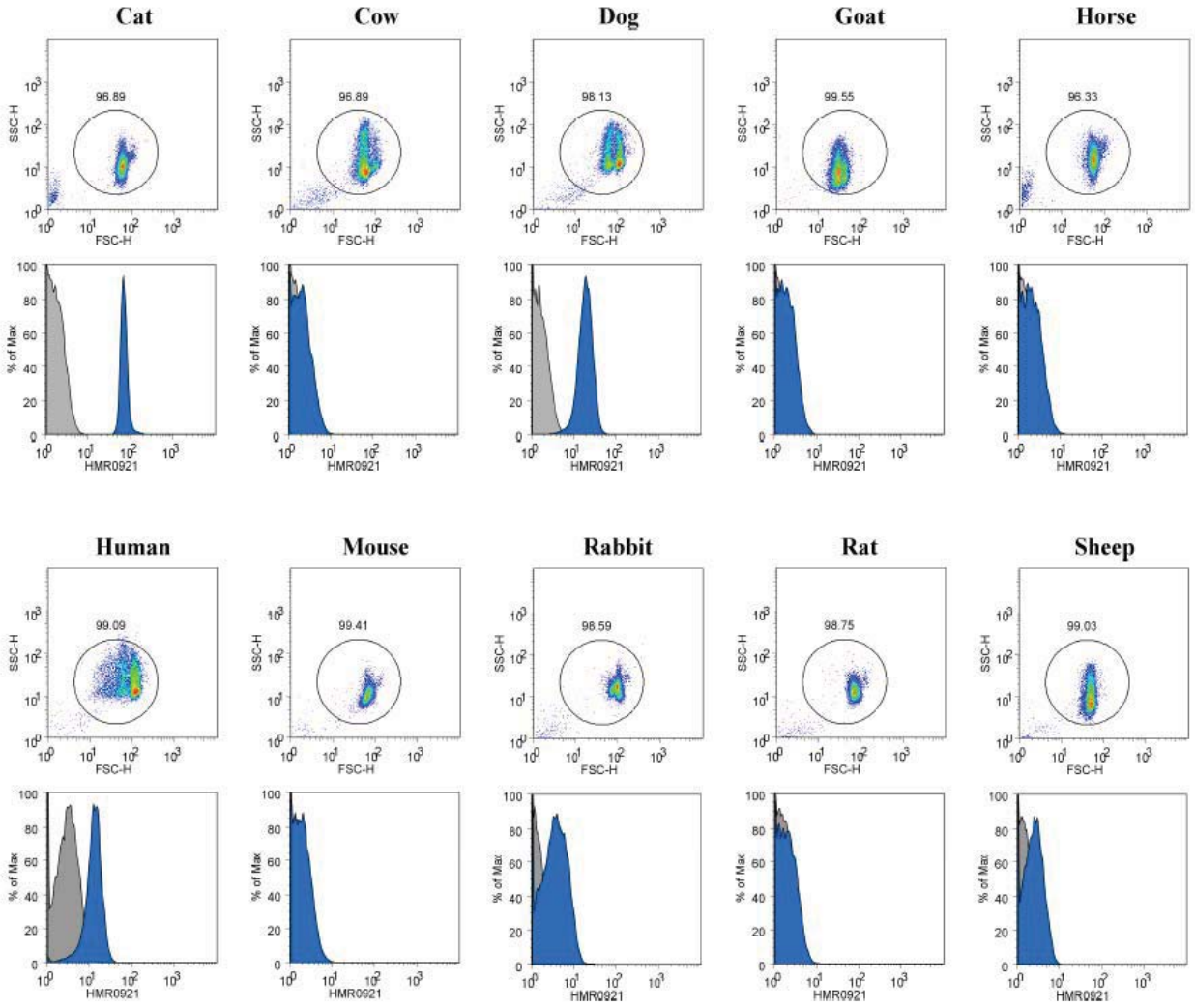
Null alleles of *ABCG2* encoding the breast cancer resistance protein define the new blood group system Junior

Carole Saison, Virginie Helias, Bryan A. Ballif, Thierry Peyrard, Hervé Puy, Toru Miyazaki, Sébastien Perrot, Muriel Vayssier-Taussat, Mauro Waldner, Pierre-Yves Le Pennec, Jean-Pierre Cartron & Lionel Arnaud

