A genetic score for the prediction of beta-thalassemia severity

Fabrice Danjou,[1] Marcella Francavilla,[1] Franco Anni,[1] Stefania Satta,[1] Franca-Rosa Demartis,[1] Lucia Perseu,[2] Matteo Manca,[2] Maria Carla Sollaino,[2] Laura Manunza,[1] Elisabetta Mereu,[1] Giuseppe Marceddu,[1] Serge Pissard,[3] Philippe Joly,[4] Isabelle Thuret,[5] Raffaella Origa,[1] Joseph Borg,[6] Gian Luca Forni,[7] Antonio Piga,[8] Maria Eliana Lai,[2] Catherine Badens,[9] Paolo Moi,[1,2] and Renzo Galanello[1,2]

[1]Dipartimento di Sanità Pubblica, Medicina Clinica e Molecolare, Università degli Studi di Cagliari, Italy; [2]Ospedale Regionale per le Microcitemie, ASL8, Cagliari, Italy; [3]Laboratoire de Génétique, Hôpital Henri Mondor, Créteil, France; [4]Laboratoire de Biochimie et Biologie Moléculaire, Hôpital Edouard Herriot, Lyon, France; [5]French reference center for Thalassemia and Service d'Onco-Hématologie Pédiatrique, Assistance Publique Hôpitaux de Marseille, France; [6]Department of Applied Biomedical Science, Faculty of Health Sciences and Laboratory of Molecular Genetics, Department of Physiology and Biochemistry, University of Malta, Msida, Malta; [7]Ematologia - Centro della Microcitemia ed Anemie Congenite, Ospedale Galliera, Genova, Italy; [8]Università di Torino, Italy; and [9]French reference center for Thalassemia and Department of Genetics, Assistance Publique Hopitaux de Marseille and Aix-Marseille University, France

 doi:10.3324/haematol.2014.113886
Manuscript received on July 16, 2014. Manuscript accepted on December 2, 2014.
Correspondence: fabdanjou@yahoo.fr

Supplementary data

Supplementary Methods

Single nucleotide polymorphisms selection:

rs7482144 (HBG2:g.-158C>T), is an XmnI restriction site situated at 158 bp upstream from *HBG2* on chromosome 11, at 5276169 bp. This variant has been for long associated with increased production of HBF and beta-thalassemia as well as sickle cell disease phenotype amelioration.[1–3]

rs9399137: located within the *HBS1L-MYB* intergenic region on chromosome 6, at position 135419018 bp, this variant has been repeatedly associated with HBF levels in different populations.[4–6] It is also in complete linkage disequilibrium with a 3 bp deletion (rs66650371) located in close proximity to four erythropoiesis-related transcription factors binding sites, functionally associated with HBF levels.[7]

rs1427407: within intron 2 of *BCL11A* gene on chromosome 2, is situated at 60718043 bp. As reported by different genome-wide association studies, rs1427407 is strongly associated with HBF levels. It is also in high linkage disequilibrium with rs766432 and rs4671393 (r2=0.96 and r2=0.92 respectively) in the CEU samples from the 1000 Genomes Project, two variants for which effects on HBF levels are also well-documented.[4,8,9]

rs10189857: within intron 2 of *BCL11A* gene on chromosome 2, is situated at 60713235 bp. This variant, together with rs7599488 (in full linkage disequilibrium in the CEU samples from the 1000 Genomes Project), are documented to have an additive independent effect from rs1427407 on HBF levels.[6,10]

Thalassemia Severity Score calculation:

The TSS is calculated as follow:

1) linear predictor score (LPS) is obtained by summing the products of each genotype marker's natural logarithm of Hazard Ratio (from Table 3), for each individual (reference categories have HR = 1):
   $$LPS = \sum \ln\left(HR_{Genotype\ Value}\right)$$
2) The TSS is then obtained from linear predictors applying the following transformation:
   $$TSS = \frac{(9.5 + LPS)^6}{80000}$$

