The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/152997

Please be advised that this information was generated on 2018-10-16 and may be subject to change.
Homozygosity Mapping and Targeted Sanger Sequencing Reveal Genetic Defects Underlying Inherited Retinal Disease in Families from Pakistan


Abstract

Background
Homozygosity mapping has facilitated the identification of the genetic causes underlying inherited diseases, particularly in consanguineous families with multiple affected individuals. This knowledge has also resulted in a mutation dataset that can be used in a cost and time effective manner to screen frequent population-specific genetic variations associated with diseases such as inherited retinal disease (IRD).

Methods
We genetically screened 13 families from a cohort of 81 Pakistani IRD families diagnosed with Leber congenital amaurosis (LCA), retinitis pigmentosa (RP), congenital stationary night blindness (CSNB), or cone dystrophy (CD). We employed genome-wide single nucleotide polymorphism (SNP) array analysis to identify homozygous regions shared by affected individuals and performed Sanger sequencing of IRD-associated genes located in the sizeable homozygous regions. In addition, based on population specific mutation data we...
performed targeted Sanger sequencing (TSS) of frequent variants in AIPL1, CEP290, CRB1, GUCY2D, LCA5, RPGRIP1 and TULP1, in probands from 28 LCA families.

Results
Homozygosity mapping and Sanger sequencing of IRD-associated genes revealed the underlying mutations in 10 families. TSS revealed causative variants in three families. In these 13 families four novel mutations were identified in CNGA1, CNGB1, GUCY2D, and RPGRIP1.

Conclusions
Homozygosity mapping and TSS revealed the underlying genetic cause in 13 IRD families, which is useful for genetic counseling as well as therapeutic interventions that are likely to become available in the near future.

Introduction
Inherited retinal diseases (IRD) refer to a clinically and genetically heterogeneous group of genetic eye disorders in which the photoreceptors and retinal pigment epithelium can be affected. There is an overlap of clinical features between different IRDs, which includes syndromic or non-syndromic conditions. In cone dystrophy (CD), only central vision is impaired, whereas in cone-rod dystrophy (CRD) peripheral vision is also compromised. In retinitis pigmentosa (RP) initially peripheral vision is affected, which later progresses to central vision defects. In contrast, congenital stationary night blindness (CSNB) only involves night vision loss due to defective rod photoreceptors. The most severe form of IRD is Leber congenital amaurosis (LCA), in which patients suffer from complete blindness in the first year of life [1–3]. In addition to clinical diversity, the genetic heterogeneity in IRDs is reflected by 221 genes that have thus far been found to be mutated in IRD (https://sph.uth.edu/retnet/). Besides clear phenotypic differences, different defects in the same gene may also be responsible for different clinical phenotypes, for example different variations in RPGRIP1 (MIM # 605446) are known to cause RP, LCA and CRD, TULP1 (MIM # 602280) mutations have been shown to cause RP, LCA or CD [4], and RPGR (MIM # 312610) variants are known to cause RP or CD [5]. It has also been observed that the inherited forms of retinal diseases follow all Mendelian modes of inheritance [3].

The prevalence of retinal dystrophies has been estimated at 1 in 3,000 individuals worldwide, with RP being the most common type affecting 1 in 4,000 individuals [6–8]. In Pakistan the prevalence of IRDs is not well defined but a hospital based study estimated that 1 in 800 patients who attended the ophthalmic outpatient department, were affected with retinal diseases, with RP as the most common phenotype [9]. However, such inherited disorders have been observed more commonly in consanguineous families than in non-consanguineous families. Hamamy et al. [10] calculated the percentage of the mode of inheritance of genetically inherited diseases and suggested that consanguinity is strongly correlated with the prevalence of autosomal recessive diseases. In addition, similar observations have been made by Bittles [11] and Nirmalan et al. [12]. In the Pakistani population more than 60% of marriages are consanguineous, and among them more than 80% are first cousin marriages [11]. For consanguineous IRD families with multiple affected individuals, the causative genetic defects can be identified using genome wide single nucleotide polymorphism (SNP)-array analysis followed by homozygosity...
mapping [13–15]. In view of the high genetic heterogeneity, homozygosity mapping in most isolated cases cannot unambiguously point to a single IRD-associated gene. Khan et al. [16] comprehensively reviewed the genetic causes of IRDs in the Pakistani population, and proposed an initial mutation screening method of IRDs by analyzing frequently occurring mutations. Since 2008, we have collected 81 consanguineous IRD families in Pakistan, and reported on the underlying genetic causes in 25 of these families [14,16–26].