- Supplementary Figures 1-5

- Supplementary Tables 1-2

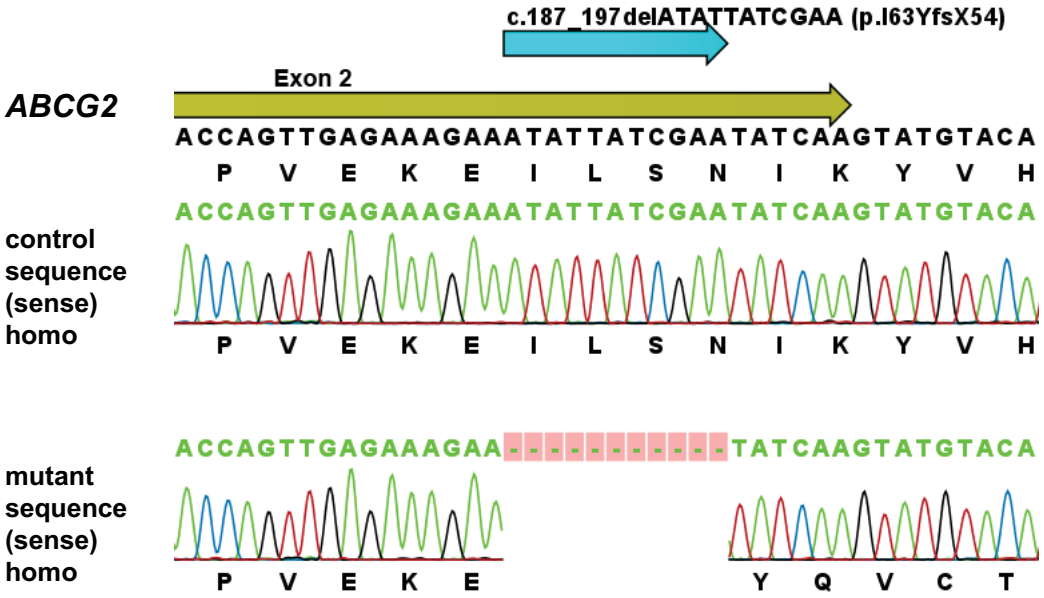


Supplementary Figure 1: Flow cytometry analysis of RBCs from 10 mammalian species with anti-Jr^a HMR0921. Blood samples were taken on EDTA and extensively washed to deplete leukocytes. RBCs were incubated with the human monoclonal HMR0921 (blue profiles) or without (grey profiles) in low-ionic strength solution supplemented with bovine serum albumin. Binding of HMR0921 to RBCs was revealed with a goat F(ab')$_2$ anti-human IgG(H+L)-PE and immediately analyzed with a FACSCalibur flow cytometer, with the same voltage settings as human RBCs on 05/15/2008. Data were analyzed with FlowJo software; the density dot plots show the FSC/SSC gating of RBCs while the overlays show their HMR0921 staining.

Species	Gene ID	Peptide ID	Peptide length	Genomic location
Homo sapiens	ENSG00000118777	ENSP00000237612	655 aa	4:89011416-89152474
Felis catus	ENSFCAG00000002841	ENSFCAP00000002614	587 aa	1506:113877-209541

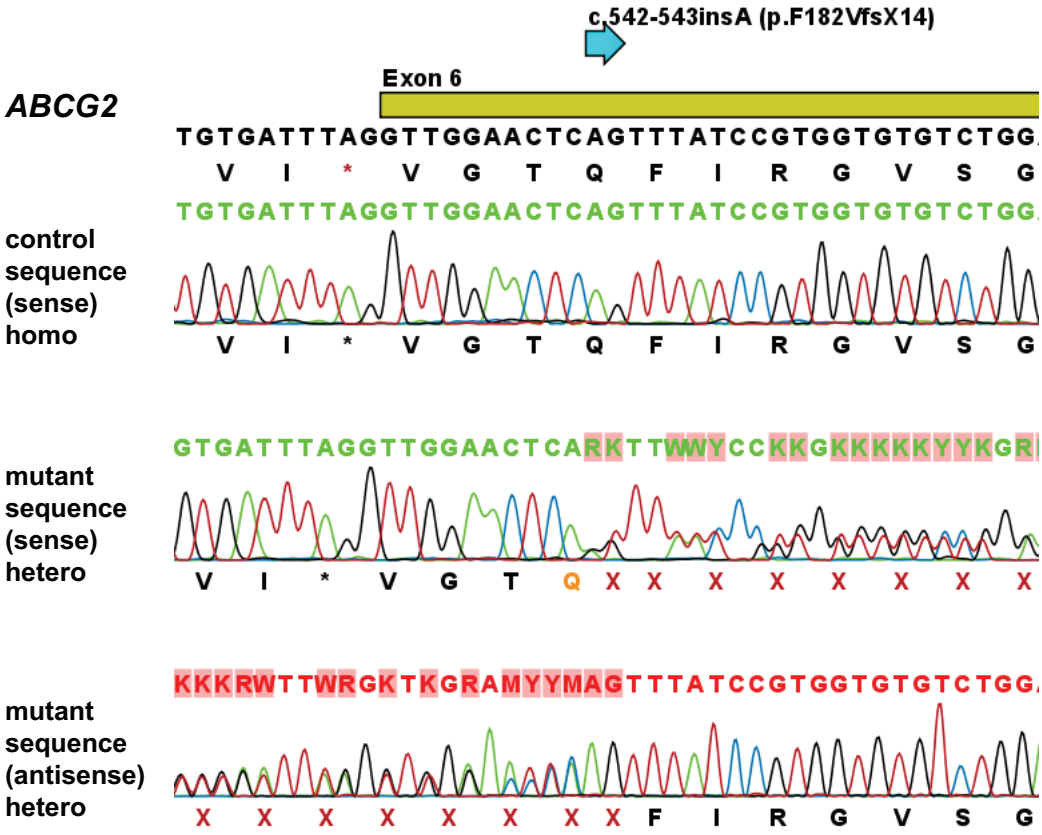
```
CLUSTAL W(1.81) multiple sequence alignment

ENSP00000237612/1-655     MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE
ENSFCAP00000002614/1-587  ------------------------------------------------------------

ENSP00000237612/1-655     KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN
ENSFCAP00000002614/1-587  --------GIMRPGLNAILGPTGGGKSSLLDVLAARKDPHGLSGDVLINGAPRPANFKCN
                                  ***:************************** ********************

ENSP00000237612/1-655     SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT
ENSFCAP00000002614/1-587  SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTTNEKNMRINRVIQELGLDKVADSKVGT
                          ****************************.****.:*** ********************

ENSP00000237612/1-655     QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF
ENSFCAP00000002614/1-587  QFIRGVSGGERKRTSIGMELITDPAILFLDEPTTGLDSSTANAVLLLLKRMSEQGRTIIF
                          ************************:****************************:*******

ENSP00000237612/1-655     SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING
ENSFCAP00000002614/1-587  SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFALMXXXXXXXXXXXXXXXXXXXX
                          ************************************
```

