Published in final edited form as:

Cornea. 2009 August ; 28(7): 801–807. doi:10.1097/ICO.0b013e31819672fb.

Exclusion of Positional Candidate Gene Coding Region Mutations in the Common Posterior Polymorphous Corneal Dystrophy 1 Candidate Gene Interval

Anthony J. Aldave, M.D.¹, Vivek S. Yellore, Ph.D.¹, Rosalind C. Vo, B.S.¹, Khairidzan M. Kamal, M.D.¹, Sylvia A. Rayner, B.Sc.¹, Christopher L. Plaisier, B.S.², Michael C. Chen, B.S.¹, Mausam R. Damani, B.A.¹, Michele N. Pham, B.A.¹, Michael B. Gorin, M.D., Ph.D.¹, Eric Sobel, Ph.D.², and Jeanette Papp, Ph.D.²

¹The Jules Stein Eye Institute, David Geffen School of Medicine at UCLA

²Department of Human Genetics, David Geffen School of Medicine at UCLA

Abstract

Purpose—Posterior polymorphous corneal dystrophy (PPCD) is an autosomal dominant disorder of the corneal endothelium associated with visually significant corneal edema and glaucoma. Statistical genetic analysis of four families with PPCD has demonstrated linkage to a 2.4 cM common support interval on chromosome 20 bordered by the markers D20S182 and D20S139. We sought to identify the genetic basis of PPCD linked to chromosome 20 (PPCD1) by screening the 26 positional candidate genes between these markers in a family previously mapped to the PPCD1 region.

Methods—The coding regions of the 26 positional candidate genes mapped to the common PPCD1 support interval were amplified and sequenced in affected and unaffected individuals from a family previously linked to the PPCD1 locus. Nine other genes positioned just outside of the common PPCD1 support interval but within the autosomal dominant congenital hereditary endothelial dystrophy (CHED1) interval were also screened.

Results—Four DNA sequence variants in three of the positional candidate genes demonstrated complete segregation with the affected phenotype: p.Thr109Thr (rs6111803) in *OVOL2*, p.Arg56Gln (novel variant - RPSnovel) in *RPS19P1*, and p.Thr85Thr (rs1053834) and p.Pro99Ser (rs1053839) in *C20orf79*. Each of these four sequence variants demonstrated significant linkage with the affected phenotype in this family ($p\text{-value} = 2.5 \times 10^{-7}$ for RPSnovel, rs1053834 and rs1053839; $p\text{-value} = 8.6 \times 10^{-7}$ for rs6111803). However, we also identified each of these four sequence variants in $\geq 9\%$ of unaffected control individuals. The haplotype upon which the disease causing mutation is segregating was found to have a population frequency of 4.2% in the CEPH HapMap trios. While a number of other previously described and novel SNPs were identified in the 35 positional candidate genes located within the PPCD1 and CHED1 intervals, none segregated with the affected phenotype.

Conclusions—We report the absence of a presumed pathogenic coding region mutation in the common PPCD1 support interval. Although minor alleles of four SNPs were identified that segregated with the affected phenotype, the relatively high frequency of each minor allele in the general population indicates that none is a candidate for the causal variant for PPCD. Instead, the causal variant is most likely a coding region deletion or a variant in a non-coding region of the PPCD1 common support interval.

Keywords

Posterior polymorphous corneal dystrophy; Candidate gene screening; Linkage; Congenital hereditary endothelial dystrophy

