Rhesus Gene Family and Blood Group Antigens—Molecular Aspects in Relation to Genomic Transfusion Medicine

Cheng-Han Huang, MD, PhD

Head, Laboratory of Biochemistry and Molecular Genetics, Lindsley F Kimball Research Institute, New York Blood Center

Abstract

The Rhesus (Rh) antigens form a blood group system of major significance in transfusion medicine due to their polymorphic nature at a population level and their potency as protein immunogens, which in response to incompatibility induce harmful hemolytic reactions. For seven decades, the Rh family has undergone extensive investigation and has served as a model for studies on membrane biology with a focus on biochemistry and genetics. The past decade has seen a rapid growth of molecular data on Rh allelic diversity and a major effort in probing the function of Rh proteins as gas channels in the membrane. It is now established that the antigen carrier RhD or RhCE and its regulator RhAG, the erythroid branches of the Rh family, dictate antigen expression, which together with RhBG and RhCG, the epithelial branches of the Rh family, penetrate all vertebrate animals and arise from a common ancestor of unicellular origin. Hematology and immunohematology, similar to other clinical disciplines, are on the horizon of genomic medicine, being transformed by the new knowledge of chromosome biology and gene expression. This article addresses the molecular aspects of the Rh protein family with an emphasis on its blood group system, relating translation research to genomic transfusion medicine.

Keywords

Rh family, erythrocyte, epithelia, gas channel, membrane transport, blood group system, gene evolution, genetic disease

Disclosure: The author's work was supported by National Institutes of Health (NIH) grants HL54459, HD62704, and HL66274 and funds from the New York Blood Center.
Acknowledgments: The author thanks Dr Olga Blumenfeld for helpful comments and Dr Mao Ye for figure illustrations.
Received: December 10, 2008 **Accepted:** June 9, 2009 *DOI: 10.17925/OHR.2009.02.0.52*
Correspondence: Cheng-Han Huang, MD, PhD, Laboratory of Biochemistry and Molecular Genetics, Lindsley F Kimball Research Institute, New York Blood Center, New York, NY 10021. E: chuang@nybloodcenter.org

The human red cell surface is coated with a plethora of reactive substances known as blood group antigens, which form 30 systems and serve compatibility roles analogous to human leukocyte antigens (HLAs). Of this vast antigen repertoire, ABO and Rhesus (Rh) embody two of the most medically significant systems as a cornerstone of immunohematology in transfusion medicine.[1–4] In contrast to ABO antigens, which are expressed in the form of carbohydrates in multiple tissues,[5] Rh antigens are membrane proteins and are erythroid-specific[6–8] with two intrinsic biological features underlying their clinical importance. They are highly polymorphic at a population level (particularly the D blood group antigen) and are most potent in inducing harmful immune reactions in counteracting conflict situations, such as maternal–fetal incompatibility.[3] In fact, the Rh-negative phenotype, a red cell hallmark of absence of the D antigen, is known as one of the most prevalent human genetic variations.[9]

In the 1990s, cloning of Rh30 (RhD and RhCE) and RhA glycoprotein (RhAG) opened a new chapter of the erythroid Rh proteins.[10–14] Subsequently, these genes have become essential tools in redefining the Rh antigen system and in correlating genotypes with phenotypes. Around 2000, the discovery of RhBG and RhCG as the epithelial branches of the Rh family[15,16] facilitated a broad interest in pursuing the nature of their physiological roles in various systems. Most notably, the past decade has seen a rapid growth of molecular data on Rh allelic diversity and a major effort in probing the function of Rh proteins as membrane gas channels.[17–22] With the new knowledge of blood group antigens, immunohematology as a subspecialty of hematology, similar to other clinical disciplines, is developing within genomic medicine built on chromosome biology and gene expression. This articles summarizes the molecular aspects of the Rh protein family with an emphasis on its blood group system in relating translational research to genomic transfusion medicine.

