCCCTCTTTACCAA(C/G)/TTC) and Int.4.cons-rev (AATCTGGA-ATACCCAGGC) allowed the amplification of a short segment (140 bp) of RHD-like introns 4 under the following conditions: 25 cycles of PCR (denaturation for 15 s, annealing for 20 s at 55°C, and extension for 10 s) were carried out using 1.25 U of Taq DNA polymerase (QIAGEN, Hilden, Germany) in a reaction medium supplemented with 20% of Q-solution additive (QIA-GEN).

Association of Introns 3 and 4 in Chimpanzee and Gorilla RH-like Genes

Primers RB46-dir (TGGCAAGAACCTGGACCT-ATCACC) and Int.3-rev (CACGCAC(C/T)TCACT-GATTCTACTCTTC) were deduced from the intron 3 sequences of human RH genes (Matassi et al. 1997) and Ex.5-rev were used to amplify DNA fragments encompassing intron 3 to the 5′ part of exon 5 from genomic DNA samples of one chimpanzee and two gorillas. These three animals were selected because they possessed introns 3 of the RHCE and RHD-like types and introns 4 of RHD-like type 1 and type 2. Amplicons recovered from agarose gel after electrophoresis were partially sequenced. Sequences of segments surrounding the intron 3 RHD-specific deletion were characterized, as were regions encompassing exon 4, intron 4, and a part of exon 5 (see fig. 1 for details).
FIG. 1.—Relative positions of primers used for PCR amplification. The black rectangles correspond to regions which are present in introns of the RHCE type and absent from introns of the RHD type. The primers used for the study of intron length polymorphisms are indicated (names of primers in italics). A pair of primers (RB.46-dir and Exon5-rev) was used to amplify long genomic fragments.

For one gorilla and one chimpanzee which possessed RHCE-like intron 3 with RHCE-like intron 4, we tested by PCR whether these two introns belonged to the same RH-like gene. For this purpose, we used RB46-dir and RHCE.int.4-rev (CCACCCTTGTTCCTTCACTCCTGG). The latter primer is specific for the RHCE intron (see fig. 1). Amplicons recovered from agarose gel after electrophoresis were partially sequenced.

Phylogenetic Analysis
Intron 4 nucleotide sequences were aligned using CLUSTAL W, version 1.7 (Thompson, Higgins, and Gibson 1994). Phylogenetic analysis was carried out with the MEGA software package (Kumar, Tamura, and Nei 1993). Rates of substitutions (K values) were calculated according to Kimura’s (1980) two-parameter method, and trees were reconstructed by the neighbor-joining method using the pairwise deletion option. One thousand resampled versions of the original data set were generated by bootstrap, and the 1,000 corresponding trees were deduced by the neighbor-joining method.

Results
Length of Intron 4 of Nonhuman Primate RH Genes

The lengths of the fragments amplified using DNA samples of various species were studied by agarose gel electrophoresis (fig. 2). From DNA samples of 105 chimpanzees and 15 gorillas, we amplified a fragment with a length (0.5 kb) similar to that of the human RHD intron 4. A second type of fragment of a length similar to that of the RHCE intron 4 (1.25 kb) was amplified from the genomic DNAs of only 17 out of 105 chimpanzees and from only 5 out of 15 gorilla genomic DNAs. With genomic DNA samples of one orangutan, two gibbons, two baboons, three macaques, and two marmosets, PCR amplification of RH-like introns 4 produced a single 1.25-kb fragment. The longest amplified products (0.3 kb longer than the human RHCE product) were obtained from PCR amplification of squirrel monkey DNA (four individuals). The length of the amplified products from two brown lemurs (0.75 kb) was between those of amplified RHCE and RHD fragments. The shortest fragment (0.35 kb) was amplified from a mouse DNA sample.

Sequences of RH Introns 4

Nucleotide sequences of RH-like introns 4 were aligned to their human counterparts (fig. 3). The exon segments 3’ and 5’ of introns 4 were homologous to the corresponding cDNA sequences previously reported (Mouro et al. 1994; Salvignol et al. 1995; Apoil and Blancher 1999).