INTRODUCTION

Posterior polymorphous corneal dystrophy (PPCD, MIM #122000) is an autosomal dominant corneal endothelial dystrophy associated with a varied affected phenotype, ranging from asymptomatic corneal endothelial changes in the majority of patients to congenital corneal edema and glaucoma in the most severely affected individuals. Locus heterogeneity has been demonstrated for PPCD. Pathogenic mutations have been identified in the zinc finger E-box binding homeobox 1 gene (*ZEB1*; MIM 189909), also known as the transcription factor 8 gene (*TCF8*), on chromosome 10 in seventeen families with PPCD (PPCD3).^{1–3} Another locus, PPCD1, was demonstrated by linkage to the pericentromeric region of chromosome 20, first by Heon and colleagues⁴ and more recently by Gwilliam and colleagues⁵ and our group.⁶ The region of chromosome 20 that is contained within the support interval for PPCD in all three studies, the PPCD1 common support interval (Figure 1), is a 2.4 cM region between markers D20S182 and D20S139 that includes 26 mapped genes (<http://www.ncbi.nlm.nih.gov/build/36.3>). Based on the assumption that the genetic basis of PPCD is the same in each of the families linked to this region of chromosome 20, we sought to identify the causal variant by screening these 26 positional candidate genes in a large Caucasian American PPCD1-linked family that we reported previously.⁶

Autosomal dominant congenital hereditary endothelial dystrophy (CHED1, MIM #121700), characterized by corneal endothelial dysfunction and stromal edema in the first decade of life, has been mapped to a region of chromosome 20 nearly identical to the PPCD1 common support interval (Figure).⁷ The 35 genes mapped to the 2.7 cM CHED1 support interval between markers D20S48 and D20S471 include the 26 genes within the PPCD1 common support interval, as well as 7 genes adjacent to the telomeric border and 2 genes adjacent to the centromeric border. As the support intervals for CHED1 and PPCD1 are nearly identical, it remains to be determined whether CHED1, which demonstrates significant clinical and histologic similarities to PPCD,^{8–10} may actually represent an early-onset variant of PPCD and not a distinct clinical entity. Therefore, we also screened the nine genes outside the common PPCD1 support interval but within the CHED1 support interval.

MATERIALS AND METHODS

The researchers followed the tenets of the Declaration of Helsinki in the treatment of the subjects reported herein. Study approval was obtained from the Institutional Review Board at the University of California, Los Angeles (UCLA IRB # 02-10-092-01).

Patient Identification / DNA Collection and Preparation

The diagnosis of PPCD was based on the presence of established clinical criteria.⁶ DNA collected from affected and unaffected members of a large Caucasian family with PPCD1 was isolated as previously described.⁶ DNA collected from unaffected, unrelated individuals served as controls.

Selection of DNA Samples for Screening

DNA samples from two affected members and two unaffected members of the family were selected for screening of the positional candidate genes. A DNA sample from an individual

who was not related to any of the members of the family or their spouses was also screened as a control specimen. If a sequence variant was identified in both affected individuals, and in neither of the unaffected family members nor in the unrelated control sample, then the variant was studied in 25 additional pedigree members: 9 affecteds, 9 unaffecteds, and 7 spouses of unknown, although presumed unaffected, status.

PCR Amplification

Each exon of the 35 genes selected for screening, including the intron/exon boundaries, was amplified using custom designed primers (Primer 3; <http://primer3.sourceforge.net>; sequences available upon request). Each 25 μ l reaction contained 50 mM Tris-HCl (pH 9.0, 25°C), 20 mM NH_4Cl , 2.5 mM MgSO_4 , 200 μ M of each dNTP, 0.5 M Betaine, 150 mM Trehalose, 0.002% Tween-20, 0.12 μ M of each primer, 0.5 units of REDTaq Genomic DNA Polymerase (Sigma-Aldrich, St. Louis, MO), and approximately 60 ng of genomic DNA. Thermal cycling was performed in an iCycler Thermal Cycler (Bio-Rad, Hercules, CA).