In the current study we report the results from an additional 13 of these previously identified families. We performed genome-wide SNP genotyping followed by homozygosity mapping and candidate gene sequencing. In addition, we analyzed several families using targeted Sanger sequencing (TSS) of frequently reported variations from Pakistani population in APL1 (MIM # 604392), CEP290 (MIM # 610142), CRB1 (MIM # 604210), GUCY2D (MIM # 600179), LCA5 (MIM # 611408), RGR1P1 and TULP1.

Materials and Methods

Subjects

Since 2008, we have recruited 81 IRD families from different regions of Pakistan, i.e. CD (2 families), CSNB (5 families), LCA (36 families), and RP (38 families) (S1 Table).

Ethics statement

The current study adheres to the declaration of Helsinki, and was approved by the Department of Biosciences Ethics Review Board of COMSATS Institute of Information Technology, Al-Shifa Eye Trust Hospital, Rawalpindi and Shifa International hospital, Islamabad. The subjects and their families were informed about the purpose of the study and their oral as well as written consent was taken.

Clinical evaluations

The subjects were clinically diagnosed as CD, CSNB, LCA and RP on the basis of detailed ophthalmic evaluations and fundus examination. The affected individuals complaining of reduced central vision with focusing error, photophobia and nystagmus were grouped as CD. The individuals experiencing non-progressive night blindness with normal day vision were categorized as CSNB. The subjects were categorized as LCA if they were congenitally blind, had nystagmus, and sluggish or non-reactive pupillary response. Finally, the cases reporting night vision loss with progressive mid-peripheral vision deterioration were grouped as RP (S1 Table).

DNA isolation

Blood samples were drawn from all available affected and unaffected individuals of the family into ethylenediamine tetra-acetic acid (EDTA)-coated vacutainers. DNA was extracted in Tris-EDTA buffer using a standard organic extraction protocol for 53 families and stored at −20°C. For the remaining 28 families, a standard salting out protocol was employed [27].

Genetic linkage analysis

Genetic linkage analysis was carried out for 53 of 81 families using microsatellite markers or whole genome SNP array platforms such as Illumina_10K, Affymetrix_6K, Human Omni express_700K and Cytoscan HD (Fig. 1, Table 1). The SNP array data were analyzed by homozygosity mapping using an online tool ‘Homozygosity Mapper’ (http://www.homozygositymapper.org/). Sanger sequencing was performed for IRD-associated genes. These genes were prioritized according to the size of the region in which they were located.
First, any mutation hotspot in the gene, if known, was sequenced followed by sequencing of other exons along with flanking intronic sequences. Novel missense mutations were also screened in ethnicity-matched controls (n = 90).

**Targeted Sanger sequencing (TSS)**

In probands from 28 families diagnosed with LCA (from a total of 36), the targeted variant screening was done using Sanger sequencing (Table 2). The variants were chosen based on their frequency in the Pakistani population [16]. In addition, we also screened our LCA panel with other frequent variants that are associated with LCA in the Caucasian population including the intronic CEP290 variant c.2991+1655A>G [28] and the GUCY2D exon 12 variant c.2302C>T [29,30]. The RPGRIP1 variant (c.3565C>T) described by Abu-Safieh et al. [31] was found to be segregating in one of our LCA families (F04) and therefore this variant was also analyzed in our LCA cohort [16].

**In silico analysis**

The pathogenicity index for the identified missense mutations was calculated in silico using Sorting Intolerant From Tolerant (SIFT) (http://sift.bii.a-star.edu.sg/), Mutation Taster (http://www.mutationtaster.org/), and Polymorphism Phenotyping V2 (PolyPhen-2) (http://genetics.bwh.harvard.edu/pph2/). The PhyloP score and Grantham distances were also recorded to check the nucleotide conservation and change in amino acid physiochemical properties. The frequency of the variant in the general population was determined using Exome Variant Server (EVS) (http://evs.gs.washington.edu/EVS/), 1000 genomes and our in-house mutation database, which contained exome sequence variant data of 2,096 persons with various human conditions. To assess the effect of a missense change on the protein structure of CNGA1 we used the HOPE server http://www.cmbi.ru.nl/hope/home).
### Table 1. Results of targeted Sanger sequencing and genetic analyses of 13 IRD families.