An Alu-Sx element was present in the RHCE-like introns 4 (1.25 kb) of chimpanzees, gorillas, rhesus monkeys, squirrel monkeys, and marmosets. All of these Alu elements belong to the Alu-Sx subfamily, as determined by analysis with CENSOR software (Jurka et al. 1996). Comparing the nucleotides of the intron 4 Alu-Sx elements with those of a consensus human Alu-Sx sequence (Batzer et al. 1996) demonstrated the close relationship existing between the intron 4 Alu-Sx sequences: 12 characteristic positions are shared by human and nonhuman primate intron 4 Alu-Sx sequences (fig. 3). Intron 4 in the squirrel monkey possesses the Alu-Sx element and an Alu-Sc inserted in a reverse orientation 3’ to the Alu-Sx element (fig. 3). The Alu-Sc el-
Fig. 3.—Comparison of human RH and primate RH-like intron 4 sequences. Sequences are presented in reference to the RHCE gene intron 4, with a dot for nucleotide identity and a dash for nucleotide deletion. Nucleotide positions are numbered according to the human RHCE intron 4 (GenBank accession number AF050636; Apoil, Roubinet, and Blancher 1999). The upper part of the figure presents the alignment for the 5' and 3' segments of the introns, excluding the Alu insert. The lower part of the figure presents the variable positions of Alu-Sx elements found in RH intron 4 of primates by reference to a consensus human Alu-Sx sequence (Batzer et al. 1996), together with the polymorphic 5' and 3' flanking sequences. An Alu-Sx element is inserted into the 5' flanking region of the squirrel monkey intron. The possible transcription orientations of Alu elements are indicated by arrows. The 12mer duplication (GAGCAGGTTCAG) which is present in the RH D type 2 introns 4 of chimpanzees and gorillas is indicated with a black triangle. Sequences of gorilla RHCE-like (AF049913) and RH D-like type 1 (AF071197) introns 4 were previously reported (Apoi, Roubinet, and Blancher 1999).
Fig. 4.—Alignment of lemur and mouse RH-like introns 4 with the human RHCE intron 4 sequence, which is taken as a reference; nucleotide identities are indicated with dots, and nucleotide deletions are indicated with dashes. The sequences of lemur and mouse RH-like introns 4 are aligned with the human RHCE sequence (AF050636) after removal of the Alu-Sx element. The point of insertion of the Alu-Sx element in the RHCE sequence is indicated. The 354-bp segment which is deleted, together with the Alu in RHD intron 4, is indicated in bold type.

Phylogenetic Study of RH-like Introns 4

A phylogenetic study was performed on the basis of the alignment of intron 4 sequences. Only nucleotide
Fig. 5.—Phylogenetic tree computed from the sequences of RH gene introns 4. Distances were calculated using Kimura’s (1980) two-parameter method on the basis of the nucleotide positions shown in figure 2 excluding the Alu insert. One thousand resampled versions of the original data sets were generated by bootstrapping, and the corresponding trees were deduced by the neighbor-joining method with MEGA software (Kumar, Tamura, and Nei 1993). The frequencies of branching orders equivalent to those of the tree presented in the figure are indicated on the branches of the tree. * Bootstrap values <50%.

positions included in segments of introns 4 present in all primate sequences were taken into account, thus excluding the Alu insert region. Lemur and mouse intron 4 sequences were excluded from the analysis, because we estimated that the divergence of intron sequences between lemurs, mice, and other primates was too great to calculate a correct estimation of distances. A phylogenetic tree was reconstructed using the neighbor-joining method (fig. 5). The tree showed no deviations from the expected branching order, i.e., New World monkeys, Old World monkeys, and the “African” group, made up of humans, chimpanzees, and gorillas. However, genetic distances are shorter than expected between sequences of humans, chimpanzees, and gorillas. These irregularities are weakly significant because bootstrap analysis produced low levels of confidence for branchings inside the “African” group of sequences.

Presence of RHD-like Introns 4 of Type 1 and Type 2 in Chimpanzees and Gorillas

The presence of RHD-like introns 4 of types 1 and 2 was investigated by PCR using a pair of primers (Int.4.cons-dir and Int.4.cons-rev) which amplified a segment centered on the 12mer repeat present in the type 2 RHD-like intron 4 (see Materials and Methods for more details). Among 105 chimpanzees, 102 possessed the two types of RHD-like intron 4, and 3 possessed only the RHD-like type 1 intron. Among 15 gorillas, 7 possessed the two types of RHD-like intron 4, and 8 possessed only type 1 RHD-like intron 4 (table 1).

The PCR amplification centered on the 12mer repeat demonstrated that all 72 human DNA samples of various ethnic origins and of the D-positive type (24 Caucasians, 24 Africans, and 24 Amerindians) possessed counterparts of chimpanzee and gorilla RHD-like type 1 introns 4 and lacked counterparts of RHD-like type 2 introns 4.

