Supplemental Online Content

Basmanav FB, Cesarato N, Kumar S, et al. Assessment of the genetic spectrum of uncombable hair syndrome in a cohort of 107 individuals. *JAMA Dermatol*. Published online August 31, 2022. doi:10.1001/jamadermatol.2022.2319

This supplemental material has been provided by the authors to give readers additional information about their work.

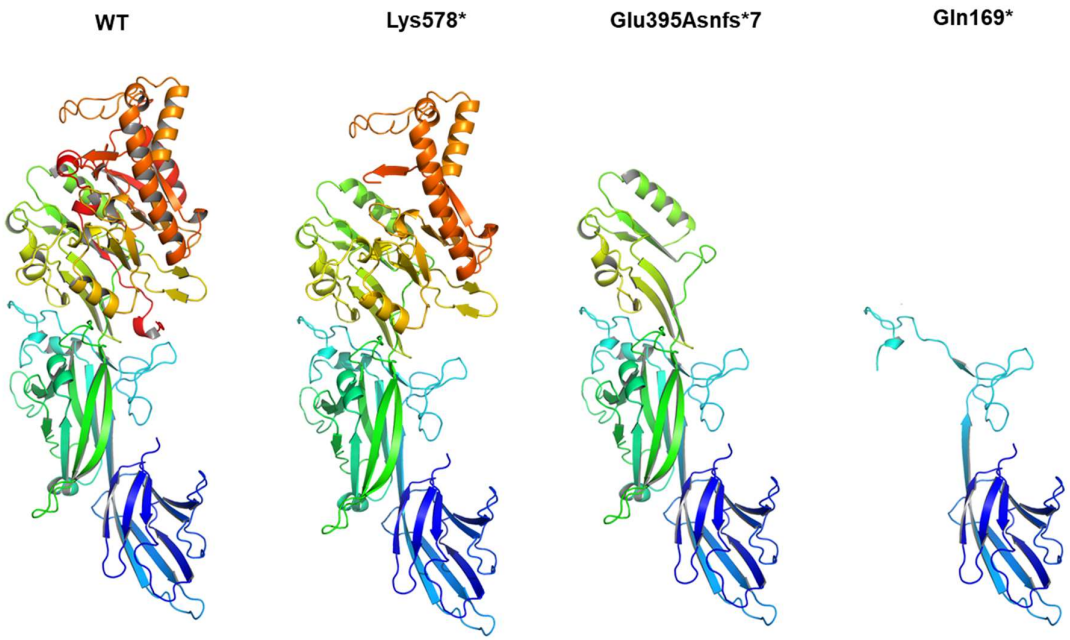
eMethods. Whole-exome sequencing and haplotype analysis

For whole-exome sequencing, we fragmented 1 µg of DNA with sonication technology (Bioruptor, diagenode, Liège, Belgium). The fragments were end-repaired and adaptor-ligated, including incorporation of sample index barcodes. After size selection and pooling, we subjected all libraries to an enrichment process with the SureSelect Human All Exon V6 or V7 kits following manufacturer's procedures (Agilent, Santa Clara, CA, USA). The final libraries were sequenced on an Illumina HiSeq 4000 or an NovaSeq 6000 sequencing instrument (Illumina, San Diego, CA, USA) with a paired-end 2×75 bp or 2x101 bp protocol. The reads were then mapped to the Genome Reference Consortium human genome build 38 (UCSC version hg20) using the Burrows-Wheeler Alignment tool (BWA-mem) (unpublished data, H.L.). The 30X coverage reached 68 - 88% of the target sequences. The Genome Analysis Toolkit (GATK) v.3.8[1] was used to mark duplicated reads, perform local realignment around short insertions and deletions (INDELs), recalibrate the base quality scores, and call SNVs and INDELs. Samtools v.1.6[2] and Platypus v.0.8.1[3] were used as independent short read variation caller. The Varbank2 pipeline v.3.3/v.3.10/v.3.14 and interface, which were developed in-house at the Cologne Center for Genomics, were used for data analysis and filtering (https://varbank.ccg.uni-koeln.de/varbank2, unpublished data, H.T., J.A., and P.N.). All variation calls were filtered for rare, high-quality variants in either a homozygous (AF>0.75) or heterozygous (AF>0.25) state that were predicted to modify a protein sequence or impair splicing. Criteria used to determine high quality GATK UnifiedGenotyper calls were for SNVs: passed VQSR filter, QD>5; MQ>50; FS<40; MQRankSum>−5; ReadPosRankSum>−5 and for INDELs: passed VQSR filter; QD>4; FS<100; ReadPosRankSum>-5. Platypus and Mpileup calls were filtered for QD>5 and MQ>50. Rare variant status was assigned on the basis of an MAF of ≤ 0.01, according to the highest value in any given population of an in-house GRCh38 lifted version of the gnomAD v.2.1 database[4].