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Disease</th>
<th>Genotyping method</th>
<th>Number of Hz regions</th>
<th>Size of Hz region, in Mb</th>
<th>Ranking of Hz region</th>
<th>Gene</th>
<th>DNA mutation</th>
<th>Predicted protein variant</th>
<th>First report of variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>F01</td>
<td>LCA</td>
<td>TSS</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>AIPL1</td>
<td>c.834G&gt;A</td>
<td>p.(W278*)</td>
<td>[32]</td>
</tr>
<tr>
<td>F02</td>
<td>LCA</td>
<td>TSS</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>AIPL1</td>
<td>c.834G&gt;A</td>
<td>p.(W278*)</td>
<td>[32]</td>
</tr>
<tr>
<td>F03</td>
<td>LCA</td>
<td>TSS</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>GUCY2D</td>
<td>c.2283del</td>
<td>p.(S762fs*22)</td>
<td>This study</td>
</tr>
<tr>
<td>F04</td>
<td>LCA</td>
<td>Cytoscan HD, SS</td>
<td>5</td>
<td>14.0</td>
<td>1</td>
<td>RPGRIP1</td>
<td>c.3565C&gt;T</td>
<td>p.(R1189*)</td>
<td>[31]</td>
</tr>
<tr>
<td>F05</td>
<td>LCA</td>
<td>Illumina_700K, SS</td>
<td>&gt;10</td>
<td>4.5</td>
<td>19</td>
<td>RPGRIP1</td>
<td>c.930+1G&gt;A</td>
<td>p.(?)</td>
<td>This study</td>
</tr>
<tr>
<td>F06</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>3</td>
<td>1.6</td>
<td>2</td>
<td>RPE65</td>
<td>c.131G&gt;A</td>
<td>p.(R44Q)</td>
<td>[36]</td>
</tr>
<tr>
<td>F07</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>3</td>
<td>5.2</td>
<td>1</td>
<td>RPE65</td>
<td>c.361del</td>
<td>p.(S121Lfs*6)</td>
<td>[36]</td>
</tr>
<tr>
<td>F08</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>4</td>
<td>18.8</td>
<td>1</td>
<td>CNGA1</td>
<td>c.1298G&gt;A</td>
<td>p.(G433D)</td>
<td>This study</td>
</tr>
<tr>
<td>F09</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>1</td>
<td>6.5</td>
<td>1</td>
<td>CNGB1</td>
<td>c.2493–2A&gt;G</td>
<td>p.(?)</td>
<td>This study</td>
</tr>
<tr>
<td>F10</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>&gt;10</td>
<td>13.4</td>
<td>2</td>
<td>CRB1</td>
<td>c.2234C&gt;T</td>
<td>p.(T745M)</td>
<td>[38]</td>
</tr>
<tr>
<td>F11</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>&gt;10</td>
<td>9.2</td>
<td>1</td>
<td>TULP1</td>
<td>c.1466A&gt;G</td>
<td>p.(K489R)</td>
<td>[40]</td>
</tr>
<tr>
<td>F12</td>
<td>RP</td>
<td>Illumina_700K, SS</td>
<td>&gt;10</td>
<td>6.5</td>
<td>1</td>
<td>PDE6A</td>
<td>c.304C&gt;A</td>
<td>p.(R102S)</td>
<td>[41]</td>
</tr>
<tr>
<td>F13</td>
<td>RP</td>
<td>Affymetrix 10K, SS</td>
<td>6</td>
<td>—</td>
<td>—</td>
<td>RPGR</td>
<td>c.2426_2427del</td>
<td>p.(E809Gfs*25)</td>
<td>[42]</td>
</tr>
</tbody>
</table>

Hz, Homozygous; Mb, Megabases; SS, Sanger sequencing entire gene; TSS, targeted Sanger sequencing; DNA, Deoxyribonucleic acid.