Polymorphism of Intron 3 in Chimpanzees and Gorillas

As human RHD and RHCE genes differ in the length of intron 3, we searched for a similar polymorphism in chimpanzees and gorillas with a PCR assay (see Materials and Methods). From all chimpanzee DNA samples (N = 105), products of lengths equivalent to those of human RHCE and RHD, respectively, were obtained. RHD-like and RHCE-like intron 3 fragments were obtained with 7 of 15 gorilla DNA samples. With the remaining eight gorilla samples, only RHCE-like fragments were obtained (table 1). Partial sequences of the chimpanzee and gorilla PCR products confirmed that they were amplified from RH-like genes (data not shown).

Types of Associations Between Introns 3 and Introns 4 and Estimated Frequencies of RH Haplotypes in Chimpanzees

In order to investigate the various combinations of introns 3 and 4 in chimpanzee and gorilla RH-like genes, amplifications of long genomic fragments were performed. Fragments were separated on the basis of their lengths and were checked by sequencing for the presence of RHD-like specific deletions in introns 3 and 4 and for the presence of the RHD-like type 2 addition in intron 4 (see Materials and Methods and fig. 1). From one chimpanzee and one gorilla, DNA fragments of 2 kb containing RHCE-like introns 3 and 4 were amplified by means of a primer in intron 3 (RB46-dir) and a primer specific to the RHCE intron 4 (RHCE.Int.4-rev). The corresponding RH-like genes were named “Chimp CE/CE” and “Gor CE/CE” (fig. 6). Partial sequencing of chimpanzee and gorilla amplicons confirmed the presence of RHCE-specific elements in intron 3 and intron 4.

Table 1

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>TYPE OF RH INTRON 3</th>
<th>TYPE OF RH INTRON 4</th>
<th>NO. OF INDIVIDUALS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CE</td>
<td>D</td>
<td>CE</td>
</tr>
<tr>
<td>Chimpanzees (105 animals)</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Gorillas (15 animals)</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Note.—The phenotypes for gorilla (D<sup>⁎⁎</sup>) are indicated. For a general review on the D<sup>⁎⁎</sup> and R-C-E-F systems, see Blancher and Socha (1997).
Genomic DNAs of four animals (two chimpanzees and two gorillas) which possessed both types of intron 3 and RH-like intron 4 of types 1 and 2 were amplified from intron 3 to exon 5 (see fig. 1). Two types of fragments were obtained of respective lengths 1.95 and 2.25 kb. Partial sequencing of the chimpanzee and gorilla 1.95-kb fragments demonstrated that they contained an RH-like intron 3 and an RH-like intron 4 of type 1 ("Gor D/D1," "Chimp D/D1"; fig. 6). Partial sequencing of the 2.25-kb fragment demonstrated that it contained RHCE-like intron 3 and an RH-like intron 4 of type 2 ("Gor CE/D2," "Chimp CE/D2"; fig. 6).

All attempts to amplify genomic fragments associating RHCE-like introns 3 with RH-like type 1 introns 4 in chimpanzees and gorillas failed. Therefore, the existence of a gorilla RH-like gene associating RHCE-like introns 3 with type 1 RH-like introns 4 is questionable. It is possible, therefore, that these two introns do not belong to the same gene (the h1 haplotypes of gorillas and chimpanzees in fig. 6).

As 105 randomly selected chimpanzees (table 1) were studied for the presence of introns 3 and 4, one can propose the existence of at least three RH haplotypes and estimate their frequencies. The presence of three RH-like genes per chimpanzee haploid genome was previously demonstrated (Blancher, Calvas, and Ruffié 1992); therefore, only three-gene haplotypes are proposed in this paragraph. Three chimpanzees possessed a phenotype suggesting that they were homozygotes for haplotype h1 (frequency = p) which associates RH-like genes [D/D1−CE/?−?/?] (frequency = p), defined by intron 3/4 combinations (question marks indicate the impossibility of assessing the type of intron). According to the Hardy-Weinberg binomial formula, the frequency of haplotype 1 is p². In the same way, 85 animals are expected to be either homozygotes of haplotype h2 (frequency = q) [D/D1−CE/D2−?/?] or heterozygotes of haplotypes h1 and h2, and the calculated frequency is q² + 2pq. Finally, 17 chimpanzees are considered either homozygotes of haplotype h3 (frequency = r) [D/D1−?/D2−CE/CE] or heterozygotes of haplotypes h2 and h3, with a frequency of r² + 2pr + 2qr. The approximate estimations for p, q, and r are 0.17, 0.75, and 0.08, respectively, and are reported in figure 6. Obviously, many other haplotype combinations can be proposed, and only the complete sequencing of the chimpanzee RH loci of numerous animals could lead to a definite description of chimpanzee haplotypes. In addition, the respective positions of RH-like genes in chimpanzee and gorilla genomes, as shown in figure 6, are arbitrary.