Origin, Distribution, and Duplication of Rhesus Genes

Although first found in human red cells, the Rh family has ancient roots and probably arose from a common ancestor of prokaryotic origin.[22–24] In general, the Rh gene is scarce in prokaryotes, but exhibits a wider distribution in unicellular eukaryotes and becomes ubiquitous in multicellular organisms from primitive metazoans to invertebrates to vertebrates and mammals. Depending on the species, the Rh family varies in member composition and copy numbers. Usually, Rh-carrying

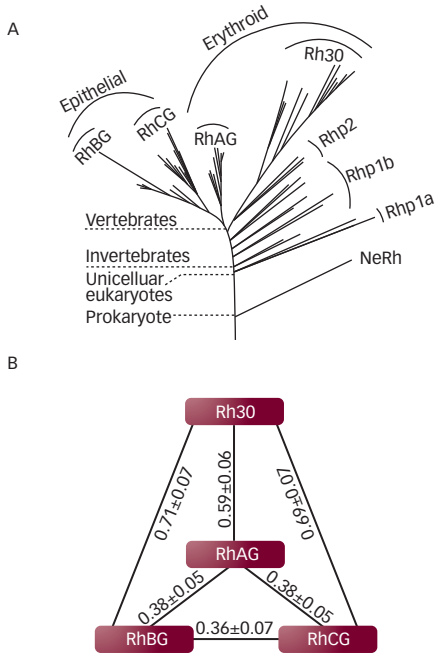
prokaryotes contain a single-copy gene, and Rh-carrying unicellular eukaryotes have one or two but seldom more than two copies, except for *Trichomonas vaginalis*. Such variations are also seen in metazoan animals in the form of either gene gain or gene loss, but the trend is an increase of copy number from lower to higher organisms.[22–24] On one hand, the water flea *Daphnia pulex* has as many as six genes; on the other hand, insects, such as fruit flies and honey bees, have only one gene. The Rh gene copy number is much less varied in vertebrates, particularly mammals.

The phylogenetic tree of the Rh family reveals three features that define the relationship of its members in metazoan animals[22–24] (see *Figure 1*).[1] Division of subgroups is clearly and only seen in vertebrates, and consists of two branches of erythroid specificity (Rh30 and RhAG) and two branches of epithelial specificity (RhBG and RhCG).[2] A novel gene is found in lower vertebrates (fish to frog) but not in mammals. Its members form a separate subgroup (Rhp2) above the invertebrate class (Rhp1), thus being the direct ancestor of the four universal vertebrate subgroups.[3] Although invertebrate homologs do not cluster into subgroups due to a greater intraspecific similarity, they individually are comparable in distance from the vertebrate subgroups. These features suggest that invertebrate Rh proteins retain the ancestral function, whereas vertebrate Rh subgroup proteins underscore a major drive of subfunctionalization coincident with vertebrate evolution.

The Rh family tree also reveals the degree of conservation of vertebrate subgroups, i.e. the epithelial branches are more conserved than the erythroid branches (RhBG>RhCG>RhAG>Rh30) (see *Figure 1*). So, RhBG and RhCG could have arisen before RhAG and Rh30, as *Trichoplax adhaerens*, a basal metazoan, has two epithelial cell types and two Rh genes,[25,26] but no hemocyte, a primitive form of vertebrate blood cells. Then, a RhAG-like gene derived from an epithelial ancestor could have emerged first in invertebrate hemocyte followed by another duplication making Rh30 in vertebrate red cells. As only one Rh gene is found in insects, it may be expressed in both epithelia and hemocytes. As in humans[27] (see *Figure 2*), epithelial and erythroid Rh genes exist as independent chromosomal loci in vertebrates. The only known exception is the duplication of RhD from RhCE in recent evolution that resulted in two genes in tandem and eight common haplotypes among human populations.[28] Thus, all the precursor genes of RhBG, RhCG, RhAG, and Rh30 must have evolved early on in accordance with vertebrate evolution and primarily via chromosomal translocation. Further insight into this view can be obtained after the identification of Rh subgroups in jawless vertebrate hagfish and lamprey.