DNA Sequencing

Purification of the PCR products was achieved by incubating 15–30 ng DNA with 5 U Exonuclease I and 0.5 U Shrimp Alkaline Phosphatase (USB Corp., Cleveland, OH) for 15 minutes at 37°C. After inactivation of the nucleases for 15 minutes at 80°C, sequencing reactions were performed by the addition of 2 μ l BigDye Terminator Mix v3.1 (Applied Biosystems, Foster City, CA), 2 μ l of SeqSaver (Sigma-Aldrich) and 0.2 μ l primer (10 pM/ μ l). Samples were denatured at 96°C for 2 minutes, then cycled 25 times at 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 minutes. Unincorporated nucleotides were removed using the CleanSeq reagent and an SPRI plate (Agencourt Bioscience Corporation, Beverly, MA), resuspended in 0.1mM EDTA and then analyzed on an ABI-3100 Genetic Analyzer (Applied Biosystems). Nucleotide sequences, read manually and with Mutation Surveyor v2.2 (Softgenetics, State College, PA), were compared with the published cDNA sequences for each gene (Table 1).

Statistical Analysis

For sequence variants that demonstrated segregation with the PPCD phenotype in the pedigree, linkage analysis was performed using the gamete competition option of the statistical genetics software package Mendel 8.0 (www.genetics.ucla.edu/software/mendel).¹¹ Gamete competition is a generalization of the TDT test for linkage. This statistical analysis can make full use of all the information in extended pedigrees and is much less sensitive to missing data than the standard TDT test.¹²

CEPH HapMap data for the SNPs rs6111803, rs1053834 and rs1053839 was downloaded from the HapMap web site and analyzed using Haploview.^{13, 14} Linkage disequilibrium in the CEPH HapMap trios is reported as r^2 .

RESULTS

Screening of the 26 Positional Candidate Genes in the Common PPCD1 Interval

Sequence variants were identified in 19 of the 26 genes located within the common PPCD1 support interval (Table 1). Fifty-one variants were identified in total, 32 of which were previously identified SNPs and 19 of which were novel variants. Initial screening identified three previously identified SNPs and one novel sequence variant that demonstrated potential segregation with PPCD in the pedigree. For each of these four SNPs, the minor frequency allele was present in both affected family members, absent in both unaffected family members, and

absent in the unrelated control sample. These four SNPs were then screened in additional pedigree members.

The first of these SNPs, c.327C>A (p.Thr109Thr; rs6111803) in *OVOL2*, was heterozygous in all 11 affected individuals screened, and homozygous in the major allele in all 11 unaffected family members and the 7 spouses. The minor allele of this SNP was found in 9% (11/118) of control individuals (all heterozygous) and in 20% of the Caucasian individuals from the CEPH collection in the International HapMap project (<http://www.hapmap.org>).

The second and third of these SNPs were identified in the same gene, *C20orf79*: c.255C>T (p.Thr85Thr; rs1053834) and c.295C>T (p.Pro99Ser; rs1053839). Both of these SNPs were heterozygous in 10 of the 11 affected individuals screened and homozygous in the minor allele in the eleventh. These two SNPs were homozygous in the major allele in all 11 unaffected family members and 6 of the 7 spouses; they were both heterozygous in the seventh spouse. The minor alleles of these two SNPs, rs1053834 and rs1053839, were found in 31% (31/100) of control individuals, slightly higher than their rates in Caucasian individuals from the CEPH collection in the International HapMap project, 23% and 24%, respectively.

The novel sequence variant was identified in *RPS19P1*: c.167G>A (p.Arg56Gln; RPSnovel). This novel SNP followed exactly the same pattern of heterozygosity and homozygosity in the individuals genotyped as the two SNPs in *C20orf79*. The minor allele of this novel SNP was found in 17% (17/100) of controls individuals.

Screening of Seven Additional Genes within the CHED1 Interval

Sequence variants were identified in 6 of the 9 genes located within the CHED1 interval but outside of the common PPCD1 interval (Table 1). Twelve variants were identified in total, 11 of which were previously identified SNPs. The minor allele of each of these sequence variants was present in either one or both of the unaffected family members that were screened. No further screening was performed with these 12 SNPs.