The 15 gorillas studied here displayed four types of combination between the three types of intron 4 and the two types of intron 3 (table 1). One of these combinations consisted in the presence of RHCE-like intron 3 and RH-like intron 4 of type 1 (six gorilla samples). From the existence of this combination, one can infer that in some gorilla genes, RHCE-like intron 3 is associated with type 1 RH-like intron 4. However, the existence of such a gorilla RH-like gene is questionable, because, as mentioned above, all attempts to amplify genomic fragments associating RHCE-like introns 3 with type 1 RH-like introns 4 in gorillas failed.

**Discussion**

We demonstrate here that an Alu-Sx element is inserted in orthologous positions in RH-like gene introns 4 of humans, great apes, New World monkeys, and Old World monkeys. These results led us to conclude that an Alu-Sx element was most probably inserted into the
RH gene of the common ancestor of Old World monkeys and New World monkeys after the separation of this ancestor from the lemur lineage. In humans, this Alu-Sx element is present only in RHCE intron 4. The human RHD gene differs from the RHCE gene by the absence of this Alu-Sx element and of 354 bp 5’ of the Alu. This 354-bp DNA segment is homologous to a region of the lemur RH-like intron 4 sequence. This homology led us to propose that the insertion of the Alu-Sx in the RH ancestor gene was 3’ of this region (see fig. 4). The 5’ and 3’ Alu flanking sequences displayed great variability in length between species (lower part of fig. 3). In marmosets and squirrel monkeys, the 5’ Alu-Sx flanking regions contain repeats of 10 (marmoset) to 28 (squirrel monkey) thymidines which are replaced in humans and apes by a 12mer poly-A sequence. The corresponding zone in the rhesus monkey consists of a mixture of A, T, and G. Moreover, the squirrel monkey intron 4 displays a second Alu element which belongs to the Alu-Sc subfamily, inserted in reverse orientation in the 5’ poly-T flanking sequence of the Alu-Sx. The insertion of the Alu-Sc occurred after the insertion of the Alu-Sx element because the expansion of the Alu-Sx subfamily in primate genomes preceded that of the Alu-Sc (Kapitonov and Jurka 1996).

Some chimpanzee and gorilla RH-like genes were shown to possess RHCE-like introns 4 encompassing an Alu-Sx element in a position orthologous to that observed in the human RHCE gene. However, whereas all humans, apart from exceptional variants (Huang 1997a), possess the RHCE gene, and thus introns 4 of the RHCE type, only 16% of the chimpanzees (17/105) and 33% of the gorillas (5/15) studied here possessed an RHCE-like intron 4.

Like humans, chimpanzees and gorillas also possess RHD-like introns 4. Both human RHD and ape RHD-like introns 4 are deprived of a 654-bp segment which is present in RHCE and RHCE-like introns 4. In addition, some chimpanzee and gorilla RHD-like introns 4 exhibit an additional 12mer repetitive DNA segment 3’ of the deleted region. Introns without the repeat were referred to as type 1 RHD-like, while RHD-like introns 4 with the 12mer segment were designated type 2 RHD-like. The chimpanzee type 2 RHD-like intron 4 is homologous to the chimpanzee sequence previously reported by Westhoff and Wylie (1996).

In humans, the absence of RHD is frequent (15% of Caucasians are D-negative and do not possess RHD introns 4 in their genomes). In African apes, absence of RHD-like introns 4, if possible, must be infrequent, as all chimpanzees and all gorillas tested so far have possessed type 2 and/or type 1 RHD-like introns 4. However, it has to be noted that in the series reported by Westhoff and Wylie (1996), one gorilla out of five and no chimpanzees out of seven were negative for the amplification of RHD-like introns 4.