### Table 2. Frequent variants pre-screened in 28 LCA families.

<table>
<thead>
<tr>
<th>Gene</th>
<th>DNA variant</th>
<th>Protein variant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIPL1</td>
<td>c.834G&gt;A</td>
<td>p.(W278*)</td>
<td>[16,32,62,63]</td>
</tr>
<tr>
<td>CRB1</td>
<td>c.2234C&gt;T</td>
<td>p.(T745M)</td>
<td>[36]</td>
</tr>
<tr>
<td>CRB1</td>
<td>c.2536G&gt;A</td>
<td>p.(G846R)</td>
<td>[64]</td>
</tr>
<tr>
<td>CRB1</td>
<td>c.2966T&gt;C</td>
<td>p.(I989T)</td>
<td>[64]</td>
</tr>
<tr>
<td>CEP290</td>
<td>c.2991T&gt;A</td>
<td>p.(?)</td>
<td>[28]</td>
</tr>
<tr>
<td>GUCY2D</td>
<td>c.2302C&gt;T</td>
<td>p.(R768W)</td>
<td>[30]</td>
</tr>
<tr>
<td>LCA5</td>
<td>c.1151del</td>
<td>p.(P384Qfs*18)</td>
<td>[63,65]</td>
</tr>
<tr>
<td>RPGRIP1</td>
<td>c.3565C&gt;T</td>
<td>p.(R1189*)</td>
<td>[31]</td>
</tr>
<tr>
<td>TULP1</td>
<td>c.1138A&gt;G</td>
<td>p.(T380A)</td>
<td>[17,40,63]</td>
</tr>
<tr>
<td>TULP1</td>
<td>c.1466A&gt;G</td>
<td>p.(K489R)</td>
<td>[39,40]</td>
</tr>
</tbody>
</table>

*In original description [64] this variant erroneously was indicated as c.3101T>C.

*In lymphoblast cells, 50% of the resulting mRNA contains a cryptic exon resulting in a predicted stop mutation and 50% of the mRNA is normal [28].
Results, Discussion and Conclusions

Clinical analyses

Typical features of RP and LCA as described in S1 Table were observed in the corresponding families and probands. The fundus pictures of the probands from selected families are given in S1 Fig.

Families F01 and F02; AIPL1

The AIPL1 exon 6 variation, c.834G>A; p.(W278*) [32] is a frequent LCA-associated variant worldwide and is responsible for 10% of the IRD cases reported so far in the Pakistani population [16]. Sanger sequencing of AIPL1 exon 6 revealed two families with this mutation, which segregated with the disease in these families (Fig. 2, Table 1).

Family F03; GUCY2D

Mutations in GUCY2D are known to cause LCA and CRD [33,34]. In our cohort of LCA, during targeted sequencing of exon 12 to search for c.2302C>T; p.(R768W) variant, we coincidentally identified a novel frame-shift mutation c.2283del; p.(S762Afs*22) in one LCA family (F03). Both affected siblings were homozygous for this 1-bp deletion whereas parents were heterozygous carriers (Fig. 2, Table 1).

Families F04 and F05; RPGRIP1

Genetic linkage analysis revealed homozygous regions harboring the LCA-associated gene RPGRIP1 in two of the families (F04 and F05) from the LCA panel. Upon sequencing RPGRIP1 in family F04, a previously identified nonsense mutation c.3565C>T; p.(R1189*) [31] was identified in exon 22, which segregated with the phenotype in the family (Fig. 2, S1 Fig.). In family F05, a novel canonical splice donor site variation (c.930+1G>A; p.(?)) in intron 7 of the gene was identified. As the canonical splice donor site is affected, intron 7 retention or skipping of exon 7 in the mRNA is most plausible. Introns retention would result in a frameshift that creates an early stop codon after 15 bp resulting in a truncated protein of 315 amino acid residues instead of the full length 1,286 amino acids (Fig. 2, Table 1). Skipping of exon 7 would not result in a frameshift but a deletion of 8 amino acid residues that might affect the three dimensional structure and thereby the function of the protein.

