NCBI dbRBC database of allelic variationsof genes encoding antigens of blood group systems

Santosh Patnaik[1]

[1]Affiliation not available

May 5, 2020

ABSTRACT

Analogous to human leukocyte antigens, blood group antigens are surface markers on the erythro-cyte cell membrane whose structures differ among individuals and which can be serologically identified. The Blood Group Antigen Gene Mutation Database (BGMUT) is an online repository of allelic variations in genes that determine the antigens ofvarious human blood group systems. The database is manually curated with allelic information collated from scientific literature and from direct submissions from research laboratories. Currently, the database documents sequence variations of a total of 1251 alleles of all 40 gene loci that together are known to affect antigens of 30 human blood group systems. When available, information on the geographic or ethnic prevalence of an allele is also provided. The BGMUT website also has general in-formation on the human blood group systems and the genes responsible for them. BGMUT is a part ofthe dbRBC resource of the National Center for Biotechnology Information, USA, and is available online at http://www.ncbi.nlm.nih.gov/projects/gv/rbc/xslcgi.fcgi?cmd=bgmut. The database should be of use to members of the transfusion medicine community, those interested in studies of genetic variation and related topics such as human migrations, and students as well as members of the general public.

INTRODUCTION

The BGMUT Blood Group Antigen Gene Mutation Database documents variations in genes that encode antigens for human blood groups. It is a part of the dbRBC resource (1) of the National Center for Biotechnology Information (NCBI) of USA and can be freely accessed online at http://www.ncbi.nlm.nih.gov/projects/gv/rbc/xslcgi.fcgi?cmd=bgmut.Recent documentation of the extent and the surprisingly high numbers of mutations in the human genome have suggested that, perhaps with the exception of identical twins, no two individuals bear exact copies of chromosomal DNA. In those studies, DNA of random subjects is compared but more often phenotypic differences observed in disease states, whether single gene inherited disorders or in association studies of complex conditions are taken as criteria for selection of individuals whose DNA is examined for sequence changes. In the latter studies, large fragments of DNA are usually examined and compared statistically to matched control individuals. Changes in blood group phenotypes are another criterion for selection of subjects who may show differences in sequences of two or more defined sets of genes. These genes encode a group of red cell membrane proteins that are polymorphic in the population and are defined as blood group antigens; in addition, these genes may encode certain glycosyl transferases that are involved in the synthesis of red cell membrane glycans whose structures also differ among individuals. The former group consists of structural molecules, channels, adhesion molecules or enzymes and, excluding their red cell membrane location, can be considered as representative of any human protein, whereas the latter are similar tomost other glycosyl transferases. Evidence suggests that sequence changes in these proteins or even their absence,such as in null phenotypes, is not, in most cases, physiologically harmful.

The proteins or the glycans fulfill their role as blood group antigens because they are polymorphic in the population and their sequence changes can be readily predictedby serological approaches; they are known as ‘antigens ’because in the course of transfusion, or pregnancy, the presence of a variant protein epitope is recognized as‘non-self’ and may ultimately result in an adverse immunological reaction. The use of transfusion being ubiquitous in the practice of medicine, populations world-wide are serologically tested and variant antigens and their genes, in contrast to many other variant genes arebeing documented in a large number of diverse populations. Although some variants occur rarely and, may onlybe observed in a single individual or family, others appearin unexpectedly large populations, such as the MiIIIphenotype encoded by the MiIIIGYPAgene, in Taiwan (incidence can be as high as 88% among Ami tribes) (2).For many alleles, the database also provides informationon the geographic or ethnic origin of alleles and/or theirassociated serological phenotypes when such informationwas presented in the publications describing the alleles. This may be of use to those interested in populationmigrations.