Gamete Competition: Testing for Linkage

We tested for distorted transmission of SNPs rs6111803, rs1053834, rs1053839 and RPSnovel. As expected, the linkage between each SNP and the disease allele was highly statistically significant with a p-value $< 10^{-6}$ for each SNP (Table 2). The results showed that the minor allele was much more likely to be transmitted to affected individuals for each SNP, with the ratio of transmission likelihoods greater than 400 to 1.

Estimation of Haplotype Frequencies

The haplotype upon which the disease causing mutation is segregating was found to have a population frequency of 4.2% in the CEPH HapMap trios (Figure 2). The linkage disequilibrium (LD) between the three PPCD associated SNPs genotyped by the HapMap consortium is shown in Figure 3. The SNPs rs1053834 and rs1053839 are in complete LD ($r^2 = 1.0$) in the CEPH HapMap trios. There is no LD between rs6111803 and either rs1053834 or rs1053839 ($r^2 = 0.022$ and 0.028 , respectively), which is expected considering that the distance between rs6111803 and the other two SNPs is approximately 770Kbp. The haplotype seen in the family extends much further than is found in the general European population.

DISCUSSION

We believe that the same genetic basis for PPCD1 is shared by the four families in which statistical analysis has demonstrated linkage to the pericentromeric region of chromosome 20. Based on this assumption, the pathogenic mutation(s) should lie within one of the 26 genes

mapped to the 2.4 cM region that is common to the linked support intervals of each of the four families. Sequencing of these 26 positional candidate genes in one of the four families revealed four sequence variants that segregated with the affected phenotype. However, each variant was also identified in at least 9% of control individuals that were screened. Since PPCD is highly penetrant, none of these four coding region variants with relatively high minor allele frequencies is a candidate for the causal variant for PPCD. Instead, it is likely that each of these four sequence variants is linked to a much less frequent, pathogenic variant. However, the demonstration of linkage of each of these SNPs with the disease allele provides additional support for our previous conclusion that the pathogenic mutation in the family we report lies within the refined PPCD1 interval.⁶

Due to the very similar linkage support intervals and clinical features shared by CHED1 and PPCD, we believe CHED1 may simply represent a severe, early-onset form of PPCD. Therefore, we also screened the 9 genes located within the CHED1 interval but immediately telomeric or centromeric to the common PPCD1 interval. Again, no presumed pathogenic coding region variants were identified. Thus, we are now able to exclude a coding region mutation in the CHED1 interval, which contains the common support interval for PPCD1. Recently, Heon and colleagues also reported the absence of coding region variants in three of the 35 genes mapped to the common PPCD1/CHED1 interval, *RBBP9*, *ZNF133* and *SLC24A3*, in several members of the original PPCD family linked to the pericentromeric region of chromosome 20.¹⁵

There are several explanations for the absence of coding region mutations in the common support interval for PPCD1: the presence of a mutation in a yet to be described coding region; locus heterogeneity with a different genetic basis for PPCD1 in the family that we report than in one of the families reported by Gwilliam and colleagues^{5, 6}; an error in the assignment of the boundaries for the common support interval; a deletion or a genomic rearrangement involving a PCR primer binding site; or a non-coding region mutation in the common support interval for PPCD1. Although it is possible that PPCD1 is caused by a mutation in a yet to be described gene, 6 years after the publication of the DNA sequence of human chromosome 20, gene annotation is essentially complete.¹⁶ Locus heterogeneity for PPCD on chromosome 20 also seems unlikely as it violates the general scientific rule of parsimony: the simplest explanation is best. Since the support intervals on chromosome 20 overlap in the published PPCD linkage studies, the great weight of the statistical evidence is that a causal variant exists in the common PPCD1 interval.