Members of the Rh family share a common membrane fold typical of 12 spanning segments as seen in human erythroid and epithelial Rh proteins.

Figure 1: The Rhesus Protein Family Consists of Four Subgroups Universal to Vertebrates



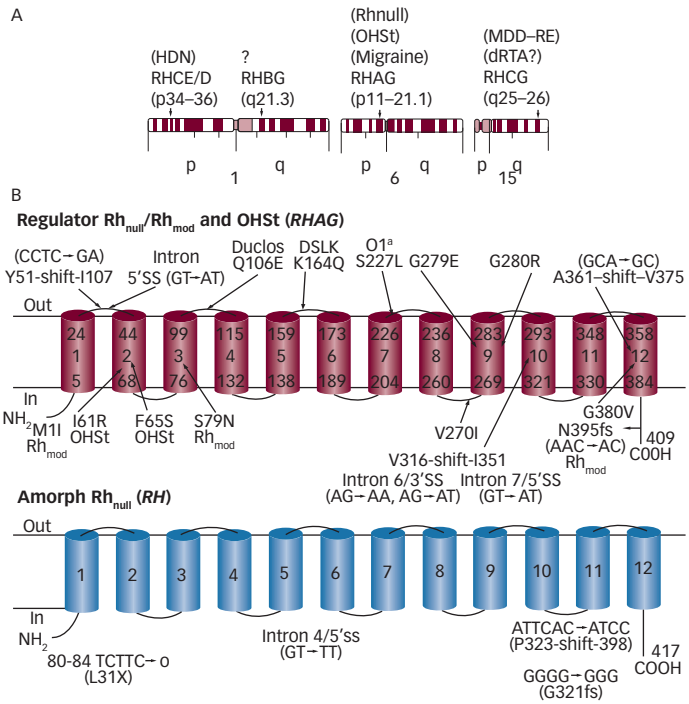
The tree contains over 100 members and is not rooted. Epithelial branches (RhBG and RhCG) and erythroid branches (RhAG and Rh30) are denoted along with Rhp2, the vertebrate Rh precursor group, and Rhp1, the invertebrate Rh group as a whole. For abbreviations and identification numbers of organisms, see supplemental information in reference 23. Rh30 include both RhD and RhCE members. The Nitrosomonas europaea *Rh (#001) and unicellular Rh homologues are situated between Rhp1 and Rhp2.*

Structure and Function of Rhesus Proteins

Members of the Rh family share a common membrane fold typical of 12 spanning segments as seen in human erythroid and epithelial Rh proteins.[10–16] The consensus primary and secondary structural features of the whole Rh family have also been identified.[23,27] In contrast to the majority of membrane proteins with a basic isoelectric point (pI), all Rh proteins except Rh30 have an acidic pI.[24] Thus, in humans and other mammals, RhAG, RhBG, and RhCG are negatively charged, while Rh30 are positively charged under physiological pH. The recently determined structure of a bacterial Rh protein from *Nitrosomonas europaea*,[29,30] an ortholog of human RhAG, RhBG, and RhCG,[23] reveals a trimer with each monomer forming a narrow pore controlled by four conserved amino acids, two Phe, and two His residues. This structure provides an atomic template for meaningful modeling of Rh proteins to obtain insight into their function.

