

Evolution of two Rh blood group-related genes of the amphioxus species *Branchiostoma floridae*

Takashi Kitano^{1*}, Masahiro Satou¹ and Naruya Saitou²

¹*Department of Biomolecular Functional Engineering, College of Engineering, Ibaraki University, 4-12-1 Nakanarusawa-cho, Hitachi 316-8511, Japan*

²*Division of Population Genetics, National Institute of Genetics, 1111 Yata, Mishima 411-8540, Japan*

(Received 7 September 2009, accepted 31 March 2010)

We determined cDNAs of two genes that belong to the Rhesus (Rh) blood group gene family in an amphioxus species (*Branchiostoma floridae*) and designated them Rh-related-1 (RhR-1) and Rh-related-2 (RhR-2). RhR-1 and RhR-2 consisted of 10 and 11 exons, respectively. 3' UTR sequences of RhR-1 were shorter (220–272 bp) than those of RhR-2 (1,505–1,650 bp). CDS lengths were 1,344 and 1,476 bp for RhR-1 and RhR-2, respectively, and the average nucleotide difference between their CDS regions was 0.33. The corresponding regions of Rh genes from exons 2 to 7 were relatively conserved among the chordate species examined in this study. Length difference numbers were in multiples of three, which implies that codon frames were conserved among them, and the same exon/intron boundary phases were observed in those regions. This region was used for the phylogenetic analyses. RhR-1 and RhR-2 formed a cluster on the phylogenetic tree of the Rh gene family. Gene duplication time of RhR-1 and RhR-2 was estimated to be ca. 500 million years ago. It is likely that the four Rh family genes in vertebrates emerged by gene duplications in the common ancestor of vertebrates, and functional differentiation has occurred after the first gene duplication.

Key words: amphioxus, blood, evolution, phylogeny, Rh

INTRODUCTION

The human Rhesus (Rh) blood group plays important roles in transfusion and clinical medicine, as it has been known to be involved in hemolytic diseases of newborns, autoimmune diseases, and mild hemolytic anemia. Human Rh blood group (also known as Rh30) genes encode membrane proteins with multiple transmembrane domains (Avent et al., 1990, 1992) and are expressed only on erythrocytes (Cherif-Zahar et al., 1990). Three other genes that are homologous to Rh blood group genes have been identified in the human genome. These are, RhAG (Rh-associated glycoprotein, also known as Rh50), RhBG (Rh family, B glycoprotein), and RhCG (Rh family, C glycoprotein). RhAG is an erythrocyte-specific protein and is known to interact with Rh. RhBG and RhCG proteins are non-erythroid members of the Rh blood group gene family and are primarily expressed in the kidney. In humans, the chromosomal locations are 1p36.11 for Rh, 6p21.1-p11 for RhAG, 1q21.3 for RhBG, and 15q25 for RhCG. Kitano and Saitou (2000) suggested that gene

duplication that produced Rh and RhAG occurred in the common ancestor of vertebrates. Moreover, they pointed out the possibility of two more duplications for Rh blood group genes in jawed vertebrates. Later, all three duplications that produced these four genes were also suggested to have occurred in the common ancestor of vertebrates (Huang and Peng, 2005; Peng and Huang, 2006).

Vertebrata, Urochordata, and Cephalochordata belong to Chordata, and Urochordata is the closest neighbor of Vertebrata (Putnam et al., 2008). The ascidian *Ciona intestinalis* is frequently used as a representative of Urochordata; its genome sequence data are available. However, it was demonstrated that ascidians have long branches on a deuterostome phylogeny constructed from 1,090 orthologous genes, indicating higher amino acid substitution levels (Putnam et al., 2008), which may pose difficulties for phylogenetic analyses. On the other hand, Cephalochordata (amphioxus) is the second closest neighbor of Vertebrata, and it does not have a long branch.