An error in the assignment of the boundaries of the support interval for PPCD1 reported by Gwilliam and colleagues⁵ could explain the absence of a presumed pathogenic mutation in the genes located between the flanking markers D20S48 and D20S139. In fact, such an error in fine mapping the candidate gene interval for another corneal dystrophy, Schnyder corneal dystrophy (SCD; MIM # 121800), was recently reported. After a screening of the coding regions of the genes mapped to the candidate interval failed to identify the genetic basis of SCD, mutations were identified in the *UBIAD1* gene, located 368 Kbp outside of the candidate gene interval.^{17–20}

As it remains most likely that the genetic basis for PPCD1 is shared by the four families in which linkage to the pericentromeric region of chromosome 20 has been demonstrated, we believe that the most plausible explanations for our not identifying a coding region mutation in the common PPCD1 support interval are a coding region deletion or a noncoding mutation within a promotor or intron that affects transcript expression and/or integrity. An alternative mechanism of overlooking a causative mutation could be due to the presence of a variant within the PCR primer binding site that would allow only the normal exon to be amplified. Likewise, a deletion that involves a PCR binding site and a portion or all of an exon would not result in

an amplified product and therefore would not be detected by targeted sequencing. Instead, only the wild-type allele, present in the hemizygous state, would be amplified. While the identification of a heterozygous polymorphism in an exon would indicate that both alleles were present, it is possible that PPCD1 is associated with a deletion or genomic rearrangement involving a PCR binding site for one of the >150 exons of the 35 positional candidate genes that we amplified in which no heterozygous SNPs were identified.

Presumed pathogenic intronic variants have been previously associated with two other corneal dystrophies: fleck corneal dystrophy (in a single family) and autosomal recessive congenital hereditary corneal endothelial dystrophy (CHED2) (in three families).^{21–23} Each of these intronic changes was predicted to inactivate the splice acceptor site and therefore result in skipping of the adjacent exon, although none were actually demonstrated to result in a truncated protein product.^{21–23} Although we screened the intron-exon boundaries of each of the 35 genes in the common PPCD1/CHED1 interval and did not identify any sequence variants involving the splice donor or acceptor sites, we acknowledge that we may have missed an intronic variant positioned further away from the intron-exon boundary.

The failure of our candidate gene screening strategy to identify a likely causative mutation provides the rationale for a more intensive approach to searching for the PPCD-related mutation. To identify a pathogenic deletion, genomic rearrangement or intronic mutation in the common PPCD1 interval, we plan to perform large-scale re-sequencing of the entire 2.4 cM region between markers D20S182 and D20S139. While quantitative real-time PCR (qRT-PCR) is a useful technique to detect a deletion involving a portion or all of the coding region of a gene of interest, as the authors have employed previously,²⁴ it would not be able to detect a pathogenic intronic variant. The recent development of large-scale re-sequencing strategies, by contrast, make sequencing of the entire 2.4 cM common PPCD1 interval possible, and thus will permit identification of any variation in the nucleotide sequence in this interval, including deletions, insertions, substitutions and genomic rearrangements. Given the demonstration of significant linkage of four SNPs with the disease allele in the common PPCD1 support interval, providing compelling evidence for the localization of the pathogenic mutation to this region, we believe that re-sequencing the entire common PPCD1 support interval will lead to the identification of the genetic basis of PPCD1.

ACKNOWLEDGEMENTS

Support provided by NIH K08 EY016079 (AJA).

REFERENCES

1. Liskova P, Tuft SJ, Gwilliam R, et al. Novel mutations in the ZEB1 gene identified in Czech and British patients with posterior polymorphous corneal dystrophy. *Hum Mutat* 2007;28:638. [PubMed: 17437275]
2. Aldave AJ, Yellore VS, Yu F, et al. Posterior polymorphous corneal dystrophy is associated with TCF8 gene mutations and abdominal hernia. *Am J Med Genet A* 2007;143:2549–2556. [PubMed: 17935237]
3. Krafchak CM, Pawar H, Moroi SE, et al. Mutations in TCF8 Cause Posterior Polymorphous Corneal Dystrophy and Ectopic Expression of COL4A3 by Corneal Endothelial Cells. *Am J Hum Genet* 2005;77:694–708. [PubMed: 16252232]
4. Heon E, Mathers WD, Alward WL, et al. Linkage of posterior polymorphous corneal dystrophy to 20q11. *Hum Mol Genet* 1995;4:485–488. [PubMed: 7795607]
5. Gwilliam R, Liskova P, Filipec M, et al. Posterior polymorphous corneal dystrophy in Czech families maps to chromosome 20 and excludes the VSX1 gene. *Invest Ophthalmol Vis Sci* 2005;46:4480–4484. [PubMed: 16303937]