Genotyping for haplotype analysis was carried out using GSAMD24v2-0 chips (Ilumina) according to manufacturer's instructions (Infinium HTS Assay protocol, Ilumina). Phasing of the samples genotypes was performed with Eagle 2.4.1[5] using 1000 Genomes Phase 3 data[6] as a reference panel. Haplotypes were reconstructed in a 10 megabase region around *PADI3* on chromosome 1. Of the 40 individuals, 16 were carriers for a pair of the four analyzed variants, 11 were homozygous carriers for one of the variants and 13 carried one of the four variants in heterozygosity (either unaffected parents or affected individuals carrying an additional *PADI3* pathogenic variant that was not assessed by haplotype analysis). In detail, haplotypes of 22 individuals (of whom 2 were homozygous carriers) were analyzed for c.335T>A (p.Leu112His), haplotypes of 26 individuals (of whom 8 were homozygous carriers) were analyzed for c.881C>T (p.Ala294Val), haplotypes of 4 heterozygous variant carriers were analyzed for c.1813C>A (p.Pro605Thr), and haplotypes of two individuals (of whom one was a homozygous carrier) were analyzed for c.505C>T (p.Gln169*).

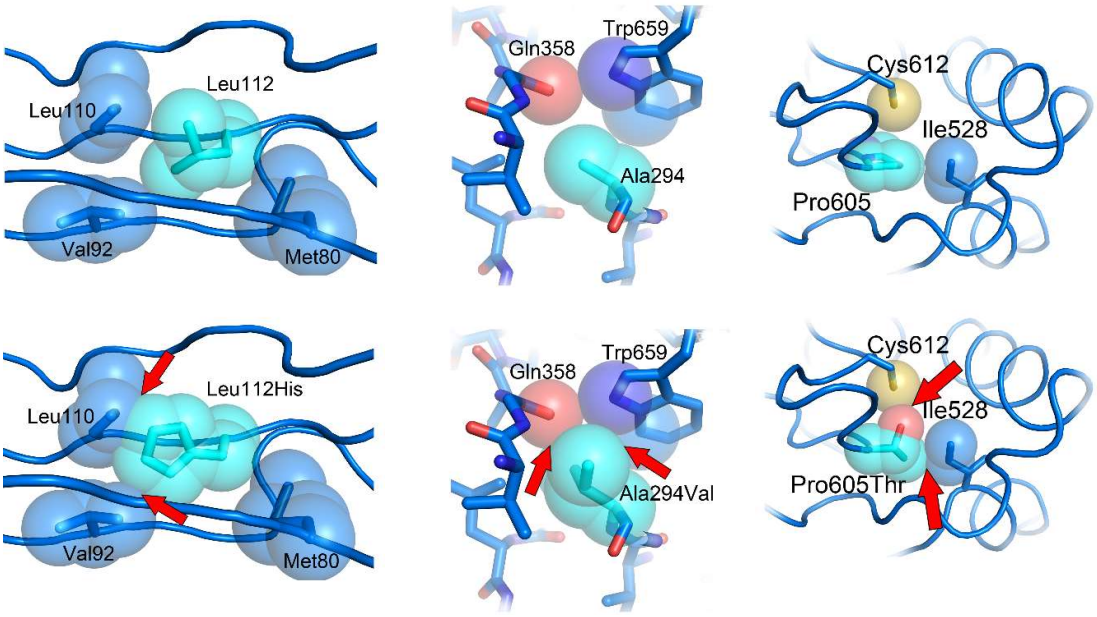
eFigure 1. Clinical manifestation of UHS in additional individuals and the respective underlying genotypes