In order to include Rh-related (RhR) genes of amphioxus for phylogenetic analysis of the Rh gene family, it is necessary to obtain clear characteristics of the genes. Data

Edited by Chung-I. Wu

* Corresponding author. E-mail: tkitano@mx.ibaraki.ac.jp

of the four genome sequences (ABEP01017616, ABEP01003689, ABEP01002794, and ABEP01062874) of RhR genes in amphioxus (*Branchiostoma floridae*), which were obtained from the whole genome shotgun sequence project (Putnam et al., 2008), were deposited in the DDBJ/EMBL/GenBank International Nucleotide Sequence Database, and some EST sequences of amphioxus, which contain a part of the RhR gene, were available. To elucidate exon and intron boundaries and coding regions of RhR genes in amphioxus, we sequenced complete cDNA sequences of six clones of RhR genes in amphioxus and conducted an evolutionary analysis of the Rh gene family.

MATERIALS AND METHODS

cDNA sequencing Six clones (bfad001a05, bfad037h16, bfad047k21, bfad049a17, bflv024h10, and bflv041d06) of Rh blood group-related genes in amphioxus (*Branchiostoma floridae*) were kindly provided by the Academia DNA Sequencing Center, National Institute of Genetics in collaboration with the Department of Zoology, Graduate School of Science, Kyoto University (Yu et al., 2007). Sequencing were performed using a BigDye Terminator v3.1 Cycle Sequencing Kit and an ABI PRISM 3130xl

Genetic Analyzer (PE Biosystems). Both strands were read using sequencing primers. Primers used for RhR-1 were as follows: bfRhR1-1, 5'-CTGGTGGCTGCCTTCGT-GCT-3'; bfRhR1-2, 5'-CGCCGCCATCTGGAATTCTG-3'; bfRhR1-3, 5'-GTTTGTCCACCTTGGAGTC-3'; bfRhR1-4, 5'-AGAAACGTCATCAGGAACC-3', and for RhR-2 were as follows: bfRhR2-1, 5'-CAACGAGTGGGTCGGACTCA-3'; bfRhR2-2, 5'-ATTAGGCAGATTCCTAATAC-3'; bfRhR2-3, 5'-GTGGACATGAACGGCAGTC-3'; bfRhR2-4, 5'-TTGTC TGTATCTAGCCTCC-3'; bfRhR2-5, 5'-CCCAAATAGG-GAAACTTGGT-3'; bfRhR2-6, 5'-CGAGTACGGAGACGA-TAAG-3'; bfRhR2-7, 5'-ATATGATCATAACATTTAACGC-3'; bfRhR2-8, 5'-TGTGCCCCACCATGTCCTC-3'; bfRhR2-9, 5'-CACTCTGCGTCCTGAATGAC-3'. The Phred/Phrap software program (Ewing et al., 1998) was used for base-calling and assembly and for obtaining quality scores for assembled data. Editing was performed using Consed (Gordon et al., 1998) to identify all low-quality bases and to check that the assembly was correct based on linking information.

Sequence analyses To compare sequences with those of other members of the Rh gene family, human (*Homo sapiens*) and torafugu (*Takifugu rubripes*) cDNA

Table 1. List of genes and accession numbers of Rh family genes used in this study

Species (Scientific binomen)	Gene	Accession number	
		mRNA	Genome
Human (<i>Homo sapiens</i>)	RhD	NM_016124	NC_000001
	RhCE	NM_020485	NC_000001
	RhAG	NM_000324	NC_000006
	RhBG	NM_020407	NC_000001
	RhCG	NM_016321	NC_000015
Torafugu (<i>Takifugu rubripes</i>)	Rh	NM_001032763	CAAB01002106
	RhAG	NM_001037867	CAAB01003585
	RhBG	AB218980	CAAB01001243
	RhCG-1	NM_001032647	CAAB01000968
	RhCG-2	NM_001032762	CAAB01003371
	RhP2-1	NM_001032644	CAAB01000115
	RhP2-2	NM_001032645	CAAB01000285
Ascidian (<i>Ciona intestinalis</i>)	Rh-A	NM_001032776	AABS01000115
	Rh-B	NM_001032787	AABS01000034
	Rh-C	NM_001032774	AABS01000044
Amphioxus (<i>Branchiostoma floridae</i>)	RhR-1	AB519682 ^a	ABEP01017616 ^a
	RhR-2	AB519685 ^a	ABEP01002794 ^a
Purple sea urchin (<i>Strongylocentrotus purpuratus</i>)	LOC584796	XM_784645	NW_001467604