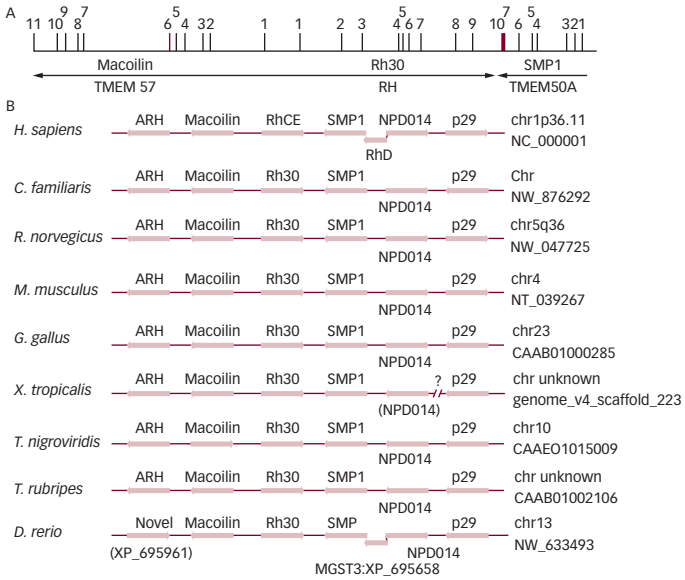
Over the past decade, both erythroid and epithelial Rh proteins have been under extensive study in the context of their structure–function relationships. These studies concern two hypotheses, one being that the physiological substrate of Rh proteins is CO_2, be it a neutral carbonic acid molecule or a bicarbonate anion (H_2CO_3/HCO_3^-).[31–35] The other hypothesis is that like microbial ammonium transport proteins (Amt), Rh proteins transport ammonia,[36] be it a gas molecule or a cationic species (NH_3/NH_4^+).[37–48] In the case of human RhAG, studies of normal red cells and Rh_{null} cells suggested that RhAG is a gas channel for CO_2,[35] while other studies suggested that it is a channel for NH_3[40] or both NH_3 and NH_4^+.[41]

Figure 2: Human Rhesus Genes are Mapped on Four Chromosomal Regions with Distinct Disease Association



A: Diagram of human chromosome 1, 6, and 15, where RH, RHBG, RHAG, *and* RHCG *are denoted along with disease association. B: Diagram of mutation in red cell Rhesus genes,* RHAG *(upper) and* RH *(lower). The changes in Rh$_{mod}$ and over-hydrated hereditary stomatocytosis (OHSt) are indicated, and others are Rh$_{null}$ mutations, as compiled in the Human Blood Group Antigen Gene Mutation Database.[17] In RhAG, the designation of transmembrane domains is based on the crystal structure of NeRh.[29,30] Three changes in the surface loops of RhAG are potentially neutral antigenic polymorphisms (Duclos, DSLK and O1^a).[58]*
HDN = hemolytic disease of the newborn; MDD-RE = recurrent early-onset major depressive disorders; dRTA = distal renal tubular acidosis; ? = unknown.

Figure 3: The *RH* Locus Sits in a Highly Conserved Block of Genes in Vertebrates from Zebrafish to Humans



A: Diagram of organization of the RH *locus and its flanking loci (common and official names are shown). Exons of the three genes are numbered in terms of their orientation. In human RhD it is a duplicate downstream of SMP1 (TMEM50A), whereas RhCE is oriented in reverse to Macoilin (TMEM57). B: Conservation of the* RH *gene block. Shown are species (left) and chromosome and accession numbers (right).*

Similarly, it remains unsolved whether human RhBG and RhCG mediate electroneutral or electrogenic ammonia passage.[42–47] The possibility has also been raised that RhAG is a promiscuous channel for other gaseous substrates such as nitric oxide and oxygen.[21] Despite the disputes, an agreement on the role of human Rh30 has emerged from the structure of *N. europaea* Rh protein,[29,30] i.e. Rh30 is less likely to retain a transport activity[20] but has a structural role participating in the assembly of Rh complex with the band 3 complex in the membrane.[21] This view is in line with fast evolution of Rh30 proteins signifying a positive selection and thus a functional modification[49–52] in accord with surface-to-volume ratio increase[32] and morphogenetic evolution of mammalian red cells.[22,24]

