

Molecular analysis of 42 patients with congenital dyserythropoietic anemia type II: new mutations in the *SEC23B* gene and a search for a genotype-phenotype relationship

Achille Iolascon,^{1,2} Roberta Russo,^{1,2} Maria Rosaria Esposito,^{1,2} Roberta Asci,¹ Carmelo Piscopo,^{1,2} Silverio Perrotta,³ Madeleine Fénéant-Thibault,⁴ Loïc Garçon,⁵ and Jean Delaunay⁶

¹CEINGE Biotecnologie Avanzate, Napoli, Italy; ²Department of Biochemistry and Medical Biotechnologies, University Federico II of Naples, Italy; ³Department of Pediatrics, Second University of Naples, Italy; ⁴Service de Biochimie, Hôpital de Bicêtre, Le Kremlin-Bicêtre, France; ⁵Service d'Hématologie Générale, Hôtel-Dieu, Paris, France, and ⁶INSERM U 779, Faculté de Médecine Paris-Sud, Univ Paris-Sud, Le Kremlin-Bicêtre, France

Citation: Iolascon A, Russo R, Esposito MR, Asci R, Piscopo C, Perrotta S, Fénéant-Thibault M, Garçon L, and Delaunay J. Molecular analysis of 42 patients with congenital dyserythropoietic anemia type II: new mutations in the *SEC23B* gene and a search for a genotype-phenotype relationship. *Haematologica* 2010;95:708-715. doi:10.3324/haematol.2009.014985

Online Supplementary Table S1. Topographical distribution of *SEC23B* mutations in CDA II patients.

Patient	Exon	Nucleotide change	Effect on protein	Mutation status
Nonsense mutations				
2; 13; 15	3	c. 235 C>T	R79X *	Comp het
8	6	c. 649 C>T	R217X *,†	Het
18	7	c. 790 C>T	R264X *	Comp het
35a, 35b	10	c. 1201 C>T	R401X *	Comp het
15; 31	14	c. 1648 C>T	R550X	Comp het
Frameshift mutations				
6	9	c. 1063 (delG)	D355IfsX8 *	Comp het
20	17	c. 1962-64 (delT)	T654TfsX13	Comp het
Missense mutations				
2; 11; 13; 16; 17; 22; 23; 27; 35a, 35b; 37	2	c. 40 C>T	R14W *,†	Comp het
40	2	c. 40 C>T	R14W	Het
19; 39	2	c. 53 G>A	R18H *	Comp het
5; 7; 12; 14; 24; 25; 30; 32; 33; 34; 36; 38	4	c. 325 G>A	E109K *,†	Hom
22; 28	4	c. 325 G>A	E109K	Comp het
10	8	c. 953 T>C	I318T *	Hom
16; 28	8	c. 953 T>C	I318T	Comp het
6	10	c.1157 A>G	Q386R *	Comp het
1a, 1b; 3	12	c.1385 A>G	Y462C *	Hom
20; 21; 26; 31	13	c. 1489 C>T	R497C *,†	Comp het
29	13	c. 1489 C>T	R497C	Hom
17	13	c. 1508 G>A	R503Q	Comp het
19; 31; 39	14	c. 1571 C>T	A524V *	Comp het
27	14	c. 1589 G>A	R530Q	Comp het
15	15	c. 1685 A>G	Y562C	Comp het
4	17	c. 1968 T>G	F656L	Het
9	18	c. 2101 C>T	R701C†	Het
21; 37	18	c. 2101 C>T	R701C	Comp het
18	19	c. 2166 A>C	K723Q	Comp het
Splicing mutations				
23	6	c. 689 +1 G>A *,†		Comp het

Hom, mutations in the homozygous state; comp het, mutations in the compound heterozygous state; del, deletions; ins, insertions. In bold, novel mutations, not contained in references 13* and 14†.