Family F06 and F07; RPE65

The SNP array data of families F06 and F07 were analyzed to identify homozygous regions carrying the genes of interest. In both families, RPE65 was identified in one of the largest homozygous regions. Upon sequencing, the most recurrent mutations, c.131G>A; p.(R44Q) [35] and 361del; p.(S121Lfs*6) [36], were identified in a homozygous state in all affected persons of families F06 and F07, respectively (Fig. 2, S1 Fig., Table 1, S2 Table).

Family F08; CNGA1

Homozygosity mapping data of family F08 revealed the arRP-associated gene CNGA1 in the largest homozygous region of ~19 Mb. Upon Sanger sequencing a novel missense mutation c.1298G>A; p.(G433D) was identified in the proband. This variant is not only absent in EVS and 1000 genomes public mutation databases, but also in our in-house WES database as well as from 90 ethnicity-matched healthy controls. Segregation analysis indicated that the mutation is present homozygously in affected individuals of the family whereas the normal individuals are
heterozygous carriers (Fig. 2, Table 1). In silico analysis supported the pathogenicity of the mutation (S2 Table). The highly conserved non-polar glycine residue at position 433 is substituted by the charged aspartate, a bigger sized amino acid that is also less flexible than glycine. The wild type residue is predicted to be buried in a coiled region on the cytoplasmic face of the ion transport domain. The 433D residue can create structural instability and can affect the ion transport function of the protein [37].
Family F09; *CNGB1*

The largest homozygous region of 6.5 Mb identified in family F09 harbored the arRP-associated gene *CNGB1*. Sequence analysis identified a novel homozygous canonical splice acceptor site mutation in intron 25 of *CNGB1*, c.2493-2A>G; p.(?), which segregated with the disease in the family (Fig. 2). This variant may result in exclusion of exon 26 from the transcript. The open reading frame would be shifted in the resulting transcript, leading to a truncated protein consisting of 831 amino acids (full length protein is 1,251 amino acids). In addition, due to this variation, a strong splice donor site is predicted that could result in the inclusion of a large part of intron 25 and exclusion of exon 26, which eventually would also lead to a premature stop codon (Fig. 2, Table 1, S1 Fig.).

Family F10; *CRB1*

Homozygosity mapping positioned the arRP- and LCA-associated gene *CRB1* in one of the homozygous regions, which was shared between the affected individuals. A previously reported missense mutation, c.2234C>T; p.(T745M) [38], affecting the Laminin-G domain, was identified that segregated with the disease phenotype in the family (Fig. 2, Table 1, S2 Table).

Family F11; *TULP1*

The largest homozygous region obtained for family F11 harbored *TULP1*, known to be associated with arRP and LCA. We identified a previously reported missense change, c.1466A>G; p.(K489R) [39,40], segregating with the disease phenotype in the family (Fig. 2, Table 1, S1 Fig., S2 Table).

Family F12; *PDE6A*

Homozygosity mapping of family F12 revealed that *PDE6A* was in the largest region which spanned 6.5 Mb. The gene was sequenced and a previously identified missense mutation, c.304C>A; p.(R102S), was found to segregate with the disease phenotype in the family [41] (Fig. 2, Table 1, S2 Table).

Family F13; *RPGR*

Family F13 was initially sampled as an autosomal recessive RP family but based on the pedigree structure (affected persons in multiple generations) and the fact that the far majority of the affected individuals are males, suggested X-linked inheritance. The analysis of SNP array data indeed pointed to *RPGR* as the candidate disease gene, as the region on the X-chromosome harboring this gene was found to be shared by all affected males. Sequence analysis of *RPGR* identified a 2-bp deletion, c.2426_2427del; p.(E809Gfs*25), in this family [42]. Interestingly, one of the affected females was also homozygous for the deletion, which is extremely rare in X-linked disorders (Fig. 2, Table 1) [43].