Additional individuals with UHS carrying bi-allelic pathogenic variants in *PADI3*. Their respective genotypes are (a) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (b) c.335T>A (p.Leu112His) homozygous, (c) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (d) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (e) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (f) c.881C>T (p.Ala294Val) & c.505C>T (p.Gln169*), (g) c.881C>T (p.Ala294Val) homozygous, (h) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (i) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (j) c.335T>A (p.Leu112His) & c.652G>A (p.Gly218Ser), (k) c.881C>T (p.Ala294Val) homozygous, (l) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (m) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (n) c.335T>A (p.Leu112His) & c.1813C>A (p.Pro605Thr), (o) c.881C>T (p.Ala294Val) homozygous, (p) c.881C>T (p.Ala294Val) homozygous, (q) c.335T>A (p.Leu112His) & c.652G>A (p.Gly218Ser), (r) c.505C>T (p.Gln169*) homozygous, (s) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val), (t) c.335T>A (p.Leu112His) & c.881C>T (p.Ala294Val).



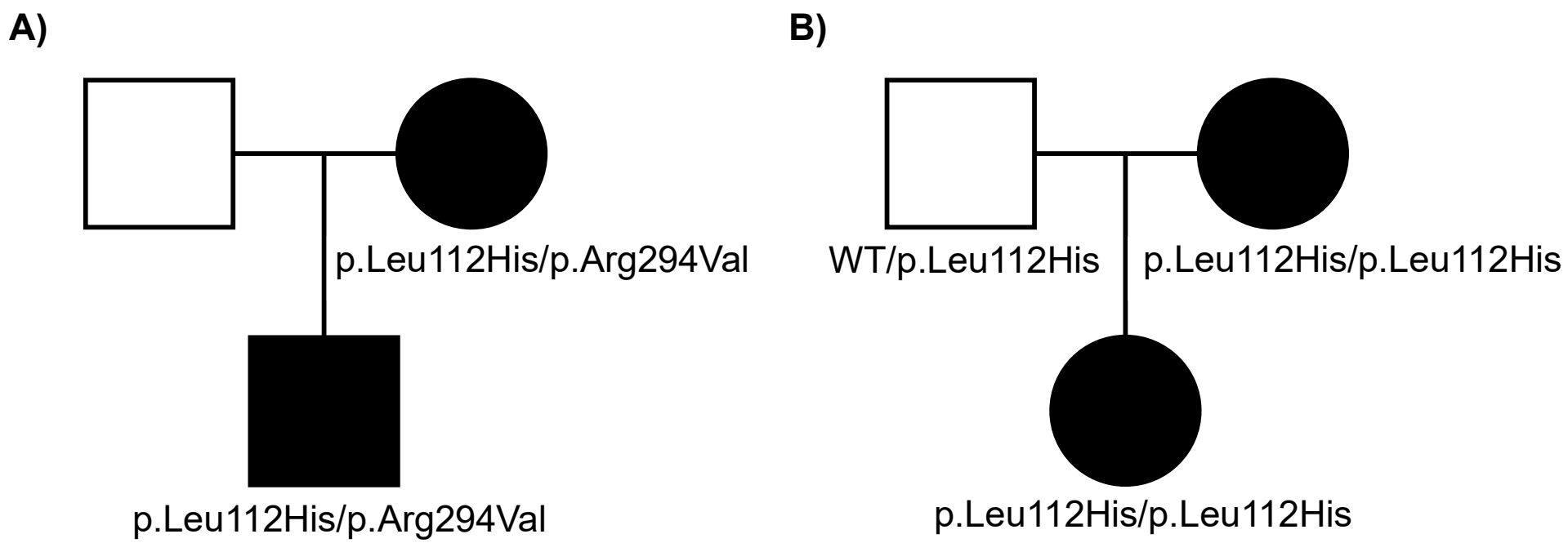
eFigure 2. 3-Dimensional protein modeling of the pathogenic stop and frameshift variants identified in *PADI3*

The wild type protein (left) in comparison to the previously identified mutant Lys578* and the newly identified mutant proteins Glu395Ansfs*7 and Gln169*, respectively. It should, however, be noted that the mutant proteins, particularly the latter two with earlier premature stop codons, may not be translated at all due to a nonsense mediated mRNA decay.



