Transfusion and Apheresis Science 44 (2011) 81–91

Contents lists available at ScienceDirect

Transfusion and Apheresis Science

journal homepage: www.elsevier.com/locate/transci

Molecular genetics and clinical applications for *RH*

Willy A. Flegel

Department of Transfusion Medicine, Clinical Center, National Institutes of Health, Bethesda, MD 20892, United States

ARTICLE INFO

Keywords:
Rhesus
Blood group
Molecular genetics
Molecular diagnostic
Transfusion

ABSTRACT

Rhesus is the clinically most important protein-based blood group system. It represents the largest number of antigens and the most complex genetics of the 30 known blood group systems. The *RHD* and *RHCE* genes are strongly homologous. Some genetic complexity is explained by their close chromosomal proximity and unusual orientation, with their tail ends facing each other. The antigens are expressed by the RhD and the RhCE proteins. Rhesus exemplifies the correlation of genotype and phenotype, facilitating the understanding of general genetic mechanisms. For clinical purposes, genetic diagnostics of Rhesus antigens will improve the cost-effective development of transfusion medicine.

Published by Elsevier Ltd.

1. Introduction

Molecular biology has been applied extensively in characterizing the genetic basis of blood group systems and developing clinical diagnostic tools for immunohematology and transfusion medicine [1–4]. There are now 51 antigens within the Rh system and more than 200 alleles for the *RHD* gene alone. *RHD* zygosity has been resolved, epitopes have been mapped, and many D variants with altered antigens have been identified. The relationship among the *RH* family members in various species contributes to our understanding of their biological importance [5].

Based on the homology of Rh polypeptides to the ammonia transporter AmpB, computational analyses have modeled the 3D structure of the RhD polypeptide to learn about additional potential functions of Rh polypeptides [6]. The reason for this interest is that *RHAG*, a gene located on chromosome 6 (6p11–p21), shares an identical exon structure and major regions of sequence identity with *RHD*/*RHCE*. Moreover, RhAG is essential for the expression of the Rh polypeptides and was identified in 2008 as the latest blood group system, No. 30, in its own right [7].

To date, the function of RhD and RhCE appears associated with membrane integrity, and possibly transport of gases like carbon dioxide. On the other hand, the Rh-associated glycoprotein (RhAG) can transport ammonia [8], but whether it does so in red blood cells (RBCs) is debated. Also, RhAG may contribute to gas exchange across the plasma membrane, and its mutations are associated with hereditary stomatocytosis [9]. Thus, expression of Rh polypeptides and associated proteins is complex, and molecular discoveries have broadened our understanding of this important blood group system. This review summarizes the progress that molecular analyses have made in furthering clinical applications for Rh.

Three clinically useful discoveries have been made since the cloning of *RHCE* and *RHD*: 1) the molecular basis underlying the common Rh-negative haplotype and the nucleotide polymorphisms associated with the common Rh antigens have been applied to predict risk for hemolytic disease of the fetus and newborn (HDFN); 2) the molecular distinction of partial D and weak D alleles, *DEL*, *RHD*-pseudogenes, and the *RHD*-deleted genome; and 3) the molecular basis of D antigen epitope expression on the RhCE polypeptide.

Recently efforts by a few independent research laboratories have begun to more fully characterize the molecular basis of RhCE variants and the allelic variation of *RHCE*. The results are applied to improve transfusion support for sickle cell disease (SCD, see a separate review in this issue) and to identify the deficiencies of monoclonal antisera in assigning antigen status accurately. It is now apparent that

E-mail address: bill.flegel@nih.gov

1473-0502/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.transci.2010.12.013

molecular analyses are the most accurate way to define the complex *RH* and other blood group systems. In a steadily increasing number of clinical settings, these molecular approaches facilitate preventing blood group incompatibilities, avoiding alloimmunizations and hemolytic transfusion reactions, and contributing to optimal RBC survival in transfusion-dependent immune disorders.

2. Molecular basis of *RH*

RH is a bigenic locus comprising *RHD* and *RHCE* positioned in a tail-to-tail orientation toward the end of the short arm of chromosome 1 (p34–36). Another gene, *SMP1*, is interspersed between both *RH* genes in close proximity to the 3′ end of *RHCE* (Fig. 1) [10]; this minute technical feature was instrumental in resolving the physical structure of the *RH* gene locus [11]. Identification of the single murine equivalent in the mouse genome project provided evidence that *RHCE* evolved from the ancestral *RH* on the basis of the position and orientation of murine genes in the region (Fig. 1). Therefore, *RHD* arose from a duplication event that predates modern humans [12]. During the duplication event, and possibly associated with its cause, two approximately 9000 base-pair-long homologous repeat sequences, termed *Rhesus boxes*, were likely introduced that flank the *RHD* gene in the genomes of modern humans. *RHD* was lost from the genome through unequal crossing over involving the *upstream Rhesus box* and *downstream Rhesus box* (Fig. 2), an event that may have occurred more than once. *RHD* and *RHCE* share sufficient sequence homology that RBCs function normally when no *RHD* gene is inherited. Why the *RHD*-deleted genomes have persisted to this day and become more prevalent is the topic of much worthwhile debate and some esoteric speculation [13,14].