Q15437	STLGIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q9B95	STLGIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q9S52	STLGIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
AZANAL	STLGIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q3382	STLGIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q8JH8	STLGIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q6A59	TLTAEIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q6D30	TLTAEIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q4TBV7	TLTAEIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	507
Q9HF9	STVAIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	509
Q9UC21	STIIVGIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	562
ALCR87	STIIVGIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	509
Q17150	STIIVGIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	511
P15303	HEVAIIEVNVNHRHAIIP009-----RGAIQFVTVVQHSSTGRIVTTTAAHWAD	511
324		
Q15437	VQGGRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q9B95	VQGGRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q9S52	VQGGRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
AZANAL	AGQGLRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q3382	VQGGRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q8JH8	AGTQRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q6A59	ASQIQRITAEFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q6D30	ASQIQRITAEFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q4TBV7	AGQGLRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q9HF9	ASIVYHRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	566
Q9UC21	AFVGGRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	622
ALCR87	PTGQ-----PALAQFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	567
Q17150	ATMELRIKTAAFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	570
P15303	LRFQFDALAAIFDQAAAVLMAIIVYFAEIE-----EGPDVLRMLDQLGLLCKQKFG	570
328		
Q15437	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	685
Q9B95	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	685
Q9S52	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	684
AZANAL	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	685
Q3382	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	685
Q8JH8	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	685
Q6A59	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	684
Q6D30	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	684
Q4TBV7	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	684
Q9HF9	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	682
Q9UC21	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	687
ALCR87	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	687
Q17150	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	689
P15303	STYFNRG-----PEFVLLDSILADRIILLMOTTFIVTLGRTTIAQRKAGYQGRHEIYERFKH	689
332		
Q15437	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	743
Q9B95	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	743
Q9S52	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	742
AZANAL	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	743
Q3382	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	743
Q8JH8	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	743
Q6A59	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	742
Q6D30	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	742
Q4TBV7	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	722
Q9HF9	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	745
Q9UC21	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	797
ALCR87	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	745
Q17150	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	747
P15303	LQAAFLDQAGEILQAAEFPHRYINTEHGGSCAQLLLEFVHFSQTRHLLTAMQGGTAAP	743
336		
Q15437	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q9B95	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q9S52	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
AZANAL	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q3382	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q8JH8	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q6A59	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q6D30	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q4TBV7	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q9HF9	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q9UC21	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
ALCR87	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
Q17150	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
P15303	MATYLETIQGNEEIDGVEFIRHVVPSRILEATNMAVFLACLITLIERE	DLFPV 54
340		
Q15437	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q9B95	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q9S52	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
AZANAL	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q3382	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q8JH8	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q6A59	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q6D30	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q4TBV7	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q9HF9	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
Q9UC21	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
ALCR87	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	115
Q17150	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	114
P15303	QTEFVLCRFTCKAVLFLCQVDYFAKLMACNFCTGRHCFPRATVGIIEVNVGPAELMVF	107
344		
Q15437	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q9B95	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q9S52	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
AZANAL	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q3382	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q8JH8	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q6A59	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 336
Q6D30	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q4TBV7	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 338
Q9HF9	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 339
Q9UC21	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 339
ALCR87	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 337
Q17150	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 341
P15303	EGTFPHYGARINLFTGGPFTGGPFGVGGVGLKLEPIRBMNH	EEDHARFNKKAATYERHIA 333
348		
Q15437	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q9B95	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q9S52	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
AZANAL	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q3382	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q8JH8	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q6A59	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 396
Q6D30	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q4TBV7	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 458
Q9HF9	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 399
Q9UC21	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 452
ALCR87	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 397
Q17150	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 401
P15303	MRTAAMRHCIDITVACALDQVGLLEKRCANLGGVWMDSEFTLSEFT	GTGRIFKDFKH 393