^aThey are representatives. AB519683 and AB519684 for the mRNA of RhR-1, AB519686 and AB519687 for the mRNA of RhR-2, ABEP01003689 for the genome of RhR-1, and ABEP01062874 for the genome of RhR-2 were also used in the study.

sequences were used as representatives of vertebrates. Three genes of ascidian (*Ciona intestinalis*) were also used. The genes and GenBank accession numbers of the mRNAs and the genomes are listed in Table 1. Exon and intron boundaries of RhR genes in amphioxus were predicted by comparing human, torafugu, and ascidian sequences. T-Coffee version 5.72 (Notredame et al., 2000) was used for multiple alignments that were made exon by exon, followed by concatenating these alignments. Amino acid sequences of exons 2–7 (exon numbers followed human RhD) were used for the following phylogenetic inferences (see RESULTS AND DISCUSSION).

Repetitive elements were detected by the RepeatMasker program (Smit et al., <http://www.repeatmasker.org>) using the cross_match search engine with slow speed/sensitivity

options, and the DNA source of *Branchiostoma floridae* species.

Phylogenetic inferences We estimated phylogenetic trees using the neighbor-joining (Saitou and Nei, 1987), the maximum likelihood (Felsenstein, 1981), and the Bayesian (Huelsenbeck et al., 2001) tree-building methods. The sequence data (XM_784645) of purple sea urchin (*Strongylocentrotus purpuratus*) was used as an outgroup. For the neighbor-joining and the maximum likelihood trees, the JTT model (Jones et al., 1992) with 1,000 bootstrap replicates was used, and a gamma shape parameter was estimated for the maximum likelihood method. MEGA version 4 (Tamura et al., 2007) and PHYML version 2.4.4 (Guindon and Gascuel, 2003) were used to construct the neighbor-joining and the maximum

Table 2. Exons and introns and their boundaries in amphioxus RhR genes

Exon	boundary 5 ^{aa}	Exon bp	boundary 3 ^{aa}	Intron bp ^b	Phase
1) RhR-1					
1		199 ^c	...CCWCTgtaag...	450 (468)	2
2	...tgcagCMTAG...	164	...CCCGAgtacg...	5,201 (1,248)	1
3	...cacagTGTTTC...	184	...AGCAGgtatt...	555 (700)	2
4	...yccagCMTGC...	151	...TCAAGgtwws...	1,235 (748)	0
5	...gacagGCSCA...	145	...GATAGgtgag...	761 (873)	1
6	...cmcagGTACG...	164	...ACATGgtaag...	511 (468)	0
7	...accagGTTCA...	138	...TTCAGgtacg...	1,321 (1,942)	0
8	...ctcagCCGTW...	128	...GAAAgtaaa...	465 (741)	2
9	...ttcagGTCCG...	157	...GGGAGgtgag...	885 (1,569)	0
10	...cgcagGTTCC...	362 ^c			
2) RhR-2					
1		186 ^c	...TTGAGgtacg...	930 (920)	0
2	...cmcagGTGGT...	136	...CCCGAgtgag...	9,661 (8,584)	1
3	...ytcagTGTTTC...	184	...ACTACgtaag...	1,201 (709)	2
4	...tgyagYCTGC...	151	...TCAAGgtcag...	411 (424)	0
5	...cacagGCCGT...	145	...GATCGgtaag...	1,849 (1,996)	1
6	...tacagGTACTION...	164	...ACATGgtaag...	715 (668)	0
7	...cacagGTTCA...	138	...TCACGgtaag...	502 (506)	0
8	...tgcagCCCTT...	134	...AACAGgtggg...	527 (785)	2
9	...cacagYGTGT...	182	...CACTGgtaag...	351 (434)	1
10	...tstagGCCTC...	74	...GGGAGgtgag...	665 (472)	0
11	...cycagATACC...	1,775 ^c			

^aEach of the five nucleotides of exon and intron are shown; and polymorphic sites among five sequences (three cDNAs and two genome sequences) are shown by abbreviations of the Nomenclature Committee of the International Union of Biochemistry.