6. Yellore VS, Papp JC, Sobel E, et al. Replication and refinement of linkage of posterior polymorphous corneal dystrophy to the posterior polymorphous corneal dystrophy 1 locus on chromosome 20. *Genet Med* 2007;9:228–234. [PubMed: 17438387]
7. Toma NM, Ebenezer ND, Inglehearn CF, et al. Linkage of congenital hereditary endothelial dystrophy to chromosome 20. *Hum Mol Genet* 1995;4:2395–2398. [PubMed: 8634716]
8. Cibis GW, Krachmer JA, Phelps CD, Weingeist TA. The clinical spectrum of posterior polymorphous dystrophy. *Arch Ophthalmol* 1977;95:1529–1537. [PubMed: 302697]
9. Krachmer JH. Posterior polymorphous corneal dystrophy: a disease characterized by epithelial-like endothelial cells which influence management and prognosis. *Trans Am Ophthalmol Soc* 1985;83:413–475. [PubMed: 3914130]
10. McCartney AC, Kirkness CM. Comparison between posterior polymorphous dystrophy and congenital hereditary endothelial dystrophy of the cornea. *Eye* 1988;2(Pt 1):63–70. [PubMed: 3261696]
11. Lange K, Sinsheimer JS, Sobel E. Association testing with Mendel. *Genet Epidemiol* 2005;29:36–50. [PubMed: 15834862]
12. Sinsheimer JS, Blangero J, Lange K. Gamete-competition models. *Am J Hum Genet* 2000;66:1168–1172. [PubMed: 10712230]
13. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265. [PubMed: 15297300]
14. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851–861. [PubMed: 17943122]
15. Hosseini SM, Herd S, Vincent AL, Heon E. Genetic analysis of chromosome 20-related posterior polymorphous corneal dystrophy: genetic heterogeneity and exclusion of three candidate genes. *Mol Vis* 2008;14:71–80. [PubMed: 18253095]
16. Deloukas P, Matthews LH, Ashurst J, et al. The DNA sequence and comparative analysis of human chromosome 20. *Nature* 2001;414:865–871. [PubMed: 11780052]
17. Orr A, Dube MP, Marcadier J, et al. Mutations in the UBIAD1 gene, encoding a potential prenyltransferase, are causal for Schnyder crystalline corneal dystrophy. *PLoS ONE* 2007;2:e685. [PubMed: 17668063]
18. Weiss JS, Kruth HS, Kuivaniemi H, et al. Genetic analysis of 14 families with Schnyder crystalline corneal dystrophy reveals clues to UBIAD1 protein function. *Am J Med Genet A* 2008;146:271–283. [PubMed: 18176953]
19. Weiss JS, Kruth HS, Kuivaniemi H, et al. Mutations in the UBIAD1 gene on chromosome short arm 1, region 36, cause Schnyder crystalline corneal dystrophy. *Invest Ophthalmol Vis Sci* 2007;48:5007–5012. [PubMed: 17962451]
20. Yellore VS, Khan MA, Bourla N, et al. Identification of mutations in UBIAD1 following exclusion of coding mutations in the chromosome 1p36 locus for Schnyder crystalline corneal dystrophy. *Mol Vis* 2007;13:1777–1782. [PubMed: 17960116]
21. Li S, Tiab L, Jiao X, et al. Mutations in PIP5K3 are associated with Francois-Neetens mouchettee fleck corneal dystrophy. *Am J Hum Genet* 2005;77:54–63. [PubMed: 15902656]
22. Sultana A, Garg P, Ramamurthy B, et al. Mutational spectrum of the SLC4A11 gene in autosomal recessive congenital hereditary endothelial dystrophy. *Mol Vis* 2007;13:1327–1332. [PubMed: 17679935]
23. Vithana EN, Morgan P, Sundaresan P, et al. Mutations in sodium-borate cotransporter SLC4A11 cause recessive congenital hereditary endothelial dystrophy (CHED2). *Nat Genet* 2006;38:755–757. [PubMed: 16767101]
24. Yellore VS, Sonmez B, Chen MC, et al. An unusual presentation of macular corneal dystrophy associated with uniparental isodisomy and a novel Leu173Pro mutation. *Ophthalmic Genet* 2007;28:169–174. [PubMed: 17896316]