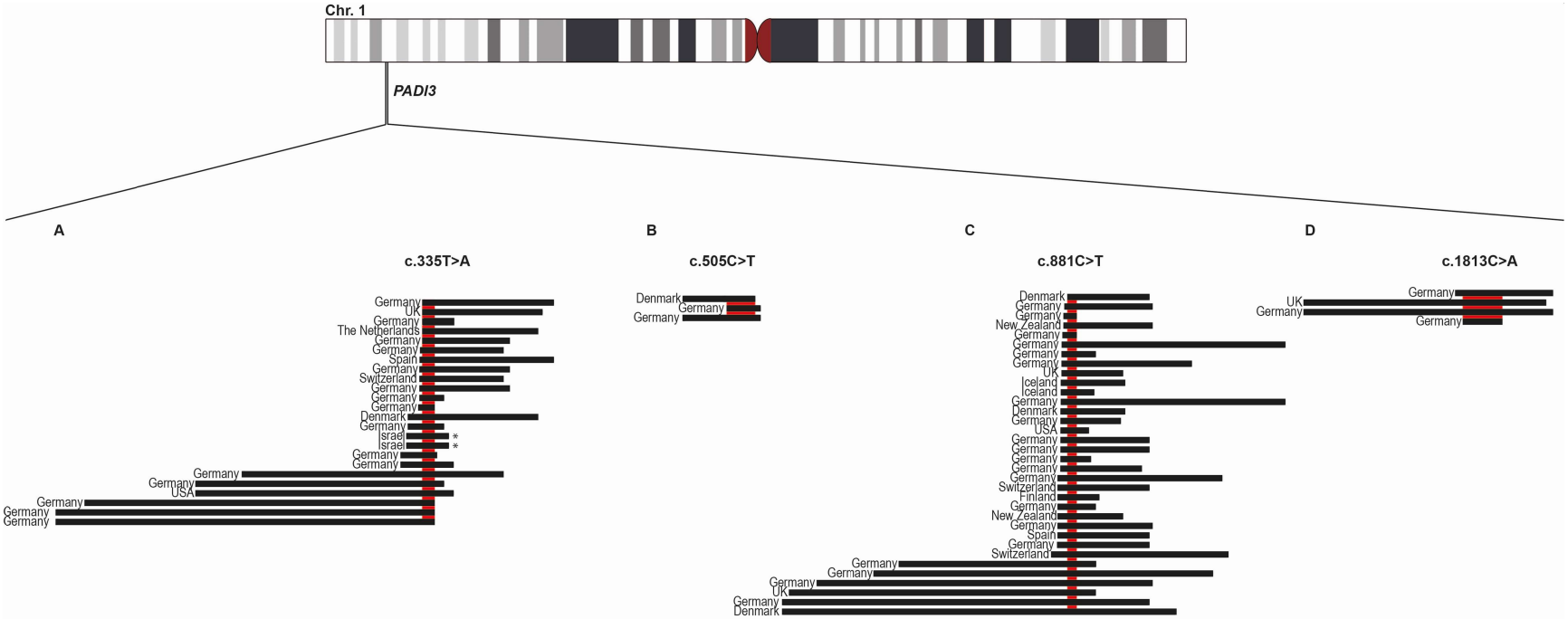
eFigure 3. 3-Dimensional protein modeling of the three most common *PADI3* variants

The top panel shows the wild type protein at the respective site of the substitution p.Leu112His (left), p.Ala294Val (middle) and p.Pro605Thr (right). The bottom panel shows the site of substitution in the respective mutant proteins.



eFigure 4. Genetic screening of two pedigrees suggesting an autosomal dominant inheritance

Mutation analysis revealed in both of the cases that affected parents also carry two pathogenic *PADI3* variants.



eFigure 5. The variable extent of sharing across haplotypes carrying the respective *PADI3* pathogenic variants.