Significance of Rhesus Proteins in Disease Association

The epithelial RhBG and RhCG proteins are highly conserved due to strong purifying selection,[23] which may account for the apparent lack of documentation of their natural mutation in humans. Nonetheless, their physiological significance has been demonstrated by the knockdown phenotypes of those orthologs in lower organisms. In the green alga *Chlamydomonas reinhardtii*, Rh1 gene knockdown impairs cell growth at high levels of CO_2 and affects the carbon concentration mechanism during growth on acetate.[32] In the worm *Caenorhabditis elegans*, Rh1 gene (*rhr-1*) knockdown causes embryonic lethality and arrests normal development.[34] Recently, the physiological importance of mammalian Rh proteins has been shown in Rhcg knockout mice that manifest pH changes in urine and epididymal fluid,[48] indicating a deficit in pH regulation primary to renal ammonium excretion. In contrast to the negative results seen in Rhbg knockout mice,[53] this work implies that human RhCG, when mutated, may involve distal renal tubular acidosis and male infertility.[48] Interestingly, human *RHCG* has also been identified as a candidate gene for early-onset major depressive disorder,[54] and human *RHAG* is linked to a subtype of migraine[55] (see *Figure 2*). These results echo a wide expression of Rhcg and a focal expression of Rhag in mouse brain,[56] suggesting that the two proteins play critical but differential roles in the central nervous system.

> *Rh antigens, particularly the D antigen, play a pivotal role in hemolytic disease of the fetus and the newborn and their mutations cause Rh$_{null}$ cells (amorph type) with mild phenotypes.*

For erythroid Rh proteins, their association with red cell disorders has been well-recognized. Rh antigens, particularly the D antigen, play a pivotal role in hemolytic disease of the fetus and the newborn[9] and their mutations cause Rh$_{null}$ cells (amorph type) with mild phenotypes.[19,57] By contrast, mutations of *RHAG* result in either recessive Rh deficiency syndrome (regulator type Rh$_{null}$ or Rh$_{mod}$)[17,57] or dominant over-hydrated hereditary stomatocytosis.[58] These two disorders exhibit chronic hemolytic anemia and share certain changes in red cell morphology and biochemical features, such as increased stomatocytosis, cation

permeability, and adenosine triphosphate (ATPase) activity, but decreased lipid asymmetry and membrane flexibility. Such changes indicate that ablation or disruption of the RhAG function also affects red cell integrity. *Figure 2* shows the sites of various mutations resulting in the two disorders and those located at the surface as potential neutral polymorphisms.

Rhesus Allelic Diversity and Relevance to Genomic Transfusion Medicine

Unlike Rh glycoproteins, human RhD and RhCE have undergone extensive allelic diversification. Nonetheless, the two-gene *RH* locus resides in a block of genes encoding highly conserved proteins and retaining conserved chromosomal synteny in vertebrates (see *Figure 3*). If the flanking genes are under conserved evolution, they could shelter the *RH* locus from interference by Rh homologous genes elsewhere in the genome, thereby facilitating homologous recombination between *RHD* and *RHCE*. Currently, about 200 allelic forms have been defined at the molecular level, emphasizing a system complexity.[17] Of such allelic variants, most have changes in the coding sequence and many occur in genetic hybrids of various patterns between *RHD* and *RHCE*.[20,59,60] Hence, *RHD* duplication appears to have served as a driving force rendering the *RH* locus susceptible to homologous recombination. This duplication may also underlie the immunopotency of D antigen and its subsequent adaptation in different populations. Given the quite recent evolutional appearance of *RHD*, the human immune system may still be at work monitoring and adapting to its novel molecular features. In this context, the incompatibility between D-positive fetus and D-negative mother in pregnancy and its attenuation through homologous recombination at the population level can be regarded as the two arms of a balancing act in positive selection.[22]

Genomic or gene-based medicine has been transforming medical practice after the revolutionary progress in decoding the human genome.[61] Similarly, since cloning of the first blood group gene, *GYPA* for M/N glycophorin,[62] the genetic definition of the first blood group variant phenotype S-s-U-[63] and the first DNA typing of M/N genotypes followed.[64] Two decades later immunohematology has been revitalized with extensive new genetic knowledge and is entering the molecular era, which will ultimately lead to genomic transfusion medicine. This knowledge is now being used in applications for patient care in the form of genotyping, as demonstrated in A Workshop on Molecular Methods in Immunohematology.[65] New ideas and methods are being rapidly disseminated to promote such applications and bring their utility to clinical settings, as reflected in recent reviews.[65,66] Although challenges still exist, Rh and ABO, with several others, will continue to serve as models for testing new methods and monitoring blood transfusion and hemolytic disease of the fetus and newborn.