Supplementary Figure 2: Partial alignment of human ABCG2 with predicted cat Abcg2 showing the peptides identified by mass spectrometry from the cat RBC membrane protein immunoprecipitated by anti-Jr^a HMR0921. The tryptic peptides identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS), on the basis of the human genome, are shown in blue in the alignment of human ABCG2 with the predicted cat Abcg2 (alignment with Clustal W software (v. 1.82), as found at http://www.ensembl.org) and the predicted trypsin cleavage sites are underlined. The tryptic peptides identified from cat Abcg2 are fully conserved in human ABCG2, as expected.

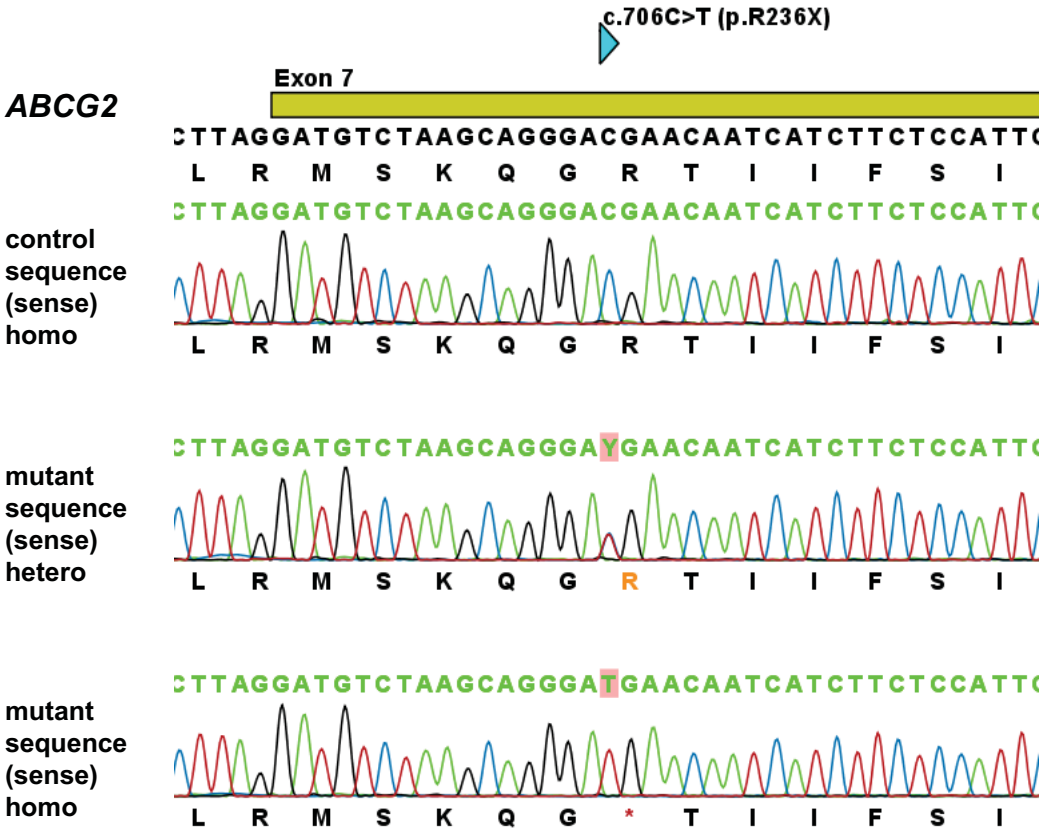


Supplementary Figure 3a: Representative sequencing traces of *ABCG2* null mutation c.187_197del11 (p.I63YfsX54)

Supplementary Figure 3b: Representative sequencing traces of *ABCG2* null mutation c.376C>T (p.Q126X)