The minimally shared regions across all haplotypes carrying a respective pathogenic variant are denoted with a dark red shade. *These two haplotypes are from an individual born to consanguineous parents and are thus inherently representing the same

eTable 1. Pathogenic *PADI3* variants underlying UHS

Variant	Countries	Minor allele frequency in gnomAD
c.60_66delinsTGCTTGG (p.Gly22Trp)	Germany (1)	*
c.335T>A (p.Leu112His)	Ireland (1), Italy (1), Germany (28), Denmark (1), USA (6), UK (3), The Netherlands (2), Switzerland (1), Belgium (1), Slovenia (1), Portugal (1), Spain (1), Israel (1)	4.52e-3
c.505C>T (p.Gln169*)	Germany (1), Denmark (2)	1.45e-4
c.556C>T (p.Arg186Trp)	Italy (1)	5.31e-5
c.652G>A (p.Gly218Ser)	Germany (1), USA (1)	2.83e-5
c.881C>T (p.Ala294Val)	Germany (31), New Zealand (1), Denmark (4), Finland (1), USA (5), Switzerland (3), Iceland (1), The Netherlands (1), UK (2), Chile (1), Belgium (1), Hungary (1), Norway (1), Spain (1)	6.63e-3
c.1114A>G (p.Arg372Gly)	Chile (1)	1.19e-5
c.1115G>T (p.Arg372Met)	Ukraine (1)	7.79e-5
c.1183del (p.Glu395Asnfs*7)	Germany (1)	4.00e-6
c.1732A>T (p.Lys578*)	Switzerland (1)	-
c.1813C>A (p.Pro605Thr)	Germany (6), UK (1), Denmark (1)	4.43e-4
c.1868C>T (p.Pro623Leu)	Finland (1)	3.29e-4

The newly identified pathogenic variants since our previous report are depicted in bold. Note that the countries of individuals who carry two different variants are represented twice (e.g. only one individual from Chile participated in the study who carried c.881C>T (p.Ala294Val) and c.1114A>G (p.Arg372Gly)). Exception to this are individuals with parents of different countries of origin and for whom we could link the inheritance of the two mutations to the respective parent by genetic screening of the whole family (e.g. c.1115G>T (p.Arg372Met)).

*c.60_66delinsTGCTTGG (p.Gly22Trp) may also represent a rare haplotype composed of three single nucleotide variants, namely, rs1360902614, rs1221559173 and rs1344213588 which have been observed in the same single individual (minor allele frequency=6.5e-6) in gnomAD.

.

eTable 2. The variable extent of sharing across individual haplotypes carrying c.505C>T (p.Gln169*)

Country	Start	End	Length	SNP Start	SNP End
Germany	17535399	17674847	139448	rs57285689	rs56698682
Denmark	17342659	17651637	308978	rs9435672	rs13375202
Germany	17342659	17674847	332188	rs9435672	rs56698682

The start and end positions are based on hg19.

eTable 3. The variable extent of sharing across individual haplotypes carrying c.335T>A (p.Leu112His)

Country	Start	End	Length	SNP Start	SNP End
Germany	17490623	17640368	149746	rs75362086	rs35943476
Germany	17459855	17610300	150446	rs910223	rs12135399
Israel*	17484754	17660468	175715	rs115806112	rs34309058
Israel*	17484754	17660468	175715	rs115806112	rs34309058
Germany	17459855	17680807	220953	rs910223	rs3094881
Germany	17542564	17896945	354382	rs113564258	rs72648415
Switzerland	17540345	17896945	356601	rs11203332	rs72648415
Germany	17552573	17923518	370946	rs150923523	rs12127699
Germany	17540345	17923518	383174	rs11203332	rs12127699
Germany	17540345	17923518	383174	rs11203332	rs12127699
The Netherlands	17552573	18043943	491371	rs150923523	rs4920625
UK	17552573	18062545	509973	rs150923523	rs12407226
Denmark	17490623	18043943	553321	rs75362086	rs4920625
Germany	17552573	18110520	557948	rs150923523	rs112353530
Spain	17541048	18110520	569473	rs57174497	rs112353530
Germany	16588211	17640368	1052158	rs183424675	rs35943476
USA	16588211	17680807	1092597	rs183424675	rs3094881
Germany	16785463	17896945	1111483	rs374130996-	rs72648415
Germany	16115747	17600317	1484571	rs2902164	rs12037653
Germany	15994251	17600317	1606067	rs41270291	rs12037653
Germany	15994251	17600317	1606067	rs41270291	rs12037653