^bLengths of introns are from ABEP01017616 for RhR-1 and ABEP01002794 for RhR-2; and lengths of introns of ABEP01003689 for RhR-1 and ABEP01062874 for RhR-2 are in parentheses.

^cFor lengths of the first and last exons, the longest ones among the three cDNAs are used.

likelihood trees, respectively. For the Bayesian method, the WAG model (Whelan and Goldman, 2001) with 300,000 generations was used. MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003) was used to construct the Bayesian tree.

RESULTS AND DISCUSSION

Structures of amphioxus RhR genes We determined six mRNAs of RhR genes in amphioxus. We also predicted the codon start position of each sequence, as only one methionine codon was observed in each putative first exon of each mRNA. The DDBJ/EMBL/GenBank International Nucleotide Sequence Database accession numbers are AB519682–AB519687. These six mRNAs were divided into two types of Rh blood group-related genes. Thus, the amphioxus species probably has two Rh blood group-related loci, and we designated them: RhR-1 and RhR-2. CDS lengths were 1,344 and 1,476 bp for RhR-1 and RhR-2, respectively. The average nucleotide difference (p-distance) in the CDS region between RhR-1 and RhR-2 was 0.33. A large difference was observed in the lengths of 3' UTRs between RhR-1 and RhR-2. 3' UTR sequences of RhR-1 were shorter (220–272 bp) than those of RhR-2 (1,505–1,650 bp).

Because the whole genome shotgun sequences of the amphioxus genome were available, we predicted exons, introns, and their boundaries in the two amphioxus RhR genes (Table 2) by comparing mRNA and genome sequences. We compared two genome sequences for each

RhR gene (ABEP01017616 and ABEP01003689 for RhR-1 and ABEP01002794 and ABEP01062874 for RhR-2). RhR-1 and RhR-2 consisted of 10 and 11 exons, respectively. These are the typical exon numbers for Rh family genes. All exon/intron boundaries were consistent with the GT/AG rule. The exon lengths were identical among the three mRNAs and in the two genome sequences for each gene, whereas the intron lengths were different among the two genome sequences. In particular, intron 2 of RhR-1 had larger differences: one was 5,201 bp and the other was 1,248 bp (Table 2). This difference was due to the insertion of 4,399 bp-long EnSpm-N4_BF type DNA transposon.

Phylogenetic relationships of amphioxus RhR genes To construct a more reliable phylogenetic tree of Rh family genes, we considered the following three points. First, to eliminate noise from phylogenetic analyses, we constructed a simpler phylogenetic tree and used only human and torafugu genes as representatives of vertebrates (Table 1). Because of tandem gene duplication, there are two Rh loci (RhD and RhCE) in humans (Blancher and Socha, 1997). In torafugu, there are two RhCG loci, probably because of a gene duplication event in the fish lineage. In addition, torafugu has two additional Rh-related genes (Rhp2-1 and Rhp2-2). Second, we used data of amino acid sequences from exons 2 to 7 (Table 3), as these regions were relatively conserved. The genes used here had the same exon/intron boundary phases at the 3' ends as exons 1–7, except exon 7 of purple