Online Supplementary Figure 15. Multiple DNA sequence alignment of the SEC23B protein. Fourteen species were aligned using Database UniProtKB release 15.0 of Mar-25-2009; program NCBI BLASTP 2.2.17 [Aug-26-2007] on services.uniprot.org (<http://services.uniprot.org/blast/>). The species' name, protein number and portions of the sequence are shown for each mutation discovered. Mutated amino acids are highlighted showing the conservation among species. *residues identical in all sequences in the alignment; . conserved substitutions observed. .semi-conserved substitutions observed


```

Q15437 GDFRMAFGATLDVKTRELKIAGAI426GPCVSLNVK433GPC-VSENELGVGGTSQWKICGLDPT 456 Homo sapiens
Q9BE95 GDFRMAFGATLDVKTRELKIAGAI426GPCVSLNVK433GPC-VSENELGVGGTSQWKICGLDPT 456 Macaca fascicularis
Q5R5G2 GDFRMAFGATLDVKTRELKIAGAI426GPCVSLNVK433GPC-VSENELGVGGTSQWKICGLDPT 456 Pongo abelii
A2AMA1 GDFRMAFGATLDVKTRELKIAGAI426GPCVSLNVK433GPC-VSENELGVGGTSQWKICGLDPS 456 Mus musculus
Q3S2N2 GNFRMAFGATLEVKTRELKVAGAI426GPCVSLNVK433GPC-VSENELGVGGTSQWKICGLDPT 456 Bos taurus
Q8JHM0 GEFRMAFGANLEVKTRELKISGAI426GPCVSLNVK433GPC-ISENELGVGGTSQWKICSLDPT 456 Rana ridibunda
Q6A298 GEFRMAFGATLEVKTRELKVS426GAI433GPCVSLN433TGK433GPC-VSDNEMGVGGTSQWKICSLSPA 455 Danio rerio
Q6DJE0 GSFRMAFGANLDVKTRELKISG426TIG433GPCVSLNVK433GPC-VSENELGVGGTSQWKICALDPT 456 Xenopus laevis
Q4TBV7 GSFRMAFGAALLEVKTREIKV426S433GAI433GPCVSL433CAK433GPC-VSEN-LGTAEPVSGRCVAWTPH 516 Tetraodon nigroviridis
Q9NPN9 NDLKMAFNATLEVRCSRELKISG426GIG433SCVSLNVK433GPC-VSDVEIGMGNTVQWKLC433TLNPS 458 Drosophila melanogaster
Q9U2Z1 GNLEKMGFNATHEVKGAGL426KIEG433VLGCCASGNVFNAN-VSDQENGIGGTCQWKFGAISPR 511 Caenorhabditis elegans
A1CRM7 DNLLMGFNASLEVLTTKELK426V433TGLIGHAVSLN433KKSS-VGETECGIGNTCAWKMGIDPA 456 Aspergillus clavatus
Q171Z0 GDLKMAFN426OTLEIKCSRELKIEG433GIG433SCVSLNVK433AS-VSDSEIGMGNTAQWKLCTHTPN 460 Aedes aegypti
P15303 GYLKMAFN426GNNAVKTSD433KLKVQGLIGHASAVK433KTDANNISESEIGIGATSTWKNASLS433SPY 453 Saccharomyces cerevisiae
. : * . * . : : : * : * : * . : . . : : : * . . : *