HISTORY AND CURRENT STATE

BGMUT was developed in 1999 as a locus-specific gene mutation database under the aegis of the Human-Variation Genome Society. It was curated under the direction of one of the authors (OOB) with original information contributed by more than a dozen blood group system experts. The database was hosted online by the Department of Biochemistry at the Albert Einstein College of Medicine, New York. BGMUT was identifiedas one of three model locus-specific databases from morethan 200 in a scholarly review (3). In 2006, BGMUTbecame a part of the dbRBC resource of the NCBI. At dbRBC, curatorship and direction for maintenanceof the database has been provided by another of the authors (WH). The number of alleles in BGMUT has ap-proximately doubled since 2004 when the database wasfirst described in a scholarly publication (4). This publica-tion has been referred to many times in the scientific literature indicating that BGMUT has been a usefulresource. Links to BGMUT database records are availableon relevant pages of the Wikipedia online encyclopedia,and on many of NCBI’s online resources. BGMUT is alsoa part of the PhenCode project which attempts to integrate genetic variation data with the UCSC Genome Browser (5).As of August 2011, BGMUT had 1251 alleles belongingto 40 genes that are together responsible for 30 humanblood group systems (Table 1). Alleles of some genes, such as ABO and RHCE/D which are, respectively, responsible for the ABO and Rh blood group systemsmost frequently examined in the populations, are morenumerous than those of others (Table 1). As per theInternational Society for Blood Transfusion, there are 30human blood group systems (http://ibgrl.blood.co.uk/ISBT%20Pages/ISBT%20Terminology%20Pages/Table%20of%20blood%20group%20systems.htm), all of whichare covered by BGMUT. The Globoside (GLOB) system iscurrently considered a part of the P1PK system inBGMUT, which additionally considers the system of T and Tn antigens as a separate blood group system.DATABASE ARCHITECTUREBGMUT is accessible online for view, search or for ad-ministration as a website in the form of HTML (hyper-text markup language) pages (Figures 1 and 2). A Microsoft SQL Server relational database is used fordata storage. Programmatic code in SQL (structuredquery language), C++and XSLT (extensible structuredlanguage transformations) languages is used for inter-action with the SQL Server database and for renderingBGMUT’s web-based interface. Raw data on all or asubset of the alleles described in BGMUT can be down-loaded as tab-delimited or comma-separated (CSV) textformats from the BGMUT website. Compilations ofallelic sequences for the ABO, H, MNS and Rh systemsin the Microsoft Excel format are also available fordownload.

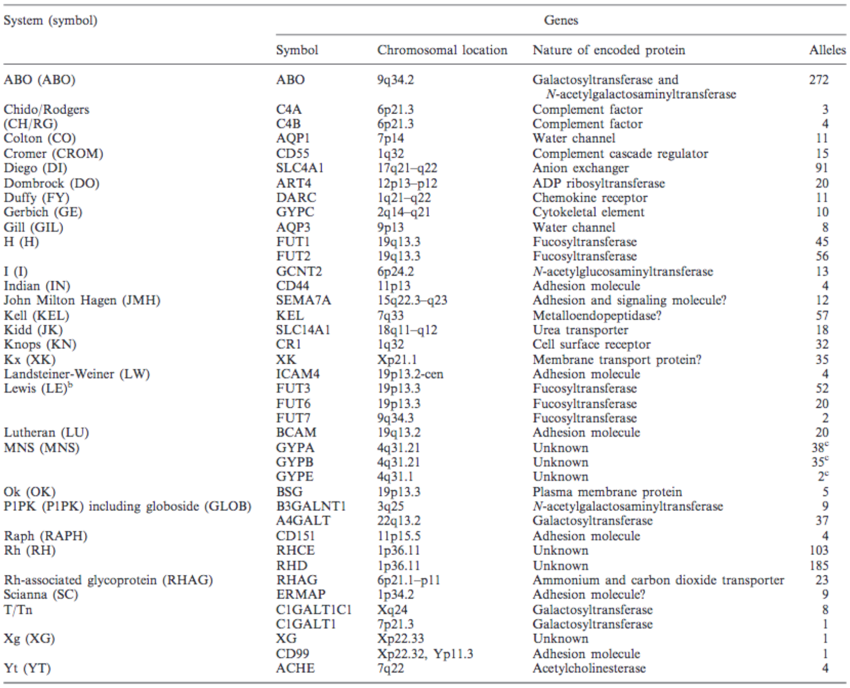
DATABASE CURATION AND ALLELE SUBMISSION

The BGMUT database is curated manually. Allelic information is periodically collated from scientific literature, asis the case for a majority of the alleles listed in BGMUT, or is obtained as direct submissions from researchers through the database website. During the process ofcuration, good quality of methods used in a study is ascertained. The new candidate allele's sequence that has been published and/or submitted to a publicly available repository such as NCBI's GenBank is compared to BGMUT's reference allele for the gene sequence. The sequence positions and the kinds of the deduced aminoacid changes are also verified. Authors are consulted incase of a question or disagreement. For direct submission of information on a new allele for inclusion in the database, a scientific publication describing the allele is not required. However, submitters are encouraged to deposit the allele's sequence in a publicly available repository. In the absence of a scientific publication, this is arequirement.