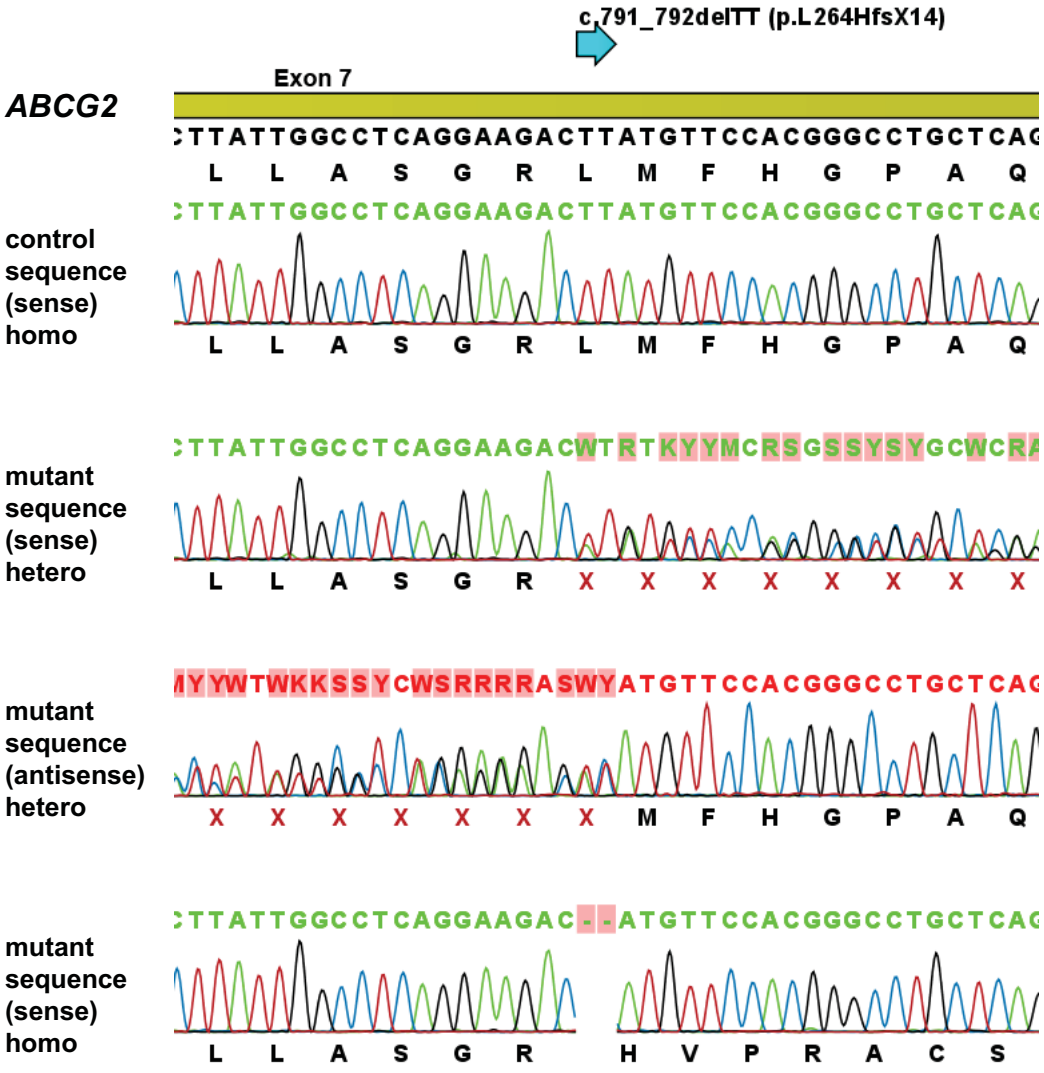


Supplementary Figure 3c: Representative sequencing traces of *ABCG2* null mutation c.542_543insA (p.F182VfsX14)

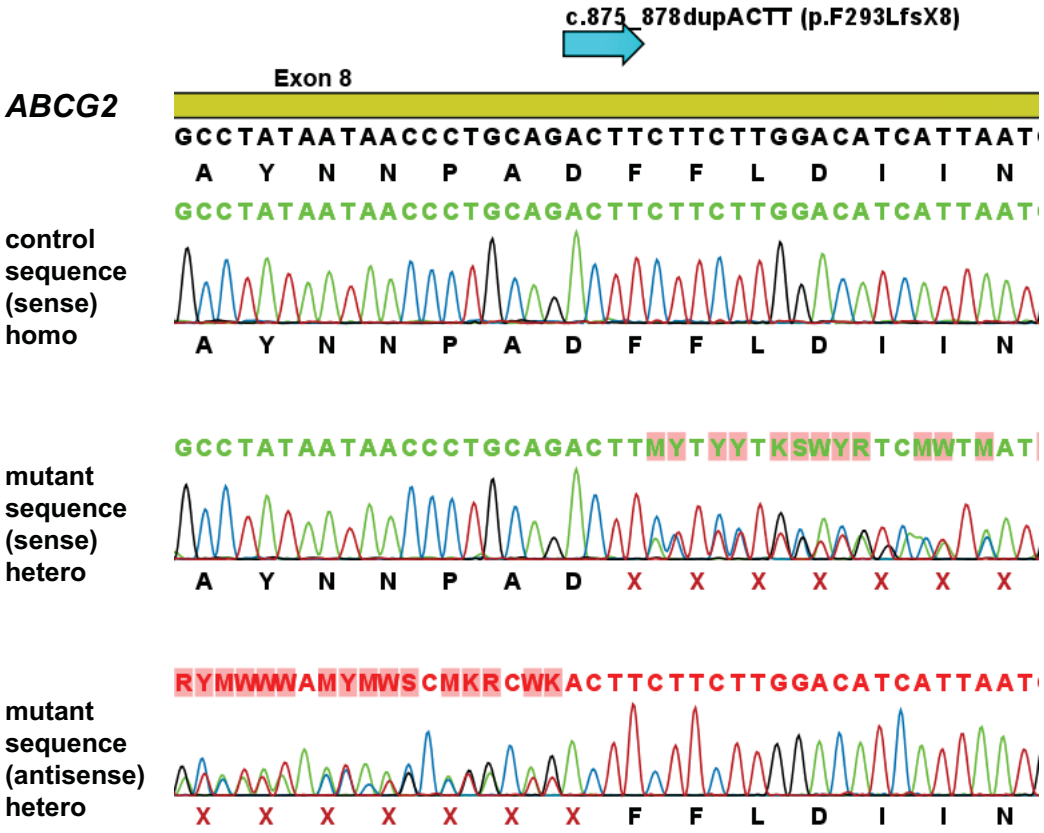


Supplementary Figure 3d: Representative sequencing traces of *ABCG2* null mutation c.706C>T (p.R236X)