More than 200 *RHD* alleles have been reported and may be grouped according to serological and molecular features (Table 1). Most of the alleles harbor either single nucleotide polymorphisms (SNPs) or present as *RHD*/*RHCE* hybrid alleles. The tail-to-tail orientation (Fig. 3) may facilitate the large number of alleles; the identification of corresponding nucleotides in both genes suggests that most hybrid alleles arise through gene conversion events. A clinical benefit of the molecular characterization of the *RH* locus is that *RHD* zygosity can be assigned with near certainty. In the past, Rh haplotype tables based on serological analyses were used to predict the 'most probable genotype' and thus *RHD* zygosity for Caucasians and African Americans.

3. Modeling of Rh polypeptides

The Rh proteins belong to the ammonium transport (*E. coli* AmtB)/methylammonium permease superfamily. Initially, molecular modeling of both Rh proteins and the RhAG protein on the crystal structure of *E. coli* AmpB suggested that the Rh proteins and RhAG form trimers in the red cell membrane, and may therefore function as ammonium transport proteins [15]. A few independent investigators have shown that RhAG facilitates ammonium transport in yeast and *Xenopus* oocytes without co-transfection of RhD and RhCE polypeptides [8]. Similar ammonium transport activity was shown directly in erythrocytes [16]. However, RhD and RhCE do not appear to transport ammonium or carbon dioxide; key amino acid substitutions in the transmembrane channel do not appear to meet the requirements for facilitating transport [17].

The model of red cell Rh proteins is consistent with the 6 exofacial loops that were first proposed on the basis of computational hydropathy plots [18]. A critical component of the structure of RhD and RhCE is the amino acids that

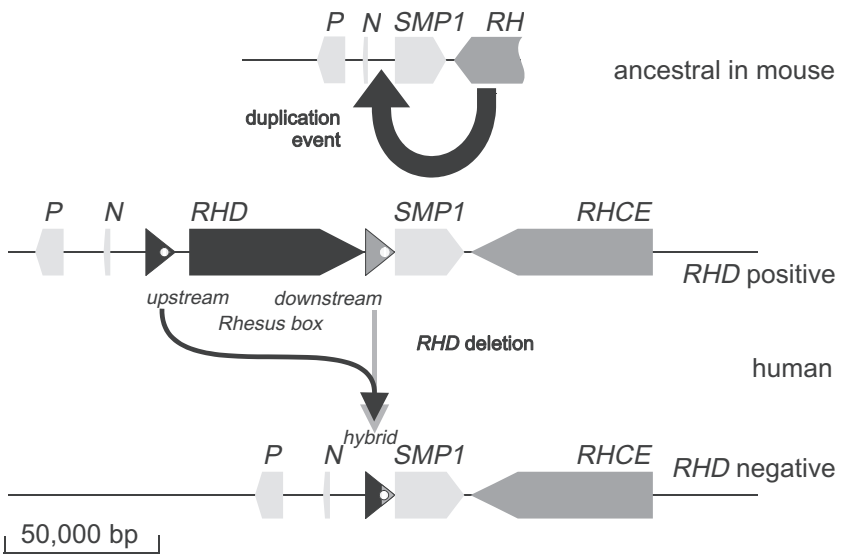

Fig. 1. Duplication of the *RH* gene and loss of the *RHD* gene. The ancestral configuration is shown as represented by the *RH* gene locus in mouse. The single *RH* gene is in close proximity to the three genes *SMP1*, *P29-associated protein* (*P*), and *NPD014* (*N*). A duplication event introduced a second *RH* gene in reverse orientation between *N* and *SMP1*. At the two break points in front and behind the *RHD* gene, DNA segments of approximately 9000 base pairs (bp) occur. Both DNA segments are flanking the *RHD* gene and dubbed "*upstream Rhesus box*" and "*downstream Rhesus box*". In the *RHD* positive haplotype, the *RHD* gene may have been lost by a recombination event (see Fig. 3).