```

```

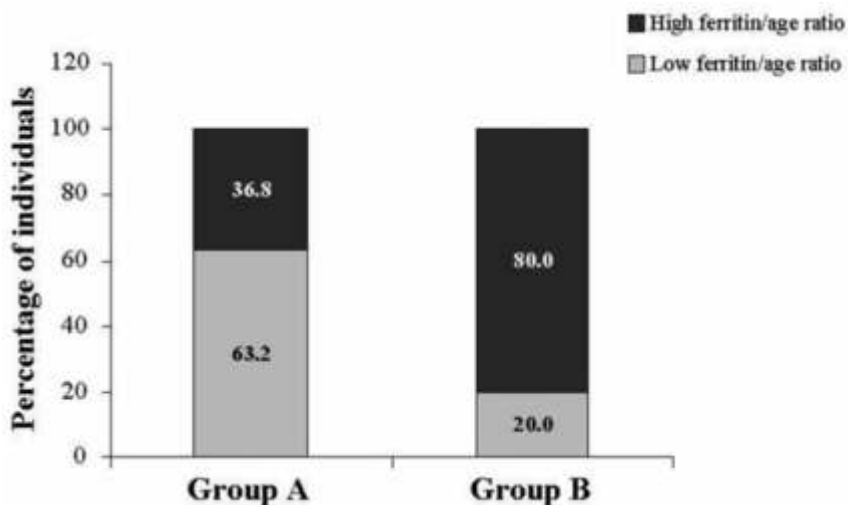
Q15437 STLGIYFEVVNQHNTP489IPQGG-----RGAIQFVIT489YQHSSTQRRIRVTTIARNWAD 507
Q9BE95 STLGIYFEVVNQHNAP489IPQGG-----RGAIQFVIT489YQHSSTQRRIRVTTIARNWAD 507
Q5R5G2 STLGIYFEVVNQHNTP489IPQGG-----RGAIQFVIT489YQHSSTQRRIRVTTIARNWAD 507
A2AMA1 STLGIYFEVVNQHNAPV489PQGG-----RGAIQFVIT489YQHSSTQKRIRVTTIARNWAD 507
Q3S2N2 TTLGIYFEVVNQHNAP489IPQGG-----RGAIQFVIT489YQHSSTQRRIRVTTIARNWAD 507
Q8JHM0 TTLGIYFEVVNQHNAP489IPQGG-----RGAIQFVIT489YQHSSTQKRIRVTTIARNWAD 507
Q6A298 TTLAHYFEVVNQHNAPV489PQGG-----RQVVQFVIT489YQHSNTQRRIRVTTIARNWAD 506
Q6DJE0 TTLAIYFEVVNQHNAP489IPQGG-----RQAVQFLT489YQHSSTQKRIRVTTIARNWAD 507
Q4TBV7 TTLALYFEVVNQHNAP489IPQGG-----RQAVQFVIT489YQHSSTQRRIRVTTIARNWAD 567
Q9NPN9 STVAYFFE489VVNQHAAP489IPQGG-----RGCIFIT489YQHPSPGQRRIRVTTIARNWAD 509
Q9U2Z1 TTIGVVFEIAAQHGS489AI489PQGG-----RQHVQFVIT489YQHADGRKRIRVTTIARNWAD 562
A1CRM7 SBYGVVFEIANQGGPAAVQ489PGPQ-----RGMHQLT489YQHSSTQRRIRVTTIARNWAD 509
Q171Z0 STMAFFFEVANQHAAP489IPQGG-----RQCLQFIT489YQHSSTQRRIRVTTIARNWAD 511
P15303 HSYAIFFEIANTAANSNP489MMSAPGSADRP489HLAYTQFIT489YQHSSTG489TN489IRVTTIARNWAD--QL 511
: . * : . . . . . * : * * * . . : : * * * * . .

```

Absolute allele and genotype frequencies

	rs41309927 G>A V426I			rs17807673 C>T P433L			rs2273526 C>G H489Q								
	Genotype n(%)			Allele n(%)		Genotype n(%)			Allele n(%)		Genotype n(%)		Allele n(%)		
	GG	GA	AA	G	A	CC	CT	TT	C	T	CC	CG	GG	C	G
Controls n=50	44 (88.0)	6 (12.0)	0 (0.0)	94 (94.0)	6 (6.0)	39 (78.0)	10 (20.0)	1 (2.0)	88 (88.0)	12 (12.0)	44 (88.0)	5 (10.0)	1 (2.0)	93 (93.0)	7 (7.0)
Cases n=40	33 (82.5)	6 (15.0)	1 (2.5)	72 (90.0)	8 (10.0)	36 (90.0)	4 (10.0)	0 (0.0)	76 (95.0)	4 (5.0)	30 (75.0)	9 (22.5)	1 (2.5)	69 (86.3)	11 (13.7)

Online Supplementary Figure 2S. Multiple DNA sequence alignment of three single nucleotide polymorphisms of SEC23B protein. Fourteen species were aligned using Database UniProtKB release 15.0 of Mar-25-2009; program NCBI BLASTP 2.2.17 [Aug-26-2007] on services.uniprot.org (<http://services.uniprot.org/blast/>). The species' name, protein number and portions of the sequence are shown for each single nucleotide polymorphism discovered. Mutated amino acids are highlighted with the red box, showing the conservation among species. * residues identical in all sequences in the alignment. : conserved substitutions observed. .semi-conserved substitutions observed. The table contains the allele and genotype frequencies in both control subjects and patients.



Online Supplementary Figure 3S. Correlation analysis of ferritin level/dosage age ratio. Patients with one nonsense + one missense mutation (group B, n=10) have a higher ferritin level/age ratio compared with patients with two missense mutations (group A, n=19) (OR=6.9, 95% CI: 1.1-41.8, P=0.03, χ^2 test).