Table 3. Exons and phases of Rh family genes

Species	Gene	Exon ^a										
		1	2	3	4	5	6	7	8	9	10	11
Human	RhD	148 (1)	187 (2)	151 (0)	148 (1)	167 (0)	138 (0)	134 (2)	80 (2)	74 (0)	27	
	RhCE	148 (1)	187 (2)	151 (0)	148 (1)	167 (0)	138 (0)	134 (2)	80 (2)	74 (0)	27	
	RhAG	157 (1)	184 (2)	151 (0)	148 (1)	167 (0)	138 (0)	122 (2)	71 (1)	74 (0)	18	
	RhBG	187 (1)	187 (2)	151 (0)	148 (1)	167 (0)	138 (0)	134 (2)	122 (1)	75 (1)	17	
	RhCG	184 (1)	187 (2)	151 (0)	148 (1)	167 (0)	138 (0)	137 (2)	125 (1)	74 (0)	129	
Torafugu	Rh	139 (1)	187 (2)	151 (0)	148 (1)	173 (0)	138 (0)	128 (2)	80 (1)	74 (0)	27	
	RhAG	193 (1)	184 (2)	151 (0)	148 (1)	167 (0)	138 (0)	122 (2)	74 (1)	74 (0)	75	
	RhBG	178 (1)	187 (2)	151 (0)	148 (1)	167 (0)	138 (0)	134 (2)	122 (1)	74 (0)	78	
	RhCG-1	223 (1)	193 (2)	151 (0)	148 (1)	167 (0)	138 (0)	137 (2)	125 (1)	74 (0)	126	
	RhCG-2	205 (1)	193 (2)	151 (0)	148 (1)	167 (0)	138 (0)	137 (2)	125 (1)	74 (0)	74 (2)	34
Ascidian	Rh-A	148 (1)	193 (2)	151 (0)	148 (1)	167 (0)	138 (0)	137 (2)	94 (0)	138		
	Rh-B	220 (1)	193 (2)	139 (0)	169 (1)	164 (0)	138 (0)	134 (2)	202			
	Rh-C	5 (2)	149 (1)	193 (2)	139 (0)	175 (1)	164 (0)	138 (0)	134 (2)	187 (0)	67	
Amphioxus	RhR-1	23 (2)	164 (1)	184 (2)	151 (0)	145 (1)	164 (0)	138 (0)	128 (2)	157 (0)	90	
	RhR-2	48 (0)	136 (1)	184 (2)	151 (0)	145 (1)	164 (0)	138 (0)	134 (2)	182 (1)	74 (0)	120
Purple sea urchin	LOC584796	154 (1)	193 (2)	151 (0)	145 (1)	164 (0)	138 (0)	127 (1)	77 (0)	72		

^aExon numbers are followed by the human RhD. Phases are shown in parentheses.

sea urchin. Moreover, length difference numbers in exons 2–7 were in multiples of three. For example, on exon 2, human RhD is of 187 bp and human RhAG is of 184 bp; thus, the difference is of 3 bp. Therefore, we can assume that there are gaps that do not affect codon frames in those regions. Exon 1 was not used because it had variable base pair lengths, suggesting a variety of translation start sites. Phases and base-pair lengths were not conserved at exon 8 and later exons. It is likely that many changes occurred in the codon frames of that region. Thus, the data of amino acid sequences from exons 2 to 7 were used to make multiple alignments of Rh family genes. Third, in order to focus attention on the

relationship among Rh family genes of Chordata (Vertebrata, Urochordata, and Cephalochordata), we used genes only from these organisms. A gene from purple sea urchin, which belongs to Echinodermata, was used as an outgroup, and other invertebrate species were not used for phylogenetic analyses.

Figure 1A shows the phylogenetic tree of chordate Rh family genes. Two RhP2 genes of torafugu formed a cluster with each other. However, the phylogenetic position of RhP2 genes was unclear with low bootstrap and posterior probabilities. RhP2-1 of torafugu is an intronless gene, and RhP2-2 of torafugu has only one intron. Their homologous genes have been observed in