For example, to calculate the TSS of a female patient homozygous for the HBB:c.118C>T mutation, with two defects at the alpha-globin genes cluster, genotypes T/T for rs7482144, T/T for rs9399137, T/T for rs1427407, and A/A for rs10189857, one should first obtain the hazards ratio values from Table 3 for each category of each predictor and compute: LPS = $\ln(1)_{(\text{female})}$ + $\ln(1)_{(\text{severe/severe})}$ + $\ln(0.252)_{(\text{two alpha defects})}$ + $\ln(0.052)_{(\text{rs7482144 T/T})}$ + $\ln(1)_{(\text{rs9399137 T/T})}$ + $\ln(0.362)_{(\text{rs1427407 T/T})}$ + $\ln(1)_{(\text{rs10189857 A/A})}$ = −5.35

Then, the LPS are transformed as follow to obtain the TSS: TSS = $(9.5+(-5.35))^6$ / 80000 = 0.06

References for Supplementary Methods

1. Labie D, Dunda-Belkhodja O, Rouabhi F, et al. The -158 site 5' to the G gamma gene and G gamma expression. *Blood*. 1985;66(6):1463–1465.
2. Gilman JG, Huisman TH. DNA sequence variation associated with elevated fetal G gamma globin production. *Blood*. 1985;66(4):783–787.
3. Nguyen TK, Joly P, Bardel C, et al. The XmnI (G)gamma polymorphism influences hemoglobin F synthesis contrary to BCL11A and HBS1L-MYB SNPs in a cohort of 57 beta-thalassemia intermedia patients. *Blood Cells. Mol. Dis.* 2010;45(2):124–127.
4. Menzel S, Garner C, Gut I, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat. Genet.* 2007;39(10):1197–1199.
5. Lettre G, Sankaran VG, Bezerra MAC, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc. Natl. Acad. Sci. U. S. A.* 2008;105(33):11869–11874.
6. Galarneau G, Palmer CD, Sankaran VG, et al. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nat. Genet.* 2010;42(12):1049–1051.
7. Farrell JJ, Sherva RM, Chen Z-Y, et al. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. *Blood*. 2011;117(18):4935–4945.
8. Solovieff N, Milton J, Hartley S, et al. Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. *Blood*. 2010;115(9):1815-1822.
9. Bhatnagar P, Purvis S, Barron-Casella E, et al. Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. *J. Hum. Genet.* 2011;56(4):316–323.
10. Bauer DE, Kamran SC, Lessard S, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science*. 2013;342(6155):253–257.

Supplementary Tables

Supplementary Table 1 Genetic modulators of beta-thalassemia phenotype: loci characteristics.