The start and end positions are based on hg19. *These two haplotypes are from an individual born to consanguineous parents and are thus inherently representing the same information.

eTable 4. The variable extent of sharing across individual haplotypes carrying c.881C>T (p.Ala294Val)

Country	Start	End	Length	SNP Start	SNP End
Germany	17548261	17599930	51669	rs12042325	rs200523725
Germany	17543844	17599930	56086	rs74058987	rs200523725
USA	17535981	17651637	115656	rs112338010	rs13375202
Germany	17535981	17660352	124371	rs112338010	rs78540609
Iceland	17536952	17674847	137895	rs2977272	rs56698682
Germany	17540345	17680807	140462	rs11203332	rs3094881
Germany	17523367	17680807	157440	rs115033991	rs3094881
Finland	17523367	17696338	172971	rs115033991	rs120483
Germany	17535981	17787315	251334	rs112338010	rs113343434
UK	17540345	17796711	256366	rs11203332	rs4284283
Iceland	17536952	17805536	268584	rs2977272	rs147190895
Denmark	17535981	17805536	269555	rs112338010	rs147190895
New Zealand	17523367	17796711	273344	rs115033991	rs4284283
Germany	17535399	17876478	341079	rs57285689	rs77544322
Denmark	17565110	17908963	343853	rs747282761	rs34984825
Germany	17552573	17922259	369686	rs150923523	rs74059336
Germany	17535981	17908963	372982	rs112338010	rs34984825
Germany	17535981	17908963	372982	rs112338010	rs34984825
New Zealand	17548261	17922259	373998	rs12042325	rs74059336
Spain	17523367	17908963	385596	rs115033991	rs34984825
Switzerland	17523367	17908963	385596	rs115033991	rs34984825
Germany	17521274	17908963	387689	rs3003469	rs34984825
Germany	17523367	17922259	398892	rs115033991	rs74059336
Germany	17540345	18089348	549003	rs11203332	rs117069891
Germany	17523367	18222257	698890	rs10888012	rs3094881
Switzerland	17495569	18247926	752357	rs143218283	rs2210811
Germany	16848016	17682100	834084	rs113723037	rs1748012
Germany	17541048	18489557	948509	rs57174497	rs74055922
Germany	17536952	18489557	952605	rs2977272	rs74055922
UK	16378219	17680807	1302588	rs121909133	rs3094881
Germany	16495998	17922259	1426261	rs924206	rs74059336
Germany	16741089	18183104	1442015	rs114515972	rs529557
Germany	16349169	17908963	1559794	rs202069202	rs34984825
Denmark	16349169	18023690	1674521	rs202069201	rs2270976

The start and end positions are based on hg19.

eTable 5. The variable extent of sharing across individual haplotypes carrying c.1813C>A (p.Pro605Thr)

Country	Start	End	Length	SNP Start	SNP End
Germany	17523367	17685905	162538	rs115033991	rs150221181
Germany	17490623	17900974	410351	rs75362086	rs7516512
UK	16848016	17871837	1023821	rs113723037	rs75174468
Germany	16848016	17900974	1052958	rs113723037	rs7516512

The start and end positions are based on hg19.

eReferences

1. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303. doi:10.1101/gr.107524.110
2. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 15 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
3. Rimmer A, Phan H, Mathieson I, et al. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet.* 2014;46(8):912-918. doi:10.1038/ng.3036
4. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 18 2016;536(7616):285-291. doi:10.1038/nature19057
5. Loh PR, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet.* 2016;48(11):1443-1448. doi:10.1038/ng.3679
6. Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74. doi:10.1038/nature15393