Pitfalls in genotyping of various blood groups have been discussed in the scope of current knowledge. What has been less explored is the power of genome- and locus-wide analyses of single nucleotide polymorphism and copy-number variation in non-coding regions at population levels. With the advanced low-cost DNA sequencing technology, genome- and locus-wide studies, when applied to human blood group systems, should reveal a new and rich source of genetic variations. Such discoveries will further enhance our understanding of human blood groups as inheritable traits and bring about their genotyping with accuracy and simplicity, thus advancing genomic transfusion medicine. ■

Cheng-Han Huang, MD, PhD, is Head of the Laboratory of Biochemistry and Molecular Genetics at Lindsley F Kimball Research Institute of the New York Blood Center in New York City. Dr Huang's studies are focused on model organisms to probe the structure–function relationships of the Rhesus (Rh) protein family, including both its erythroid members and non-erythroid homologs widely expressed in mammalian epithelial tissues. Dr Huang received his MD from Hunan Medical University (formerly Yale-in-China Medical College) in 1982 and his PhD in biochemistry from the Albert Einstein College of Medicine in New York in 1990.

1. Landsteiner K, *Zentbl Bakt Orig*, 1900;27:357–62.
2. Landsteiner K, *Wien Klin Wochenschr*, 1901;14:1132–4.
3. Levine P, Stetson RE, *JAMA*, 1939; 113:126–7.
4. Landsteiner K, Weiner AS, *Proc Soc Exp Biol Med*, 1940;43:223.
5. Watkins WM, *Science*, 1966;152:172–81.
6. Moore S, Woodrow CF, McClelland DBL, *Nature*, 1982;295: 529–31.
7. Gahmberg CG, *FEBS Lett*, 1982;140:93–7.
8. Moore S, Green C, *Biochem J*, 1987;244:735–41.
9. Klein HG, Anstee DJ, *Mollison's blood transfusion in clinical medicine, 11th ed*, Oxford: Blackwell Publishing, 2005;496–545.
10. Avent ND, et al., *Biochem J*, 1990;271:821–5.
11. Cherif-Zahar B, et al., *Proc Natl Acad Sci U S A*, 1990;87:6243–7.
12. Le Van Kim C, et al., *Proc Natl Acad Sci U S A*, 1992;89:10925–9.
13. Arce MA, et al., *Blood*, 1993;82:651–5.
14. Ridgwell K, et al., *Biochem J*, 1992;287:223–8.
15. Liu Z, et al., *J Biol Chem*, 2000;275:25641–51.
16. Liu Z, et al., *J Biol Chem*, 2001;276:1424–33.
17. Blumenfeld OO, Blood Group Gene Mutation Database. Available at: www.ncbi.nlm.nih.gov/gv/rbc/xslcgi.fcgi?cmd=bgmut/systems_info&system=rh
18. Avent ND, et al., *Expert Rev Mol Med*, 2006;8:1–20.
19. Le Van Kim C, et al., *Blood Rev*, 2006;20:93–110.
20. Westhoff CW, *Semin Hematol*, 2007;44:42–50.
21. Burton NM, Anstee D, *Curr Opin Hematol*, 2008;15:625–30.
22. Huang CH, *Hematol*, 2008;2:149–57.