ALLELES IN BGMUT

Alleles in BGMUT are grouped by the blood group system that the genes they belong to affect. For each allele in the database, BGMUT provides details on the nucleotide changes and the deduced amino acid changes in the protein encoded by the gene the allele belongs to. These changes are in context of a 'reference' allele that itself is included in BGMUT, and is the same for all alleles of agene. Besides the information on the sequence changes, BGMUT also details for an allele the frequency of occurrence, the associated blood group phenotype, references to the studies that identified and characterized the allele and accession numbers of the relevant sequences in NCBI GenBank when such information is available. GenBank accession numbers, however, are not available for many alleles because though they have been described in published literature, their sequences were not deposited inthe repository by the authors. When known, the regions of the gene or cDNA that were sequenced to identify the allele, the prevalence of the allele in different geographical regions or ethnic populations and association of the allele with diseases are also noted. Often, a name is also provided for an allele. Names make it easier to refer to alleles and can indicate the associated phenotype and/or nucleotide or amino acidvariation.

Table 1. Blood group systems in the BGMUT database



System (symbol)	Genes			
	Symbol	Chromosomal location	Nature of encoded protein	Alleles
ABO (ABO)	ABO	9q34.2	Galactosyltransferase and *N*-acetylgalactosaminyltransferase	272
Chido/Rodgers (CH/RG)	C4A	6p21.3	Complement factor	3
	C4B	6p21.3	Complement factor	4
Colton (CO)	AQP1	7p14	Water channel	11
Cromer (CROM)	CD55	1q32	Complement cascade regulator	15
Diego (DI)	SLC4A1	17q21–q22	Anion exchanger	91
Dombrock (DO)	ART4	12p13–p12	ADP ribosyltransferase	20
Duffy (FY)	DARC	1q21–q22	Chemokine receptor	11
Gerbich (GE)	GYPC	2q14–q21	Cytokeletal element	10
Gill (GIL)	AQP3	9p13	Water channel	8
H (H)	FUT1	19q13.3	Fucosyltransferase	45
	FUT2	19q13.3	Fucosyltransferase	56
I (I)	GCNT2	6p24.2	*N*-acetylglucosaminyltransferase	13
Indian (IN)	CD44	11p13	Adhesion molecule	4
John Milton Hagen (JMH)	SEMA7A	15q22.3–q23	Adhesion and signaling molecule?	12
Kell (KEL)	KEL	7q33	Metalloendopeptidase?	57
Kidd (JK)	SLC14A1	18q11–q12	Urea transporter	18
Knops (KN)	CR1	1q32	Cell surface receptor	32
Kx (XK)	XK	Xp21.1	Membrane transport protein?	35
Landsteiner-Weiner (LW)	ICAM4	19p13.2-cen	Adhesion molecule	4
Lewis (LE)^{b}	FUT3	19p13.3	Fucosyltransferase	52
	FUT6	19p13.3	Fucosyltransferase	20
	FUT7	9q34.3	Fucosyltransferase	2
Lutheran (LU)	BCAM	19q13.2	Adhesion molecule	20
MNS (MNS)	GYPA	4q31.21	Unknown	38^{c}
	GYPB	4q31.21	Unknown	35^{c}
	GYPE	4q31.1	Unknown	2^{c}
Ok (OK)	BSG	19p13.3	Plasma membrane protein	5
P1PK (P1PK) including globoside (GLOB)	B3GALNT1	3q25	*N*-acetylgalactosaminyltransferase	9
	A4GALT	22q13.2	Galactosyltransferase	37
Raph (RAPH)	CD151	11p15.5	Adhesion molecule	4
Rh (RH)	RHCE	1p36.11	Unknown	103
	RHD	1p36.11	Unknown	185
Rh-associated glycoprotein (RHAG)	RHAG	6p21.1–p11	Ammonium and carbon dioxide transporter	23
Scianna (SC)	ERMAP	1p34.2	Adhesion molecule?	9
T/Tn	C1GALT1C1	Xq24	Galactosyltransferase	8
	C1GALT1	7p21.3	Galactosyltransferase	1
Xg (XG)	XG	Xp22.33	Unknown	1
	CD99	Xp22.32, Yp11.3	Adhesion molecule	1
Yt (YT)	ACHE	7q22	Acetylcholinesterase	4