Figure 1.

Ideogram of the pericentromeric region of chromosome 20, showing the reported candidate gene intervals for PPCD1 and CHED1 and the positions of the markers that define each candidate gene region. The region of chromosome 20 that lies within all three regions to which PPCD has been mapped, the common PPCD1 interval, is defined by the markers D20S182 and D20S139. The four DNA sequence variants identified within the 2.4 cM common PPCD1 interval that demonstrated complete segregation with the affected phenotype in a family previously linked to the PPCD1 locus are shown.

rs6111803 rs1053834 rs1053839			
Haplotype			Frequency
G	C	C	0.758
G	T	T	0.125
T	C	C	0.075
T	T	T	0.042

Figure 2.

European population haplotype frequencies of three associated SNPs genotyped in the CEPH HapMap trios. The rare T-T-T haplotype (rs6111803, rs1053834, rs1053839) that is associated with PPCD has a population frequency of 0.042 (4.2%).

Figure 3.

European population linkage disequilibrium (LD) of the three associated SNPs genotyped in the CEPH HapMap trios. LD is given as the R-squared (r^2) between SNPs. The SNPs rs1053834 and rs1053839 are in complete LD ($r^2 = 1$) in the CEPH HapMap trios. There is no LD between rs6111803 and either rs1053834 or rs1053839 ($r^2 = 0.022$ and 0.028 , respectively).

Table 1
Sequence variants identified in candidate genes within the common PPCD1 and CHED 1 regions.

Candidate interval	Gene (refSeq)	Nucleotide Change	AA Change	refSNP ID	Family Members	
					Affected (N=2)	Unaffected (N=2)
CHED1	<i>PCSK2</i> (NM_002594)	No Changes				
	<i>TCTE1P</i> (NG_001059)	c.149T>A	p.Leu50X	-	0	1 (homozygous)
	<i>BFSPI</i> (NM_001195)	c.1033G>A	p.Gly345Ser	rs6080719	2 (1 heterozygous)	2 (heterozygous)
		c.1500G>A	p.Ala500Ala	rs6136118	2 (1 heterozygous)	2 (heterozygous)
		c.1749A>G	p.Pro583Pro	rs6080718	2 (homozygous)	2 (homozygous)
		c.1926C>T	p.Thr642Thr	rs6080717	1 (heterozygous)	1 (heterozygous)
	<i>RPS27AP2</i> (NG_000995)	c.282G>A	p.Arg94Arg	rs2023224	2 (1 heterozygous)	1 (homozygous)
	<i>DSTN</i> (NM_006870)	No Changes				
	<i>RBBP1</i> (NM_004587)	c.1829T>A	p.Leu610His	rs6034867	2 (homozygous)	2 (homozygous)
		c.2297C>T	p.Ser766Leu	rs11960	2 (1 heterozygous)	2 (heterozygous)
		c.2672G>T	p.Arg891Leu	rs1132274	0	1 (heterozygous)
CHED1 and common PPCD1	<i>BANF2</i> (NM_178477)	c.233C>G	p.Thr78Ser	rs1053993	0	1 (heterozygous)
	<i>SNX5</i> (NM_152227)	c.543G>A	p.Glu181Glu	rs2273448	1 (heterozygous)	2 (heterozygous)
	<i>SNORD17*</i> (NR_003045)	98T>A	-	-	1 (heterozygous)	2 (heterozygous)
	<i>C20orf72</i> (NM_018152)	c.43A>T	p.Ser15Cys	rs11551768	2 (heterozygous)	0