Loci		Position (GRCh37.p10)	classification [model codification]		Genotype (international name)	Genotype (conventional name)	Nb of cases	Frequencies
HBB gene defects*		chr11:5246696-5248301	mild/mild	b+_mild/b+_mild	HBB:c.92+6T>C/HBB:c.92+6T>C	IVS-I-6(T->C)/IVS-I-6(T->C)	12	1.40%
				b+_silent/b0_severe	HBB:c.-151C>T/HBB:c.118C>T	-101(C->T)/Codon39(C->T)	3	0.35%
				b+_mild/b+_mild	HBB:c.-79A>G/HBB:c.-79A>G	-29(A->G)/-29(A->G)	1	0.12%
				b+_silent/b0_severe	HBB:c.82G>T/HBB:c.118C>T	HbKnossos/Codon39(C->T)	1	0.12%
				b0_mild/b0_mild	HBB:c.118C>T+HBG1:c.-249C>T/HBB:c.118C>T+HBG1:c.-249C>T	deltabeta0/deltabeta0	1	0.12%
				b+_mild/b0_mild	NG_000007.3:g.64336_77738del13403/HBB:c.92+6T>C	deltabeta0/IVS-I-6(T->C)	1	0.12%
				b+_mild/b0_mild	NG_000007.3:g.70060_71452del1393/HBB:c.75T>A	1.393bpdeletion/Codon24(T->A)	1	0.12%
			mild/severe	b+_mild/b0_severe	HBB:c.92+6T>C/HBB:c.118C>T	IVS-I-6(T->C)/Codon39(C->T)	35	4.07%
				b+_mild/b+_severe	HBB:c.93-21G>A/HBB:c.92+6T>C	IVS-I-110(G->A)/IVS-I-6(T->C)	13	1.51%
				b+_mild/b0_severe	HBB:c.92+6T>C/HBB:c.92+1G>T,A	IVS-I-6(T->C)/IVS-I-1(G->T.A)	8	0.93%
				b+_mild/b0_severe	HBB:c.-137C>A,G,T/HBB:c.118C>T	-87(C->A.G.T)/Codon39(C->T)	6	0.70%
				b+_mild/b0_severe	HBB:c.92+6T>C/HBB:c.315+1G>C,A,T	IVS-I-6(T->C)/IVS-II-1(G->C.A.T)	5	0.58%
				b0_mild/b0_severe	HBB:c.118C>T+HBG1:c.-249C>T/HBB:c.118C>T	deltabeta0/Codon39(C->T)	5	0.58%
				b+_mild/b0_severe	HBB:c.79G>A/HBB:c.126_129delCTTT	HbE/Codons41/42(-TTCT)	2	0.23%
				b+_mild/b0_severe	HBB:c.92+6T>C/HBB:c.20delA	IVS-I-6(T->C)/Codon6(-A)	2	0.23%
				b+_mild/b+_severe	HBB:c.92+6T>C/HBB:c.316-106C>G	IVS-I-6(T->C)/IVS-II-745(C->G)	2	0.23%
				b+_mild/b+_severe	HBB:c.79G>A/HBB:c.316-197C>T	HbE/IVS-II-654(C->T)	1	0.12%
				b+_mild/b0_severe	HBB:c.92+1G>T,A/HBB:c.-137C>A,G,T	IVS-I-1(G->T.A)/-87(C->A.G.T)	1	0.12%
				b+_mild/b0_severe	HBB:c.92+2T>G,A,C/HBB:c.-137C>A,G,T	IVS-I-2(T->G.A.C)/-87(C->A.G.T)	1	0.12%
				b+_mild/b0_severe	HBB:c.92+6T>C/HBB:c.135delC	IVS-I-6(T->C)/Codon44(-C)	1	0.12%
				b+_mild/b0_severe	HBB:c.92+6T>C/HBB:c.47G>A,HBB:c.48G>A,HBB:c.46delT	IVS-I-6(T->C)/Codon15(G->A).Codon15(-T)	1	0.12%
				b+_mild/b+_severe	HBB:c.93-21G>A/HBB:c.-137C>A,G,T	IVS-I-110(G->A)/-87(C->A.G.T)	1	0.12%
				b+_mild/b+_severe	HBB:c.93-21G>A/HBB:c.-138C>G,T,A	IVS-I-110(G->A)/-88(C>G.T.A)	1	0.12%
				b+_mild/b+_severe	HBB:c.93-21G>A/HBB:c.-79A>G	IVS-I-110(G->A)/-29(A->G)	1	0.12%
				b+_mild/b+_severe	NG_000007.3:g.63290_70702del/HBB:c.92+6T>C	HbLepore-Hollandia/IVS-I-6(T->C)	1	0.12%
			severe/severe	b0_severe/b0_severe	HBB:c.118C>T/HBB:c.118C>T	Codon39(C->T)/Codon39(C->T)	563	65.54%
				b+_severe/b0_severe	HBB:c.93-21G>A/HBB:c.118C>T	IVS-I-110(G->A)/Codon39(C->T)	48	5.59%
				b0_severe/b0_severe	HBB:c.20delA/HBB:c.118C>T	Codon6(-A)/Codon39(C->T)	30	3.49%
				b+_severe/b+_severe	HBB:c.93-21G>A/HBB:c.93-21G>A	IVS-I-110(G->A)/IVS-I-110(G->A)	23	2.68%
				b0_severe/b0_severe	HBB:c.92+1G>T,A/HBB:c.118C>T	IVS-I-1(G->T.A)/Codon39(C->T)	14	1.63%
				b+_severe/b0_severe	HBB:c.93-21G>A/HBB:c.92+1G>T,A	IVS-I-110(G->A)/IVS-I-1(G->T.A)	10	1.16%
				b0_severe/b0_severe	HBB:c.230delC/HBB:c.118C>T	Codon76(-C)/Codon39(C->T)	8	0.93%
				b0_severe/b0_severe	HBB:c.92+1G>T,A/HBB:c.92+1G>T,A	IVS-I-1(G->T.A)/IVS-I-1(G->T.A)	6	0.70%
				b+_severe/b+_severe	HBB:c.93-21G>A/HBB:c.316-106C>G	IVS-I-110(G->A)/IVS-II-745(C->G)	6	0.70%
				b+_severe/b0_severe	HBB:c.316-106C>G/HBB:c.118C>T	IVS-II-745(C->G)/Codon39(C->T)	5	0.58%
				b0_severe/b0_severe	HBB:c.20delA/HBB:c.20delA	Codon6(-A)/Codon6(-A)	5	0.58%