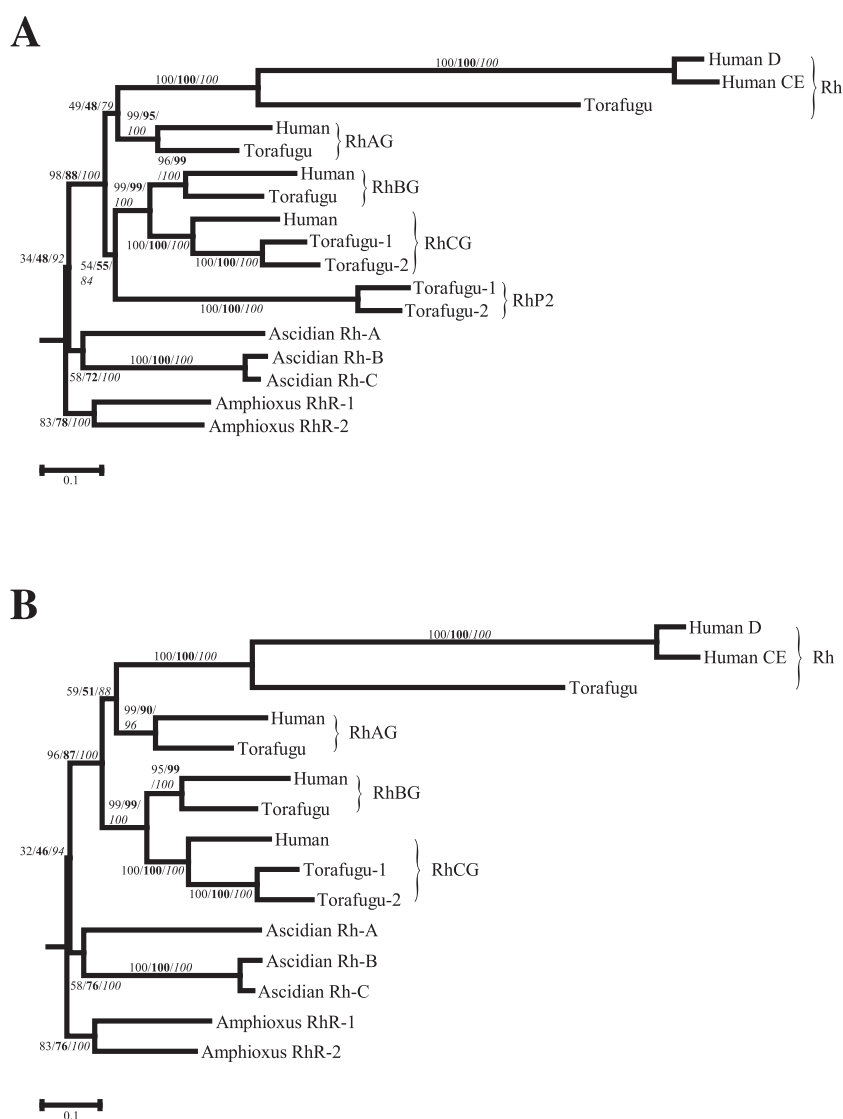


Fig. 1. Phylogenetic trees of chordate Rh family genes, including (A) and excluding (B) two RhP2 genes of torafugu. The root was determined using purple sea urchin as an outgroup. These phylogenetic trees were constructed by the neighbor-joining method. The scale is 0.1 substitutions per site. The maximum likelihood and the Bayesian methods constructed the same topologies. Numbers on each branch show bootstrap values calculated by the neighbor-joining method (normal font), by the maximum likelihood method (bold), and posterior probabilities by the Bayesian method (italic).

other fishes, such as spotted green pufferfish (*Tetraodon nigroviridis*; intronless, DQ013062) and zebrafish (*Danio rerio*; one intron, NM_131547), and in western clawed frog (*Xenopus tropicalis*; intronless, NM_001045792), and in platypus (*Ornithorhynchus anatinus*; two introns, XM_001507513). RhP2 gene had probably emerged by insertion of spliced mRNA by reverse transcription. Because the phylogenetic position of RhP2 genes was unclear, two RhP2 genes of torafugu were excluded from further analysis.

