

Supplemental Table S1. PCR conditions and primers for selected SNPs.

SNP ID	Primers (5'-3')	Anneling temperature	SNP detection
rs1043184969	F- gcccttactgacgtgacat R- ctggaatgcagccagaatca	60°C	RFLP: <i>AjiI</i> Thermo scientific
rs954450605	F- gccagcattttattgccta R- ggggtgagcaggttcagtaa	62°C	Sanger sequencing
rs2113453	F- caggaactccccacatgc R- ccagcactgagctgctctc	60°C	RFLP: <i>EcoI30I</i> Thermo Scientific
rs11207020	F- cacacactcaggcactgtt R-ctgagcttggggctttatg	60°C	RFLP: <i>SsiI</i> Thermo scientific
rs192485043	F- gagggcccttcattctctt R- tgggttgacaggtacagcaa	64°C	Sanger sequencing
rs145097803	F- gggaggcattaatgtgctttt R- cccgggattaatgcactctt	60°C	RFLP: <i>EcoRV</i> New England Biolabs
rs929412570	F- ctgtggacaactggaagca R- cctgaagccagatttcaac	60°C	RFLP: <i>EcoI30I</i> Thermo scientific

Supplemental Table S2. Haplotypes in pure and interrupted non-pathogenic alleles.

		Pure									Interrupted					
SNP	Position	(ATTTT) ⁷⁻²¹						(ATTTT) ⁶⁶	(ATTTT) ¹³⁶⁻⁻³⁰⁰		(ATTTT) ⁴³⁻⁸⁰			(ATTTT) ¹⁰⁰⁻⁴⁰⁰		
	chr1 (hg19)															
rs1043184969	57491966	T	T	T	T	T	T	T	T	G	T	T	T	T	T	T
rs514412	57551569	G	G	A	A	G	A	G	G	G	G	G	A	G	A	G
rs954450605	57551605	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G
rs2113453	57833689	T	C	T	C	T	T	T	T	T	T	T	T	T	T	T
rs11207020	57834035	T	C	C	C	C	T	T	T	T	T	T	T	T	T	T
rs192485043	57926567	T	T	T	T	T	T	T	T	A	T	T	T	T	T	T
rs145097803	58201704	T	T	T	T	T	T	T	T	A	T	T	T	T	T	T
rs929412570	58215160	C	C	C	C	C	C	C	C	G	G	C	C	G	C	C
Absolute frequency		10	25	10	1	17	2	1	11	2	6	8	2	2	1	2
Haplotype ID		I	III	VI	VII	II	V	I	I	VIII	IV	I	V	IV	V	I

Supplemental Table S3. Haplotypes sharing the 5 rare SNP alleles from haplotype VIII (SCA37 haplotype) available in 1000g phase 3.

		African						European						Asian
Position	SNP	African	Caribbean in	Esan in Nigeria	Gambian in	western	Mende in Sierra	Leone	UTAH residents	(CEPH)	Tuscani from	Italy	Sri Lanka Tamil	from Uk
chr1	(hg19)													
57491966	rs1043184969	T	T	T	T	T	T	T	T	T	T	T	T	T
57551569	rs514412	G	A	G	A	A	G	A	A	G	G	G	G	G
57551605	rs954450605	G	G	G	G	G	G	G	G	G	G	G	G	G
57832716	(ATTTT) _n													
57833689	rs2113453	C	C	T	C	T	T	C	T	C	C	T	C	T
57834035	rs11207020	T	T	C	C	C	T	C	C	C	C	C	T	C
57926567	rs192485043	T	T	T	T	T	T	T	T	T	A	T	T	T
58201704	rs145097803	T	A	A	A	A	A	A	A	A	T	A	A	A
58215160	rs929412570	C	C	C	C	C	C	C	C	C	C	C	C	C
Haplotypes (n)		1	1	1	1	1	3	1	1	1	1	1	1	2
Alleles (n)		192	198	226	170	198	214	204						

Bold- Rare SNPs; Green- SNP alleles from SCA37 haplotype (VIII).