23. Huang CH, Peng J, *Proc Natl Acad Sci U S A*, 2005;102:15512–17.
24. Peng J, Huang CH, *Transf Clin Biol*, 2006;13:85–94.
25. Schierwater B, *Trichoplax adhaerens*, *BioEssays*, 2006;27: 1294–1302.
26. Srivastava M, et al., *Nature*, 2008;454: 955–60.
27. Huang CH, Liu PZ, *Blood Cells Mol Dis*, 2001;27:90–101.
28. Carritt B, Kemp TJ, Poulter M, *Hum Mol Genet*, 1997;6:843–50.
29. Li X, et al., *Proc Natl Acad Sci U S A*, 2007;104:19279–84.
30. Lupo D, et al., *Proc Natl Acad Sci U S A*, 2007;104:19303–8.
31. Soupene E, et al., *Proc Natl Acad Sci U S A*, 2002;99:7769–73.
32. Soupene E, et al., *Proc Natl Acad Sci U S A*, 2004;101:7787–92.
33. Kustu S, Inwood B, *Transf Clin Biol*, 2006;13:103–10.
34. Ji Q, et al., *Proc Natl Acad Sci U S A*, 2006;103:5881–6.
35. Endeward V, *FASEB J*, 2008;22:64–73.
36. Marini AM, et al., *Trends Biochem Sci*, 1997;22:460–61.
37. Marini AM, et al., *Nat Genet*, 2000;26:341–4.
38. Westhoff CM, et al., *J Biol Chem*, 2002; 277:12499–502.
39. Hemker MB, et al., *Br J Haematol*, 2003;122:333–40.
40. Ripoche P, et al., *Proc Natl Acad Sci U S A*, 2004;101:17222–7.
41. Benjelloun F, et al., *Pflugers Arch*, 2005;450:155–67.
42. Ludewig U, *J Physiol (Paris)*, 2004;559:751–9.
43. Bakouh NL, et al., *J Biol Chem*, 2004;279:15975–83.
44. Nakhoul NL, et al., *Am J Physiol Renal Physiol*, 2005;288: F170–81.
45. Zidi-Yahiaoui N, et al., *Biochem J*, 2005;391(Pt 1):33–40.
46. Mak DO, et al., *Am J Physiol Renal Physiol*, 2006;290:F297–305.
47. Mayer M, et al., *J Gen Physiol*, 2006;127:133–44.
48. Biver S, et al., *Nature*, 2008;456:339–43.
49. Kitano T, et al., *Biochem Biophys Res Commun*, 1998;249:78–85.
50. Matassi G, et al., *J Mol Evol*, 1999;48:151–9.
51. Huang CH, et al., *J Mol Evol*, 2000;51:76–87.
52. Kitano T, Saitou N, *Immunogenet*, 2000;51:856–62.
53. Chambrey R, et al., *Am J Physiol Renal Physiol*, 2005;289: F1281–F1290.
54. Verma R, et al., *Biol Psychiatry*, 2008;63:1185–9.
55. Norberg A, et al., *Neurosci Lett*, 2006;396:137–42.
56. Lein ES, et al., *Nature*, 2007;446:168–76.
57. Huang CH, Liu PZ, Cheng JG, *Semin Hematol*, 2000;37:150–65.
58. Bruce LJ, et al., *Blood*, 2009;113(6):1350–57.
59. Wagner EF, Flegel WA, *Immunohematol*, 2004;20:23–36.
60. Wagner FF, Moulds JM, Flegel WA, *Transfusion*, 2005;45: 338–45.
61. Guttmacher AE, Collins FS, *N Engl J Med*, 2002;347:1512–20.
62. Siebert PD, Fukuda M, *Proc Natl Acad Sci U S A*, 1986;83:1665–9.
63. Huang CH, et al., *Blood*, 1987;70:1830–35.
64. Huang CH, et al., *Blood*, 1991;77:381–6.
65. A Workshop on Molecular Methods in Immunohematology, *Transfusion*, 2007;47:1S–100S.
66. Avent ND, *Br J Haematol*, 2009;144(1):3–13.