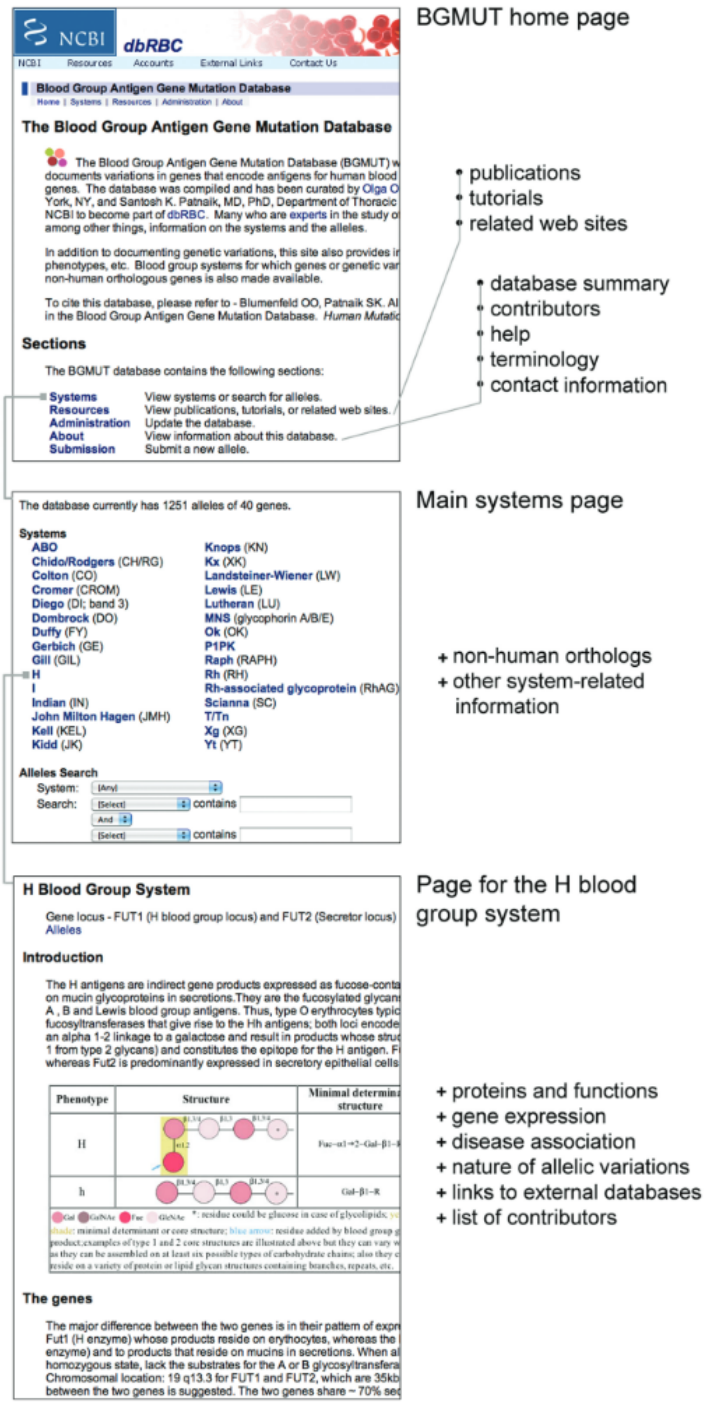
a) As of August 2011; total 30 systems, 40 genes and 1251 alleles.

b)TheFUT3,FUT6andFUT7genes are not expressed in erythroid cells.

c) Including hybrid alleles with another member of theGYPA/B/Egene family

BGMUT uses the allele names given by the dis-coverers of the alleles and/or those that are commonly used by blood group system experts. In the absence of such a name, the curator may provide an arbitrary name. It should be noted that every allele entry in BGMUT has a unique database identification number associated with it. The naming of alleles has presented a problem com-pounded by names given to serological phenotypes ofantigens they encode. At this time, efforts by the transfusion medicine community are in progress to standardize and simplify this nomenclature (6). In the database, tomore generally identify each allele, in most cases, 'name'includes the gene name followed by the DNA change particular to the allele; to facilitate the recognition of the phenotype//genotype connection, under 'alias' is given the serological designation as used by the original authors. In some cases, such as for the Rh system,because of traditional reasons, this arrangement is reversed; in others, such as ABO, only the phenotype des-ignation marked by consecutive numbers is given; this is because the large number of sites of DNA changes in many alleles makes their listing under 'name' unwieldy.