Inherited retinal diseases represent a diverse group of eye disorders that are heterogeneous both at the genotype and phenotype level. So far, mutations in 221 genes have been associated with syndromic and non-syndromic inherited retinal dystrophies, and still more are to be identified. This study underscores the genetic diversity of IRD as we report mutations in 10 different genes causing IRD in 13 families. To come to these results, we performed homozygosity mapping and candidate gene sequencing. This approach is successful for most of the consanguineous families. In outbred families this approach is only successful in a small proportion of families [13]. For such families, a pre-screening of frequently reported mutations can be an alternative method before starting with any high throughput analysis like next generation sequencing.
sequencing (NGS). To test this, we performed TSS of frequently found causative variants from seven IRD genes (Table 2) in 28 LCA probands. Despite being frequent in Pakistani and the Caucasian populations, most of them were not found in 25 families except the AIPL1 variant p. (W278*), which was present in two families (F01 and F02). We were unable to find the recurrent exon 12 variant, p.(R768W) in GUCY2D. However, in the same exon, a novel variant (p. (S762Af’s ‘22)) was identified in the proband of family F03. As we were able to solve only three families using TSS, this does not seem to be the best approach. For other populations this approach only makes sense if the frequent population-specific mutations are known.

Basic research in genetics has not only elucidated the underlying mutations in the causative genes but also provided initial information helpful for designing gene therapy. It has been estimated that 81.5% of all the gene therapy trials in the world are focused on cancer, cardiovascular diseases and monogenic inherited disorders. Other broadly targeted areas for gene therapy include infectious diseases, neurological disorders, ocular diseases, inflammatory diseases and diseases such as chronic renal disease, diabetes, etc [44]. The pre-clinical studies in model organisms, before initiation of any human trials, have provided detailed information not only on the therapeutic efficacy but also about safety and toxicity issues. Moreover, choosing the right model organism, which can provide as much information as possible for human trials is equally important [45]. In case of retinal disease gene therapy trials, a number of successful animal models have been described, for example, AIPL1, GUCY2D, RPGRIP1 and TULP1 knock out mouse models have already been reported in which gene therapy was explored [46–49]. In addition to mouse models, dog models for RPGR and RPGRIP1 gene therapy are also known [50]. Similarly, PDE6A and PDE6B gene therapy proof-of-principle in mouse models were reported by Wert et al. [51]. The delivery method of a recombinant gene construct is important. For example, AAV-based gene therapy has been shown to be successful in a CRD dog model and in humans with RPE65-associated LCA [52–57], as well as in choroideremia subjects [58]. The major limitation of AAV-vector based gene therapy is that these vectors cannot carry inserts larger than 4.9 kb, and therefore other methods, viral and non-viral, are needed. Other types of treatments are based on antisense oligonucleotides for CEP290-associated retinal degeneration [59,60]. Besides these genetic approaches, an oral drug therapy based on 9-cis-retinoid was successful in persons with RPE65 and LRAT mutations [61]. Thus, finding new associations for the IRD will not only add scientific knowledge but will also provide critical information for therapeutics.

In addition to gene therapy, another important aspect is genetic counseling. In the X-linked and autosomal recessive families, unaffected persons can be tested for carriership of the causal variants. Early genetic counseling may include advice on choosing appropriate studies and professions, improving their quality of life. Through proper genetic counseling the prevalence of the respective diseases in these families may decrease.

In conclusion, using homozygosity mapping, Sanger sequencing and TSS approaches we were able to identify the underlying genetic causes in 13 IRD families from Pakistan, and identified four novel variations in CNGA1, CNGB1, GUCY2D and RPGRIP1 in four different families.

Supporting Information

S1 Fig. Fundus photographs of selected probands from families F04, F06, F09 and F11. Arrows mark the vessel attenuation, arrowheads represent changes in macula and a block arrows mark the pigmentary changes. (TIF)
S1 Table. Overview of the Pakistani IRD cohort and clinical characteristics. (DOCX)

S2 Table. In silico analysis of the identified missense mutations. (DOCX)

Acknowledgments
We acknowledge all the participating families and are thankful for their cooperation in this study. We thank Hanka Venselaar, Center for Molecular and Biomolecular Informatics, Radboud University Medical Center, Nijmegen, The Netherlands, for her help in predicting the effect of missense variation on protein structure.

Author Contributions
Conceived and designed the experiments: MAz AIdH RWJC MIK RQ FPMC. Performed the experiments: MM MAj MAz MIK. Analyzed the data: MM MAj MAz LHW MIK. Contributed reagents/materials/analysis tools: NKW SNS BM HA LA SA SM AH STAS SHBA WA YMK AIdH LHW. Wrote the paper: MM MAj MAz LHW RWJC MIK RQ FPMC.

References