Figure 1B shows the phylogenetic tree of chordate Rh family genes, excluding two RhP2 genes of torafugu. It is important to emphasize that the same topologies were constructed by the three tree-building methods. Rh and RhAG genes, which have erythrocyte specific expression, formed a cluster, although branch supporting probabilities were low (bootstrap value of 59% by the neighbor-joining method, bootstrap value of 51% by the maximum likelihood method, and posterior probability of 88% by the Bayesian method), and RhBG and RhCG genes formed a cluster with high probabilities (bootstrap value of 99% by the neighbor-joining method, bootstrap value of 99% by the maximum likelihood method, and posterior probability of 100% by the Bayesian method). Rh blood group genes had longer branches than other Rh related genes in the phylogenetic tree. This pattern is consistent with previous studies (Kitano et al., 1998; Kitano and Saitou, 1999; Huang and Peng, 2005). In ascidian, three genes (Rh types A, B, and C glycoproteins) formed a cluster. They probably arose by two gene duplication events (Huang and Peng, 2005), independent from those in vertebrates. In amphioxus, two genes (RhR-1 and RhR-2) formed a cluster arisen by a gene duplication. This gene duplication was probably independent of those in vertebrate and ascidian lineages.

Emergence time of each Rh family gene To estimate duplication times between RhR-1 and RhR-2 genes in amphioxus, and among RhAG, RhBG, and RhCG genes in vertebrates, we reconstructed a linearized tree (Fig. 2) assuming equal evolutionary rates in all lineages (Takezaki et al., 2004). Three Rh genes (human D and CE, and torafugu) were excluded from the tree, as they have higher evolutionary rates and longer branches, as indicated above. The divergence time (T) between human and torafugu was assumed to be 450 million years ago (MYA) (Kumar and Hedges, 1998). Evolutionary rates (λ) for RhAG, RhBG, and RhCG were estimated to be 3.6×10^{-10} , 3.4×10^{-10} , and 3.8×10^{-10} , respectively. Because these three values were similar, the mean λ (3.6×10^{-10}) was used as the rate of amino acid change of Rh family genes. Thus, gene duplication time of the two amphioxus genes (RhR-1 and RhR-2) was estimated to be ca. 500 MYA [$= d/\lambda = 0.1852/3.6 \times 10^{-10}$]. Whether or not these duplicated genes have functional differentiation

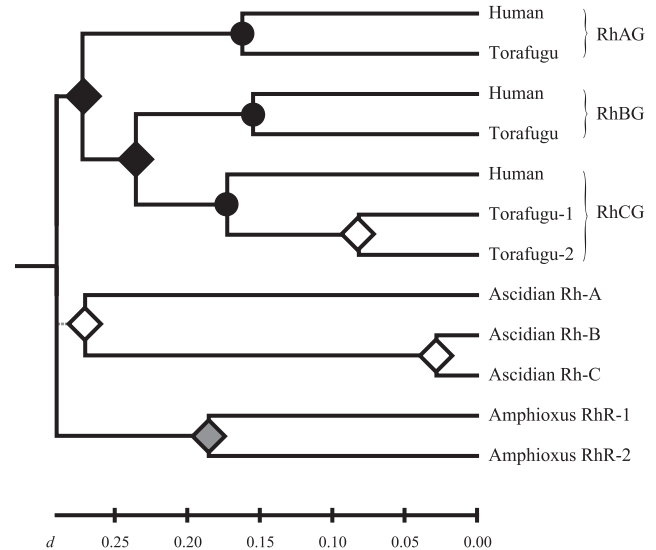


Fig. 2. A linearized tree of chordate Rh family genes. The root was determined using purple sea urchin as an outgroup. Divergence nodes between human and torafugu are shown by filled circles. Two gene duplications in the common ancestor of vertebrates and the gene duplication in the amphioxus lineage are shown by filled and gray diamonds, respectively. Other gene duplications are shown by open diamonds. Because the common ancestral branch of three ascidian genes had no amino acid changes in the linearized tree, this branch is shown by a broken gray line.

requires further investigation, as these two Rh family genes have existed for quite a long time in amphioxus.

Two gene duplication times for RhBG and RhCG and for RhAG and RhBG-RhCG in the vertebrate lineage were estimated to be ca. 650 MYA and ca. 750 MYA, respectively. This period roughly corresponds to two rounds of genome duplications in the common ancestor of vertebrates. However, clear synteny around these four Rh family genes have not been observed. Thus, further investigation is required in order to clarify whether or not these four Rh family genes were duplicated by two rounds of genome duplications in the common ancestor of vertebrates. In either case, because Rh and RhAG genes have erythrocyte-specific expression and RhBG and RhCG proteins are non-erythroid members of the Rh gene family, it is reasonable to suppose that these functional differentiations occurred after the first gene duplication.