Alignments of DNA changes in alleles of a single geneor a gene family may uncover patterns that provide some insights into the process of diversification, including gene rearrangements and meiotic or unequal recombination (7).An attempt to correlate DNA changes in the gene(s)encoding the antigen(s), with the serological phenotype or the glycosyltransferase activity, shows that even thoughvariation in common alleles occurs within the samesequence stretches that define the common epitopicregions (4), in rare alleles, changes that affect the sero-logical response are seen throughout the exons. The result-ing amino acid alterations may cause altered folding that interferes with epitope presentation or gives rise to newepitopes. As expected, null phenotypes usually result fromabsence of an intact protein and are caused by nonsensemutations, insertions or deletions, or absence of thecommon epitope due to recombination. More generally,in the systems documented in the database, the phenotype reflects a particular sequence alteration be it the result of a single or multiple molecular events (e.g. the Rh neg, ABOO and MNS Sta alleles). Clearly, in contrast to many global studies, data presented in BGMUT allow to focuson DNA changes in a single gene and a closer look at itsresulting alleles, to ask how they came about and toconsider the properties of the resulting proteins. This knowledge will only expand as complete sequence content of variant alleles become more readily available.


NCBI
dbRBC
NCBI Resources Accounts External Links Contact Us
Blood Group Antigen Gene Mutation Database
Home | Systems | Resources | Administration | About
BGMUT home page
publications
tutorials
related web sites
database summary
contributors
help
terminology
contact information
Main systems page
+ non-human orthologs
+ other system-related information
Page for the H blood group system
+ proteins and functions
+ gene expression
+ disease association
+ nature of allelic variations
+ links to external databases
+ list of contributors
System: [Any]
Search: [Select] contains
And
[Select] contains

The Blood Group Antigen Gene Mutation Database

The Blood Group Antigen Gene Mutation Database (BGMUT) w
documents variations in genes that encode antigens for human blood
genes. The database was compiled and has been curated by Olga O
York, NY, and Santosh K. Patnaik, MD, PhD, Department of Thoracic
NCBI to become part of dbRBC. Many who are experts in the study o
among other things, information on the systems and the alleles.

In addition to documenting genetic variations, this site also provides i
phenotypes, etc. Blood group systems for which genes or genetic var
non-human orthologous genes is also made available.

To cite this database, please refer to - Blumenfeld OO, Patnaik SK. Al
in the Blood Group Antigen Gene Mutation Database. *Human Mutatio*

Sections

The BGMUT database contains the following sections:

Systems	View systems or search for alleles.
Resources	View publications, tutorials, or related web sites.
Administration	Update the database.
About	View information about this database.
Submission	Submit a new allele.

The database currently has 1251 alleles of 40 genes.

Systems

ABO
Chido/Rodgers (CH/RG)
Colton (CO)
Cromer (CROM)
Diego (DI; band 3)
Dombrock (DO)
Duffy (FY)
Gerbich (GE)
Gill (GIL)
H
I
Indian (IN)
John Milton Hagen (JMH)
Kell (KEL)
Kidd (JK)
Knops (KN)
Kx (XK)
Landsteiner-Wiener (LW)
Lewis (LE)
Lutheran (LU)
MNS (glycophorin A/B/E)
Ok (OK)
P1PK
Raph (RAPH)
Rh (RH)
Rh-associated glycoprotein (RhAG)
Scianna (SC)
T/Tn
Xg (XG)
Yt (YT)

Alleles Search

H Blood Group System

Gene locus - FUT1 (H blood group locus) and FUT2 (Secretor locus)
Alleles

Introduction

The H antigens are indirect gene products expressed as fucose-conta
on mucin glycoproteins in secretions.They are the fucosylated glycan
A , B and Lewis blood group antigens. Thus, type O erythrocytes typic
fucosyltransferases that give rise to the Hh antigens; both loci encode
an alpha 1-2 linkage to a galactose and result in products whose stru
1 from type 2 glycans) and constitutes the epitope for the H antigen. F
whereas Fut2 is predominantly expressed in secretory epithelial cells