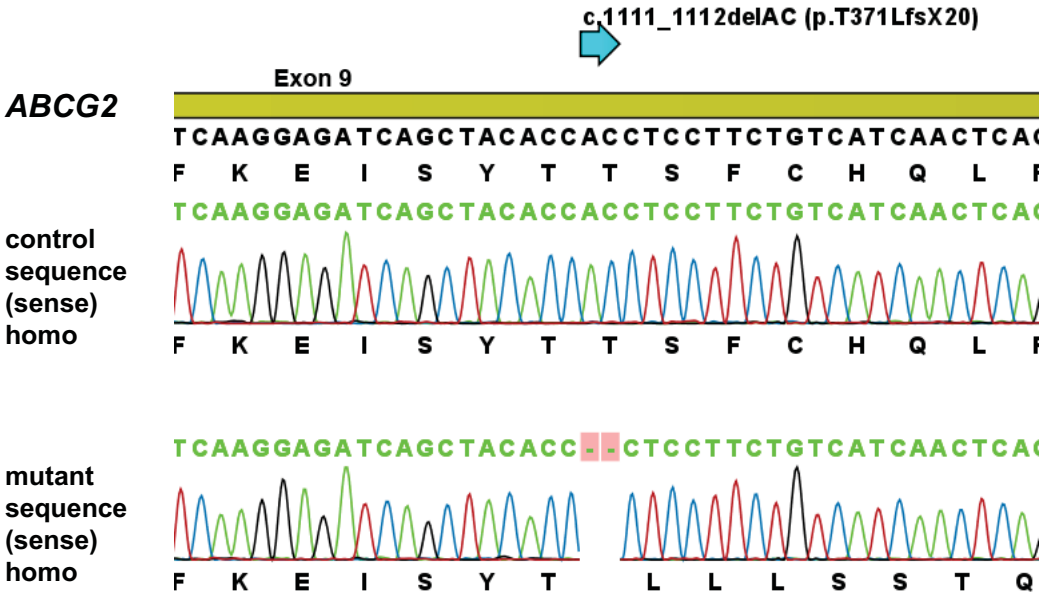
Supplementary Figure 3e: Representative sequencing traces of *ABCG2* null mutation c.730C>T (p.Q244X)



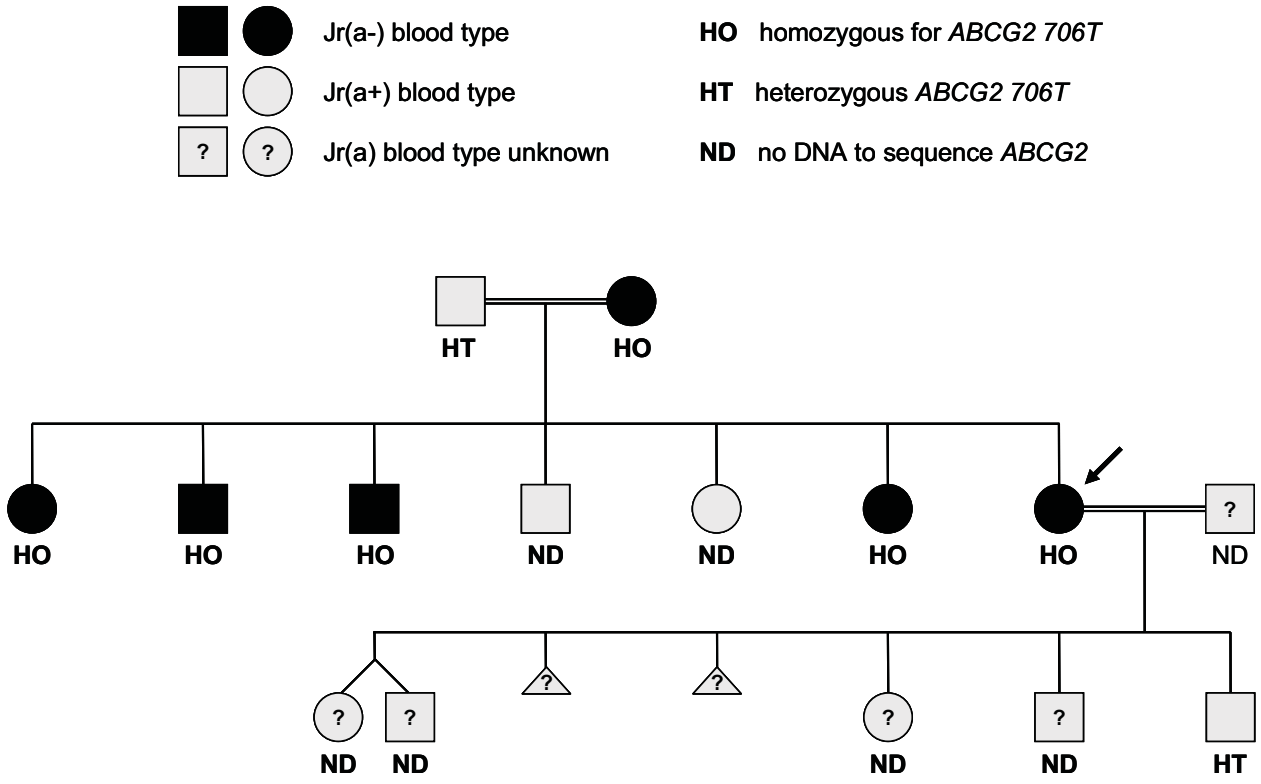
Supplementary Figure 3f: Representative sequencing traces of *ABCG2* null mutation c.791_792delTT (p.L264HfsX14)