				b+_severe/b0_severe	HBB:c.93-21G>A/HBB:c.20delA	IVS-I-110(G->A)/Codon6(-A)	4	0.47%
				b0_severe/b0_severe	HBB:c.315+1G>C,A,T/HBB:c.118C>T	IVS-II-1(G->C.A.T)/Codon39(C->T)	3	0.35%
				b0_severe/b0_severe	HBB:c.92+1G>T,A/HBB:c.315+1G>C,A,T	IVS-I-1(G->T.A)/IVS-II-1(G->C.A.T)	2	0.23%
				b+_severe/b0_severe	HBB:c.92+1G>T,A/HBB:c.316-106C>G	IVS-I-1(G->T.A)/IVS-II-745(C->G)	2	0.23%
				b0_severe/b0_severe	HBB:c.92+2T>G,A,C/HBB:c.20delA	IVS-I-2(T->G.A.C)/Codon6(-A)	2	0.23%
				b0_severe/b0_severe	HBB:c.92+2T>G,A,C/HBB:c.92+2T>G,A,C	IVS-I-2(T->G.A.C)/IVS-I-2(T->G.A.C)	2	0.23%
				b+_severe/b+_severe	HBB:c.93-21G>A/HBB:c.316-197C>T	IVS-I-110(G->A)/IVS-II-654(C->T)	2	0.23%
				b0_severe/b0_severe	HBB:c.112delT/HBB:c.112delT	Codons36/37(-T)/Codons36/37(-T)	1	0.12%
				b0_severe/b0_severe	HBB:c.315+1G>C,A,T/HBB:c.25_26delAA	IVS-II-1(G->C.A.T)/Codon8(-AA)	1	0.12%
				b0_severe/b0_severe	HBB:c.92+1G>T,A/HBB:c.20delA	IVS-I-1(G->T.A)/Codon6(-A)	1	0.12%
				b0_severe/b0_severe	HBB:c.25_26delAA/HBB:c.25_26delAA	Codon8(-AA)/Codon8(-AA)	1	0.12%
				b0_severe/b0_severe	HBB:c.315+1G>C,A,T/HBB:c.135delC	IVS-II-1(G->C.A.T)/Codon44(-C)	1	0.12%
				b+_severe/b+_severe	HBB:c.316-106C>G/HBB:c.316-106C>G	IVS-II-745(C->G)/IVS-II-745(C->G)	1	0.12%
				b+_severe/b0_severe	HBB:c.316-197C>T/HBB:c.118C>T	IVS-II-654(C->T)/Codon39(C->T)	1	0.12%
				b+_severe/b0_severe	HBB:c.316-197C>T/HBB:c.126_129delCTTT	IVS-II-654(C->T)/Codons41/42(-TTCT)	1	0.12%
				b0_severe/b0_severe	HBB:c.52A>T/HBB:c.126_129delCTTT	Codon17(A->T)/Codons41/42(-TTCT)	1	0.12%
				b+_severe/b0_severe	HBB:c.92+1G>T,A/HBB:c.316-3C>A	IVS-I-1(G->T.A)/IVS-II-848(C->A)	1	0.12%
				b0_severe/b0_severe	HBB:c.92+2T>G,A,C/HBB:c.118C>T	IVS-I-2(T->G.A.C)/Codon39(C->T)	1	0.12%
				b+_severe/b+_severe	HBB:c.92+5G>C/HBB:c.316-3C>A	IVS-I-5(G->C)/IVS-II-848(C->A)	1	0.12%
				b+_severe/b0_severe	HBB:c.93-21G>A/HBB:c.135delC	IVS-I-110(G->A)/Codon44(-C)	1	0.12%
				b+_severe/b0_severe	HBB:c.93-21G>A/HBB:c.315+1G>C,A,T	IVS-I-110(G->A)/IVS-II-1(G->C.A.T)	1	0.12%
				b+_severe/b+_severe	HBB:c.93-21G>A/HBB:c.92+5G>C	IVS-I-110(G->A)/IVS-I-5(G->C)	1	0.12%
				b0_severe/b0_severe	HBB:c.93G>C/HBB:c.93G>C	Codon30(AGG->AGC)[IVS-I-130(+1)]/Codon30(AGG->AGC)[IVS-I-130(+1)]	1	0.12%
				b+_severe/b0_severe	NG_000007.3:g.63632_71046del/HBB:c.118C>T	HbLepore-Boston-Washington/Codon39(C->T)	1	0.12%
			Total				859	100.00%
HBA genes defects**		chr16:226679-227520	No defect			aa/aa	560	65.19%
			1 defect		NG_000006.1:g.34164_37967del3804	-a3.7type/aa	192	22.35%
					HBA2:c.2T>C	aNcoIa/aa	31	3.61%
					HBA2:c.95+2_95+6delTGAGG	aHphIa/aa	7	0.81%
			2 defects		NG_000006.1:g.34164_37967del3804/HBA2:c.2T>C	-a3.7type/aNcoIa	10	1.16%
					NG_000006.1:g.34164_37967del3804/HBA2:c.95+2_95+6delTGAGG or HbVarID=1079/HBA2:c.95+2_95+6delTGAGG	-a3.7type/aHphIa or –a4.2/aHphIa	2	0.23%
					NG_000006.1:g.34164_37967del3804/NG_000006.1:g.34164_37967del3804 or NG_000006.1:g.34164_37967del3804/HbVarID=1079 or NG_000006.1:g.15164_37864del22701	-a3.7type/-a3.7type or -a3.7type/-a4.2 or –(alpha)20.5/aa	57	6.64%
			Total				859	100.00%
HBG2	rs7482144 (HBG2:g.-158C>T)	chr11:5276169	C/C (-/-)				812	94.53%
			C/T (-/+)				41	4.77%
			T/T (+/+)				6	0.70%
			Total				859	100.00%
HBS1L-MYB intergenic region	rs9399137	chr6:135419018	T/T				556	64.73%
			C/T				292	33.99%
			C/C				11	1.28%
			Total				859	100.00%
BCL11A	rs1427407	Chr2:60713235	G/G				535	62.28%
			G/T				284	33.06%
			T/T				40	4.66%
			Total				859	100.00%
	rs10189857	Chr2:60718043	A/G				418	48.66%
			A/A				310	36.09%
			G/G				131	15.25%
			Total				859	100.00%