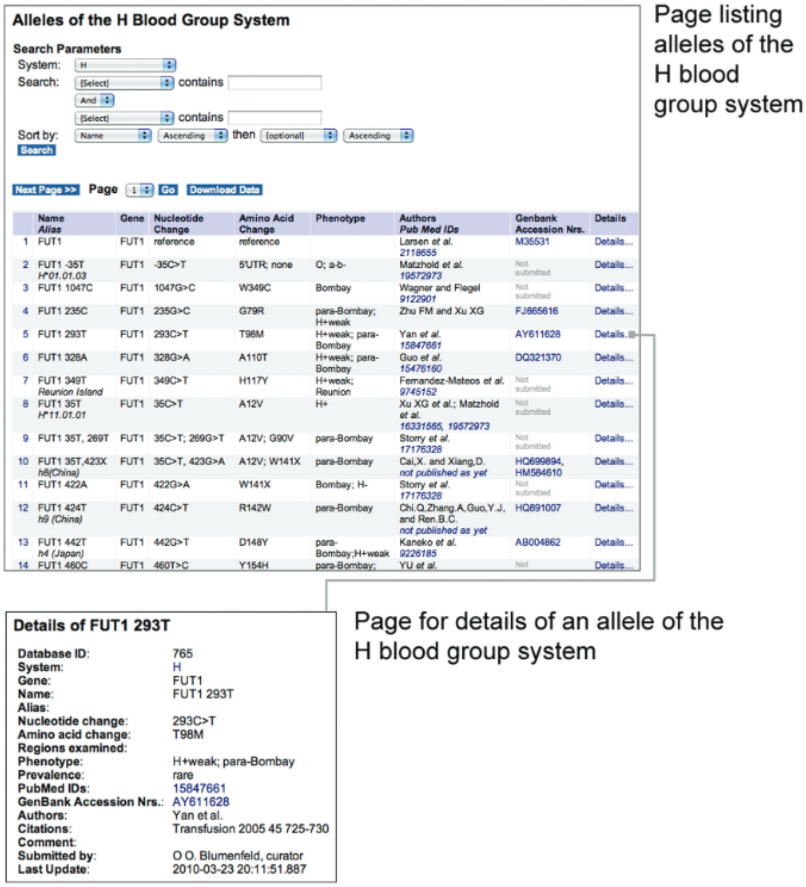
Phenotype	Structure	Minimal determina structure
H		Fuc–α1→2–Gal–β1–
h		Gal–β1–R

● Gal ● GalNAc ● Fuc ● GlcNAc *: residue could be glucose in case of glycolipids; ye
shade: minimal determinant or core structure; blue arrow: residue added by blood group g
product;examples of type 1 and 2 core structures are illustrated above but they can vary w
as they can be assembled on at least six possible types of carbohydrate chains; also they c
reside on a variety of protein or lipid glycan structures containing branches, repeats, etc.

The genes

The major difference between the two genes is in their pattern of expr
Fut1 (H enzyme) whose products reside on erythocytes, whereas the
enzyme) and to products that reside on mucins in secretions. When al
homozygous state, lack the substrates for the A or B glycosyltransfera
Chromosomal location: 19 q13.3 for FUT1 and FUT2, which are 35kb
between the two genes is suggested. The two genes share ~ 70% se

Figure 1. A montage of screenshots of the web-based interface of BGMUT depicting the hierarchically ordered home page, the page listing theblood group systems in the database and the page for a specific blood group system (H). The nature of content of the pages is visible in thescreenshots or is noted in the figure.



Alleles of the H Blood Group System

Search Parameters

System: H

Search: [Select] contains

And

[Select] contains

Sort by: Name Ascending then [optional] Ascending

	Name *Alias*	Gene	Nucleotide Change	Amino Acid Change	Phenotype	Authors *Pub Med IDs*	Genbank Accession Nrs.	Details
1	FUT1	FUT1	reference	reference		Larsen *et al.* *2118655*	M35531	Details...
2	FUT1 -35T *H*01.01.03*	FUT1	-35C>T	5'UTR; none	O; a-b-	Matzhold *et al.* *19572973*	Not submitted	Details...
3	FUT1 1047C	FUT1	1047G>C	W349C	Bombay	Wagner and Flegel *9122901*	Not submitted	Details...
4	FUT1 235C	FUT1	235G>C	G79R	para-Bombay; H+weak	Zhu FM and Xu XG	FJ665616	Details...
5	FUT1 293T	FUT1	293C>T	T98M	H+weak; para-Bombay	Yan *et al.* *15847661*	AY611628	Details...
6	FUT1 328A	FUT1	328G>A	A110T	H+weak; para-Bombay	Guo *et al.* *15476160*	DQ321370	Details...
7	FUT1 349T *Reunion Island*	FUT1	349C>T	H117Y	H+weak; Reunion	Fernandez-Mateos *et al.* *9745152*	Not submitted	Details...
8	FUT1 35T *H*11.01.01*	FUT1	35C>T	A12V	H+	Xu XG *et al.*; Matzhold *et al.* *16331565, 19572973*	Not submitted	Details...
9	FUT1 35T, 269T	FUT1	35C>T; 269G>T	A12V; G90V	para-Bombay	Storry *et al.* *17176328*	Not submitted	Details...
10	FUT1 35T,423X *h8(China)*	FUT1	35C>T, 423G>A	A12V; W141X	para-Bombay	Cai,X. and Xiang,D. *not published as yet*	HQ699894, HM584610	Details...
11	FUT1 422A	FUT1	422G>A	W141X	Bombay; H-	Storry *et al.* *17176328*	Not submitted	Details...
12	FUT1 424T *h9 (China)*	FUT1	424C>T	R142W	para-Bombay	Chi.Q,Zhang.A,Guo,Y.J, and Ren.B.C. *not published as yet*	HQ691007	Details...
13	FUT1 442T *h4 (Japan)*	FUT1	442G>T	D148Y	para-Bombay;H+weak	Kaneko *et al.* *9226185*	AB004862	Details...
14	FUT1 460C	FUT1	460T>C	Y154H	para-Bombay;	YU *et al.*	Not	Details...