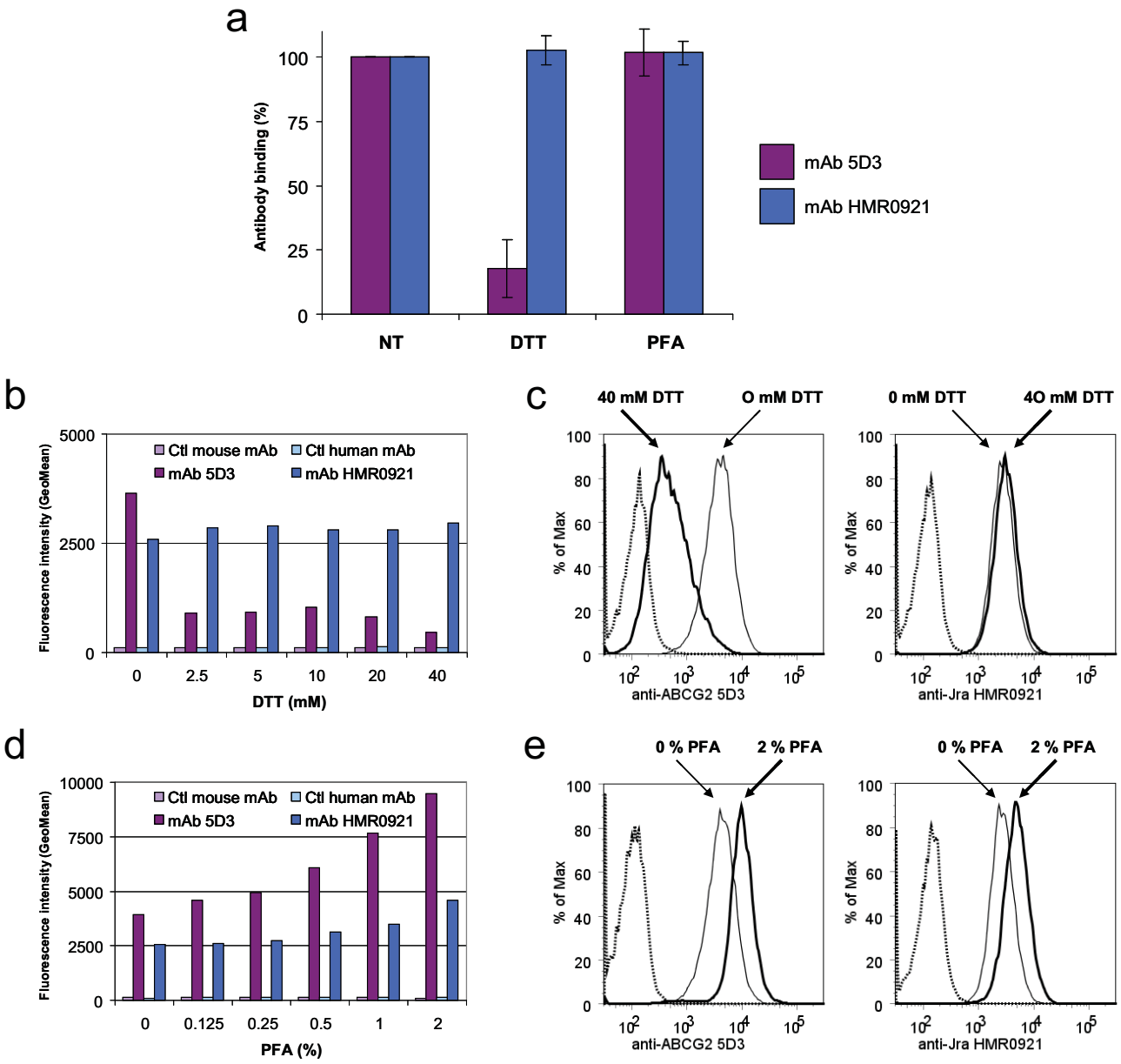
Supplementary Figure 3g: Representative sequencing traces of *ABCG2* null mutation c.875_878dupACTT (p.F293LfsX8)



Supplementary Figure 3h: Representative sequencing traces of *ABCG2* null mutation c.1111_1112delAC (p.T371LfsX20)



Supplementary Figure 4: Example of a pedigree of a Jr(a-) subject analyzed in this study. The Jr(a-) proband (indicated by an arrow) was identified after developing an anti-Jr^a induced by pregnancy.