The great homology between human RHD intron 4 and RHD-like (type 1 and type 2) introns 4 of gorillas and chimpanzees suggests a common ancestry of the corresponding genes. One can hypothesize that the common ancestor gene of RHD and RHD-like genes already displayed the deletion encompassing the Alu-Sx, together with 354 bp 5’ of the Alu. Although excision of Alu elements is a rare event, examples have previously been described by others (Meagher, Jorgensen, and Deeb 1996; Satta, Mayer, and Klein 1996). It is noteworthy that only chimpanzees and gorillas possess type 2 RHD-like intron 4, which differs from type 1 RHD-like intron 4 by the presence of a 12mer repeat (fig. 3). Indeed, using a pair of primers adapted to the detection of RHD-like intron 4 polymorphisms (presence or absence of the 12mer repeat), we demonstrated by PCR the absence of a counterpart of the apes’ type 2 RHD intron 4 in 72 human genomic samples. Taken together, the results suggest that the common ancestor of chimpanzees and gorillas, which is supposed to also be the ancestor of humans, possessed the type 2 RHD intron 4. This led us to conclude that humans, unlike chimpanzees and gorillas, most probably lost the gene(s) harboring type 2 RHD intron 4. To verify this hypothesis, it remains to identify the relicts of this deletion in the human genome and to study a larger number of individuals, because the RHD type 2 intron could be present in humans at very low frequency and perhaps only in some peculiar populations.

As reported elsewhere, gorillas express a D-like polymorphic antigen called D°°r (Roubinet et al. 1993). The expression of the D°°r antigen is associated with restriction length polymorphisms revealed by probing Southern blots with an exon–4–specific probe and with a PCR length polymorphism of the RHD-like intron 3 (Apol, Roubinet, and Blancher 1999). In the present study, we demonstrate that the expression of D°°r is also associated with the presence in the gorilla genome of the type 2 RHD-like intron 4 and, therefore, with the presence of a haplotype associating the “Gor D/D1” and “Gor CE/D2” genes (fig. 6 and table 1). However, it is not possible to ascribe one of these two genes to the expression of D°°r. Gorilla RH-like genes associated with type 1 RHD intron 4 are also functional because D°°r-negative animals are agglutinated by some anti-D reagents and express Rh-like proteins at the surfaces of their red blood cells (RBCs) (Blancher and Socha 1997; Apol, Roubinet, and Blancher 1999). Although it is not possible to assess whether or not the gorilla Gor CE/CE gene is functional, partial exon sequences of this gene established its homology with the human RchE cDNA (Apol, Roubinet, and Blancher 1999). In chimpanzees, the absence of association between the intron 3 and 4 PCR patterns and the R-C-E-F types or the restriction patterns (data not shown) did not allow us to specify the functionality of the RH-like genes described here. However, all of the animals studied here were agglutinated by some anti-D monoclonal reagents and express Rh-like proteins at the surfaces of their RBCs.

Chimpanzees and gorillas displayed a great variety of combinations between the various types of RH-like introns 3 and 4. This variety is indirect evidence that intergenic exchanges between RH-like genes in these two species were not infrequent (Blancher and Socha 1997). However, if these intergenic exchanges were frequent, they would have homogenized RH sequences in
each species. This would result in a clustering of sequences by species. In fact, this is not observed in the phylogenetic tree presented in figure 5.

In conclusion, the numerical chromosomal polymorphism of RH genes (the human locus displays either one or two genes, the chimpanzee possesses three, and the gorilla possesses two) suggests that unequal crossing over most probably arose after the original duplication of the RH ancestor gene. Intergenic exchanges by various mechanisms (double crossing over, homologous recombination, or gene conversion) between RH genes have occurred in humans and the two other species possessing more than one RH gene (i.e., chimpanzees and gorillas). However, despite these intergenic exchanges, the coding regions of human RHD and RHCE genes (417 codons) differ by 41 nucleotide substitutions, with 35 being nonsynonymous. This suggests that the differentiation of the two human RH genes (RHD and RHCE) was maintained because it represented a selective advantage.

Acknowledgments

This work was realized with funds from MESR (Contract Jeune Equipe 1966) and from Agence Francaise du Sang (Contract 65001731231). We are indebted to Stephanie Despiau for her very efficient technical assistance in sequencing. All experiments described in this article were performed in agreement with French laws and regulations currently in force. We thank Doctor Naury Saitou for helpful discussion and critical review. The RH and RH-like gene intron 4 nucleotide sequences reported in this paper have been submitted to the EMBL/GenBank nucleotide sequence databases and have been assigned the following accession numbers: chimpanzee RHCE-like, RH-like type 1, and RH-like type 2—AF045542, AF071199, and AF071198, respectively; gorilla RH-like type 2—AF071200; rhesus monkey—AF045541; squirrel monkey—AF081795; marmoset—AF045540; brown lemur—AF045544; mouse—AF045543.

LITERATURE CITED


Naruya Saitou, reviewing editor

Accepted September 27, 1999