We thank Dr. Yuji Kohara of the Academia DNA Sequencing Center, National Institute of Genetics and Dr. Noriyuki Satoh of the Department of Zoology, Graduate School of Science, Kyoto University for providing us cDNA clones of amphioxus (*Branchiostoma floridae*). We also thank two anonymous reviewers for their helpful and valuable suggestions. This study was supported by a Grant-in-Aid for Scientific Research (18770002) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan to TK.

REFERENCES

- Avent, N. D., Butcher, S. K., Liu, W., Mawby, W. J., Mallinson, G., Parsons, S. F., Anstee, D. J., and Tanner, M. J. (1992) Localization of the C termini of the Rh (rhesus) polypeptides to the cytoplasmic face of the human erythrocyte membrane. *J. Biol. Chem.* **267**, 15134–15139.
- Avent, N. D., Ridgwell, K., Tanner, M. J., and Anstee, D. J. (1990) cDNA cloning of a 30 kDa erythrocyte membrane protein associated with Rh (Rhesus)-blood-group-antigen expression. *Biochem. J.* **271**, 821–825.
- Blancher, A., and Socha, W. W. (1997) The Rhesus system. In: *Molecular biology and evolution of blood group and MHC antigens in primates* (eds.: A. Blancher, J. Klein, and W. W. Socha), pp. 147–218. Springer-Verlag, Berlin, Heidelberg, New York.
- Cherif-Zahar, B., Bloy, C., Le Van Kim, C., Blanchard, D., Bailly, P., Hermand, P., Salmon, C., Cartron, J. P., and Colin, Y. (1990) Molecular cloning and protein structure of a human blood group Rh polypeptide. *Proc. Natl. Acad. Sci. USA* **87**, 6243–6247.
- Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185.
- Felsenstein, J. (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376.
- Gordon, D., Abajian, C., and Green, P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res.* **8**, 195–202.
- Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704.
- Huang, C. H., and Peng, J. (2005) Evolutionary conservation and diversification of Rh family genes and proteins. *Proc. Natl. Acad. Sci. USA* **102**, 15512–15517.
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R., and Bollback, J. P. (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**, 2310–2314.
- Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282.
- Kitano, T., and Saitou, N. (1999) Evolution of Rh blood group genes have experienced gene conversions and positive selection. *J. Mol. Evol.* **49**, 615–626.
- Kitano, T., and Saitou, N. (2000) Evolutionary history of the Rh blood group-related genes in vertebrates. *Immunogenetics* **51**, 856–862.
- Kitano, T., Sumiyama, K., Shiroishi, T., and Saitou, N. (1998) Conserved evolution of the Rh50 gene compared to its homologous Rh blood group gene. *Biochem. Biophys. Res. Commun.* **249**, 78–85.
- Kumar, S., and Hedges, S. B. (1998) A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920.
- Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217.
- Peng, J., and Huang, C. H. (2006) Rh proteins vs Amt proteins: an organismal and phylogenetic perspective on CO₂ and NH₃ gas channels. *Transfus. Clin. Biol.* **13**, 85–94.
- Putnam, N. H., Butts, T., Ferrier, D. E., Furlong, R. F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J. K., et al. (2008) The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**, 1064–1071.
- Ronquist, F., and Huelsenbeck, J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.
- Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.
- Smit, A. F. A., Hubley, R., and Green, P. (1996–2004) RepeatMasker Open-3.0. <http://www.repeatmasker.org>.
- Takezaki, N., Rzhetsky, A., and Nei, M. (1995) Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.* **12**, 823–833.
- Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599.
- Yu, J. K., Satou, Y., Holland, N. D., Shin-I, T., Kohara, Y., Satoh, N., Bronner-Fraser, M., and Holland, L. Z. (2007) Axial patterning in cephalochordates and the evolution of the organizer. *Nature* **445**, 613–617.
- Whelan, S., and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699.