Supplementary Figure 5: Comparison of the reactivity of monoclonal antibody 5D3 and HMR0921. **(a)** Reactivity of 5D3 (purple) or HMR0921 (blue) with human RBCs non-treated (NT), treated with 200 mM dithiothreitol (DTT) or treated with 1 % paraformaldehyde (PFA), as analyzed by flow cytometry (bars represent the means of binding, and error bars the s.d. (n=3)); notably, the epitope recognized by HMR0921 is different from the epitope recognized by 5D3 since only the latter is sensitive to the reducing agent DTT. **(b)** Reactivity of 5D3 (purple) or HMR0921 (blue), as well as control mouse mAb (clone MCP11; light purple) or control human mAb (clone T27S; light blue) with Hela cells treated with different concentrations of DTT for 10 min at 21 °C, as analyzed by flow cytometry (bars represents the geometric mean of fluorescence intensity); DTT treatment of Hela cells reduces the binding of 5D3, but not of HMR0921. **(c)** Flow cytometry profile of HeLa cells treated with 40 mM DTT (heavy line) or 0 mM DTT (light line), and labeled with 5D3 (left overlay) or HMR0921 (right overlay) as in (b). **(d)** Same as (b) but with Hela cells treated with different concentrations of PFA; PFA treatment of HeLa cells increases the binding of both 5D3 and HMR0921. **(e)** Flow cytometry profile of HeLa cells treated with 2 % PFA (heavy line) or 0 % PFA (light line), and labeled with 5D3 (left overlay) or HMR0921 (right overlay) as in (d). Flow cytometry data were acquired with a FACSCanto II flow cytometer and analyzed with FlowJo software.

	ABCG2 exon / intron			Exon 2				Exon 4	Exon 5	Exon 6
	Mutation (in NM_004827.1)			**c.19G>A**	**c.34G>A**	**c.166C>A**	**c.187_197delATATTATCGAA**	**c.376C>T**	**c.393G>T**	**c.542_543insA**
	Reported in dbSNP build 132				rs2231137			rs72552713		
	Type of mutation			**missense**	**missense**	**silent**	**frameshift**	**nonsense**	**missense**	**frameshift**
	Position on NT_016354.19			13608850C>T	13608835C>T	13608703C>A	13608682_13608672delTATAATAGCTT	13600678G>A	13600072C>A	13590653insA
Subjects	**Phenotype**	**Sex**	**Origine**							
BENA	Jr(a-)	F	Northwestern France	G/G	G/G	C/C	HO	-	G/G	-
YAN	Jr(a-)	F	Korea	G/G	G/G	**C/A**	-	HO	G/G	-
KAN	Jr(a-)	F	Korea	G/G	G/G	C/C	-	HO	G/G	-
LEV	Jr(a-)	F	Korea	G/G	G/G	C/C	-	HO	G/G	-
BER	Jr(a-)	F	Central France	G/G	G/G	C/C	-	-	G/G	HT
BOU	Jr(a-)	F	Maghreb	G/G	G/G	C/C	-	-	G/G	-
GIM	Jr(a-)	F	SW France gypsy	G/G	G/G	C/C	-	-	**T/T**	-
PAT	Jr(a-)	F	SW France gypsy	G/G	G/G	C/C	-	-	**T/T**	-
BENO	Jr(a-)	F	SW France gypsy	G/G	G/G	C/C	-	-	**T/T**	-
REI	Jr(a-)	F	SW France gypsy	G/G	G/G	C/C	-	-	**T/T**	-
REN	Jr(a-)	F	SW France gypsy	G/G	G/G	C/C	-	-	**T/T**	-
KAR	Jr(a-)	F	NE Italy gypsy	G/G	G/G	C/C	-	-	**T/T**	-
CAM	Jr(a-)	F	SW France gypsy	G/G	G/G	C/C	-	-	**G/T**	-
CHA	Jr(a-)	F	Southeastern France	G/G	G/G	C/C	-	-	G/G	-
KER	Jr(a-)	F	Turkey ?	G/G	G/G	C/C	-	-	G/G	-
CIN	Jr(a-)	F	Turkey ?	G/G	G/G	C/C	-	-	G/G	-
GUE	Jr(a-)	F	French Antilles	**G/A**	**G/A**	C/C	-	-	G/G	-
HAF	Jr(a-)	F	Pakistan	G/G	G/G	C/C	-	-	G/G	-

Supplementary Table 1 (first part): Sequencing results of *ABCG2* in the 18 Jr(a-) subjects analyzed in this study. Only the positions showing a difference with reference *ABCG2* genomic sequence NT_016354.19 in at least one subject are indicated. The mutations responsible for the Jr(a-) blood group phenotype are shown in red (HO for homozygous, HT for heterozygous).