Page listing alleles of the H blood group system

Details of FUT1 293T

Database ID:	765
System:	H
Gene:	FUT1
Name:	FUT1 293T
Alias:	
Nucleotide change:	293C>T
Amino acid change:	T98M
Regions examined:	
Phenotype:	H+weak; para-Bombay
Prevalence:	rare
PubMed IDs:	15847661
GenBank Accession Nrs.:	AY611628
Authors:	Yan et al.
Citations:	Transfusion 2005 45 725-730
Comment:	
Submitted by:	O O. Blumenfeld, curator
Last Update:	2010-03-23 20:11:51.887

Page for details of an allele of the H blood group system

Figure 2. Screenshots of the BGMUT web pages showing a partial list of alleles and a specific allele of the H blood group system

OTHER RESOURCES WITHIN BGMUT

Besides the alleles of 30 human blood group systems, BGMUT also provides information on the systems themselves (Figure 1). Such information includes that on as ystem's genes, such as their sequences, chromosomal lo-cations and the functions of proteins encoded by them.Links to database records elsewhere, such as the Online Mendelian Inheritance in Man (OMIM) (8) and PubMed resources of NCBI, are provided. Brief descriptions of diseases associated with a gene's variations, if any, and a summary of the nature of the variations are also provided. Content for such pages for the various blood group systems has been contributed by various experts and is continually updated. The BGMUT website also has a few introductory tutorials on blood group systems and provides links to online resources on blood groups as well as on human gene mutations and polymorphisms in general. Information on the nature of the genes of the H,MNS, Rh and RhAG systems in non-humans, primarily primates and their allelic variations is also provided. BGMUT also makes available brief descriptions of other carbohydrate antigens such as the Tk and Sid (Sda) antigens on erythrocytes.

ADDITIONAL TOOLS OF INTEREST AT THEPARENT dbRBC RESOURCE

Besides BGMUT, the dbRBC resource contains tools andother resources developed at the NCBI to provide accessto publicly available information on human blood groupsystems. Two of the tools are briefly described here. The alignment viewer tool (http://www.ncbi.nlm.nih.gov/pro-jects/gv/rbc/align.fcgi?cmd=aligndisplay) provides an interface to choose alleles of genes of many of the bloodgroup systems for sequence alignment, though allelic entries in BGMUT cannot be directly chosen in thecurrent interface. Chosen sequences can be displayed asgenomic DNA, cDNA or amino acid sequences. Positionsknown to bear gene polymorphisms can be highlighted,and so can be specific differences within a set of allelic sequences. The reference sequence itself can be chosen.Besides the alignment, a user can switch to a display ofselected sequences in FASTA format. Nucleotide orprotein sequences for individual exons and/or intronscan be displayed for each allele in alignment to the refer-ence sequence. When the three dimensional structure of agene's protein product is known, amino acid differences ofvarious alleles can be projected onto the protein model forviewing with the CN3D model viewer (9).dbRBC's sequencing-based typing tool (http://www.ncbi.nlm.nih.gov/projects/gv/rbc/sbt.cgi?cmd=main) isof use to evaluate the allelic composition of sequencing-based typing results of cDNA or genomic sequences (10).