* deltabeta0 mutation reported here result in a mild phenotype and are of Sicilian type [NG_000007.3:g.64336_77738del13403] or Sardinian type, i.e.: association of Codon39(C->T) mutation [HBB:c.118C>T] with an HPFH Agamma nt-196(C>T) mutation [HBG1:c.-249C>T].

** Were investigated: 3.7 kb rightward type I, II and III, 4.2 kb leftward, Mediterranean type I and 20.5 kb large deletions as well as HphI small deletion and NcoI polymorphism.

<u>Supplementary Table 2</u> Kaplan-Meier survival analysis of transfusion free survival prediction by TSS in training and testing sets.

TSS		Means and Medians for Survival Time								p-value (Breslow test)
		Mean				Median				
		Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval		
				Lower Bound	Upper Bound			Lower Bound	Upper Bound	
Training set	Low severity [TSS≤3]	298.988	30.288	239.623	358.353	57.016	7.802	41.725	72.308	$1.03x10^{-42}$
	Mild severity [3<TSS≤5]	98.907	18.151	63.331	134.482	18.623	1.845	15.007	22.238	
	High severity [5<TSS≤7]	42.293	10.374	21.961	62.625	8.984	0.604	7.800	10.167	
	Very high severity [7>TSS]	16.325	4.180	8.133	24.518	5.967	0.481	5.024	6.911	
	Overall	109.197	10.106	89.389	129.005	13.607	1.034	11.581	15.633	
Testing set	Low severity [TSS≤3]	343.635	60.085	225.869	461.401	79.967	34.610	12.131	147.803	$3.20x10^{-10}$
	Mild severity [3<TSS≤5]	133.779	30.338	74.316	193.243	23.213	5.687	12.066	34.360	
	High severity [5<TSS≤7]	67.940	18.914	30.868	105.012	11.672	2.132	7.494	15.851	
	Very high severity [7>TSS]	22.839	9.395	4.425	41.254	6.033	0.770	4.523	7.543	
	Overall	142.190	19.096	104.762	179.618	13.705	1.776	10.225	17.185	