	Exon 7			Exon 8	Exon 9	Intron 9	Intron 11	Intron 12	Intron 13		Intron 14
	c.706C>T	c.730C>T	c.791_792delTT	c.875_878dupACTT	c.1111_1112delAC	c.1195-60A>T	c.1367+20A>G	c.1492+49G>T	c.1647+40T>C	c. 1648-21C>T	c1738-46A>G
	rs2231148	rs2231153	rs2231156	rs2231157	rs2231162	rs2231164					
	nonsense	nonsense	frameshift	frameshift	frameshift	non-coding	non-coding	non-coding	non-coding	non-coding	non-coding
	13587117C>T	13587105G>A	13587032_13587031delAA	13583899_13583895dupTGAA	13582259_13582258delTG	13576199T>A	13570083T>C	13568148A>C	13566286A>G	13564503G>A	13563578T>C
Subjects											
BENA	-	-	-	-	-	**T/T**	A/A	G/G	T/T	C/C	A/A
YAN	-	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
KAN	-	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
LEV	-	-	-	-	-	**A/T**	A/A	G/G	**T/C**	C/C	A/A
BER	HT	-	-	-	-	**A/T**	A/A	G/G	**T/C**	C/C	A/A
BOU	HO	-	-	-	-	**T/T**	A/A	G/G	T/T	C/C	A/A
GIM	HO	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
PAT	HO	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
BENO	HO	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
REN	HO	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
REI	HO	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
KAR	HO	-	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
CAM	HT	-	HT	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
CHA	-	HO	-	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
KER	-	-	HO	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
CIN	-	-	HO	-	-	A/A	A/A	G/G	**C/C**	C/C	A/A
GUE	-	-	HT	HT	-	A/A	**G/G**	G/G	T/T	**T/T**	**G/G**
HAF	-	-	-	-	HO	A/A	A/A	**T/T**	T/T	C/C	**G/G**

Supplementary Table 1 (second part)

Exon	Primer	Sequence	Orientation	Position on NT_016354.19	Reference
2	ABCG2-6	CTGCTCATTGCCGCACATTT	sense	13608998-13608979	*Lee et al.*
	ABCG2-7	GCCAAAACCTGTGAGGTTCA	antisense	13608599-13608618	*Lee et al.*
3	ABCG2-8	GTTGGTTTGTGCTTGTGTTC	sense	13601635-13601616	*Lee et al.*
	ABCG2-9	GCGTTGCAAATGCTCAATAA	antisense	13601342-13601361	*Lee et al.*
4	ABCG2-1b	TGGATTCAAAGTAGCCATGAGA	sense	13600902-13600881	*Lee et al.*
	ABCG2-2b	ATTCTCCCTGCCTTTTCACA	antisense	13600501-13600520	*Lee et al.*
5	ABCG2-10	GGTTCATCATTAGCTAGAACTTTACC	sense	13600216-13600191	*Lee et al.*
	ABCG2-11	TGGAAAGCAACCATTTTTGA	antisense	13599814-13599833	*Lee et al.*
6	ABCG2-3b	TCTTACAGGACTGGCACACG	sense	13590750-13590731	*Lee et al.*
	ABCG2-4b	CCTTCCCTACATTCTTACCTGCT	antisense	13590325-13590347	*Lee et al.*
7	ABCG2-12	TCAGGCTGAACTAGAGCAAACA	sense	13587240-13587219	*Lee et al.*
	ABCG2-13	AGCACCAAATGGAACAAACA	antisense	13586834-13586853	*Lee et al.*
8	ABCG2-14	CGTGGGAAGAAGAGAGAAAGAAA	sense	13584075-13584053	*Lee et al.*
	ABCG2-15	CAAAAACACCAACAGCACTCA	antisense	13583664-13583684	*Lee et al.*
9	ABCG2-5b	GGTGTTAGGGAAGCATCCAA	sense	13582523-13582504	*Lee et al.*
	ABCG2-6b	TGAAGCAGATGATAACAGAACCA	antisense	13582111-13582133	*Lee et al.*
10	ABCG2-30	GCCAAGCCATTGAGTGTTTA	sense	13576274-13576255	*Itoda et al.*
	ABCG2-33	CTGACTCATCCTACCCTCAA	antisense	13575924-13575943	*Itoda et al.*
11	ABCG2-34	TGTGGAAAGAGTTTTGTGGGTA	sense	13570274-13570253	*Bäckström et al.*
	ABCG2-35	CCCAACCCCAGATGTAATCA	antisense	13570024-13570043	*Bäckström et al.*
12	ABCG2-20	GGTCTAGCCCTGAGGATGTG	sense	13568457-13568438	*Lee et al.*
	ABCG2-21	GAGTGCAAAATGGACAGGTG	antisense	13568055-13568074	*Lee et al.*
13	ABCG2-22	AGGGTGGTTGGAGAGTGGAT	sense	13566622-13566603	*Lee et al.*
	ABCG2-23	AGCAGAGCCCCATTTACAGA	antisense	13566211-13566230	*Lee et al.*
14	ABCG2-38	TGGTAGGGACTTGAAGAGGGTA	sense	13564567-13564546	*Bäckström et al.*
	ABCG2-39	AGCTCATGGTCAGGGAAATG	antisense	13564301-13564320	*Bäckström et al.*
15	ABCG2-26	TCTTGATTGCCAGGGAAAAT	sense	13563704-13563685	*Lee et al.*
	ABCG2-27	CGCGCACAACTCACTTTATG	antisense	13563301-13563320	*Lee et al.*
16	ABCG2-28	TGACGGATGCTAGGAATGAA	sense	13561302-13561283	*Lee et al.*
	ABCG2-29	CCCATGGTTACTGTCTGAGGA	antisense	13560873-13560893	*Lee et al.*

Supplementary Table 2: Description of the primers used in this study for amplifying and sequencing *ABCG2*.