It allows for the comparison of result sequences with se-quences of known alleles of genes of various human bloodgroup systems. The results of the comparison are returnedas a table of potential allele hits, along with the respectivenucleotide changes, and links for closer examination of thealignments within the dbRBC alignment viewer tool.With a continuous increase in the number of availablealleles, dbRBC is testing the implementation of an allelicnomenclature system based on one in use for the HLAsystem (11). Thus, for example, the reference ABO gene allele sequence is named ABO*A1.01.01.01. The nameconsists of the gene name, followed by an asterisk, andfor blocks separated by periods. The leading block is togroup sequences first by their main serological phenotypes (e.g. A1, A2 or Ax phenotypes of the ABO system).Sequences that share the same serological phenotype,but differ in their amino acid sequences, are differentiatedby the second block. The third block lists alleles that havean identical amino acid sequence, but show synonymousmutations in their exonic DNA sequence. The last blockdifferentiates sequences that share an identical DNAsequence in the coding region, but differ in non-codingregions. Alleles of the ABO, FUT1, FUT2 and FUT3 genes have been given additional names as per this system as of July 2011.

FUTURE PLANS

Currently, allelic information compiled in BGMUT is notintegrated with that outside the BGMUT but within thedbRBC resource. The latter set of alleles of genes forhuman blood group systems have been programmaticallycollected from various NCBI resources. Plans for an inte-gration so that BGMUT allelic entries are available foruse within the alignment viewer and the sequencing-basedtyping tools of dbRBC are under consideration. As a part of this integration, BGMUT alleles will be assigned add-itional names as per the nomenclature system outlined inthe above section if a decision is made to implement itafter consultation with various experts.

ACKNOWLEDGEMENTS

We thank Raymond Dunivin and Douglas Hoffman atNCBI for help with the implementation and maintenanceof the BGMUT database.

FUNDING

Funding for developing and maintaining the Database,was provided by the Department of Biochemistry,Albert Einstein College of Medicine of YeshivaUniversity, N.Y., 10461, USA, and from the DavidOpochinsky/Blumenfeld Family Fund (to O.O.B.).Intramural funding of the US National Institutes ofHealth (to NCBI and W.H.)Conflict of interest statement. None declared.

REFERENCES

1. Sayers, E.W., Barrett,T., Benson,D.A., Bolton,E., Bryant,S.H.,Canese,K., Chetvernin,V., Church,D.M., DiCuccio,M.,Federhen,S.et al. (2011) Database resources of the NationalCenter for Biotechnology Information.Nucleic Acids Res.,39,D38–D51.

2. Broadberry, R.E. and Lin,M. (1996) The distribution of the MiIII(Gp.Mur) phenotype among the population of Taiwan.Transfus.Med.,6, 145–148.

3. Claustres, M., Horaitis,O., Vanevski,M. and Cotton,R.G. (2002)Time for a unified system of mutation description and reporting:a review of locus-specific mutation databases.Genome Res.,12,680–688.

4. Blumenfeld, O.O. and Patnaik,S.K. (2004) Allelic genes of bloodgroup antigens: a source of human mutations and cSNPsdocumented in the Blood Group Antigen Gene MutationDatabase.Hum. Mutat.,23, 8–16.

5. Giardine, B., Riemer,C., Hefferon,T., Thomas,D., Hsu,F.,Zielenski,J., Sang,Y., Elnitski,L., Cutting,G., Trumbower,H.et al.(2007) PhenCode: connecting ENCODE data with mutations andphenotype.Hum. Mutat.,28, 554–562.

6. Storry,J.R., Castilho,L., Daniels,G., Flegel,W.A., Garratty,G.,Francis,C.L., Moulds,J.M., Moulds,J.J., Olsson,M.L., Poole,J.et al. (2011) International Society of Blood Transfusion WorkingParty on red cell immunogenetics and blood group terminology:Berlin report.Vox Sang.,101, 77–82.

7. Patnaik, S.K. and Blumenfeld,O.O. (2011) Patterns of humangenetic variation inferred from comparative analysis of allelicmutations in blood group antigen genes.Hum. Mutat.,32,263–271.

8. Hamosh, A., Scott,A.F., Amberger,J.S., Bocchini,C.A. andMcKusick,V.A. (2005) Online Mendelian Inheritance in Man(OMIM), a knowledgebase of human genes and genetic disorders.Nucleic Acids Res.,33, D514–D517.

9. Wang, Y., Geer,L.Y., Chappey,C., Kans,J.A. and Bryant,S.H.(2000) Cn3D: sequence and structure views for Entrez.Trends Biochem. Sci.,25, 300–302.

10. Helmberg, W., Dunivin,R. and Feolo,M. (2004) Thesequencing-based typing tool of dbMHC: typing highlypolymorphic gene sequences.Nucleic Acids Res.,32, W173–W175.

11. Robinson, J., Mistry,K., McWilliam,H., Lopez,R., Parham,P. andMarsh,S.G. (2011) The IMGT/HLA database.Nucleic Acids Res.,39, D1171–D1176