Candidate interval	Gene (refSeq)	Nucleotide Change	AA Change	refSNP ID	Family Members	
					Affected (N=2)	Unaffected (N=2)
		c.794C>T	p.Thr265Ile	-	0	1 (heterozygous)
		c.846T>C	p.Asp282Asp	rs2845509	1 (heterozygous)	2 (heterozygous)
		c.348C>G	p.Val116Val	-	2 (homozygous)	2 (homozygous)
		c.362C>G	p.Thr121Arg	-	2 (homozygous)	2 (homozygous)
<i>PTMAP3</i> (NG_0011180)		c.415T>C	p.Leu139Leu	-	2 (homozygous)	2 (homozygous)
		c.553T>C	p.Cys185Arg	rs3748493	2 (heterozygous)	2 (homozygous)
		c.843C>T	p.Cys281Cys	rs3748492	2 (heterozygous)	2 (homozygous)
		c.327C>A	p.Thr109Thr	rs6111803	2 (heterozygous)	0
		No Changes				
		c.1199T>G	p.Val400Gly	rs1205193	2 (1 heterozygous)	2 (1 heterozygous)
		c.1437C>T	p.Pro479Pro	rs1205194	0	1 (heterozygous)
		c.1499C>T	p.Ala500Val	-	1 (heterozygous)	1 (heterozygous)
		c.2223T>C	p.Pro741Pro	rs2747404	2 (1 heterozygous)	2 (1 heterozygous)
		No Changes				
		c.89T>C	p.Leu30Pro	rs1883938	2 (1 heterozygous)	2 (homozygous)
		c.40T>C	p.Leu14Leu	rs6035051	2 (heterozygous)	2 (homozygous)
		c.801T>C	p.Cys267Cys	rs6035042	2 (1 heterozygous)	2 (heterozygous)
		c.1410G>C	p.Leu470Leu	rs6075337	2 (1 heterozygous)	1 (heterozygous)
		No Changes				

Candidate interval	Gene (refSeq)	Nucleotide Change	AA Change	refSNP ID	Family Members	
					Affected (N=2)	Unaffected (N=2)
	<i>POLR3F</i> (NM_006466)	c.840A>G	p.Thr280Thr	rs1055171	2 (1 heterozygous)	2 (heterozygous)
	<i>RPL21P3</i> (NG_000976)	c.346delA	p.Arg116GlyfsX28	-	2 (1 heterozygous)	2 (heterozygous)
	<i>RBBP9</i> (NM_006606)	No Changes				
	<i>RPS19P1</i> (NG_001295)	c.136C>G	p.Pro46Ala	rs2009092	2 (1 heterozygous)	2 (heterozygous)
		c.167G>A	p.Arg56Gln	-	2 (heterozygous)	0
	<i>SEC23B</i> (NM_032985)	c.490G>A	p.Val164Met	-	0	1 (heterozygous)
		c.1198T>C	p.Phe400Leu	-	0	1 (heterozygous)
		c.1276G>A	p.Val426Ile	-	1 (heterozygous)	1 (heterozygous)
		c.1467C>G	p.His489Gln	rs2273526	1 (heterozygous)	2 (heterozygous)
	<i>LOC388789</i> (XM_939954)	c.48A>G	p.Arg16Arg	rs12624935	1 (heterozygous)	2 (heterozygous)
	<i>DTD1</i> (NM_080820)	No Changes				
	<i>DUXAP7</i> (NG_004846)	c.447A>C	p.Gln149His	rs6132094	2 (1 heterozygous)	2 (heterozygous)
		c.664A>G	p.Thr222Ala	rs6081271	2 (homozygous)	2 (homozygous)
		c.665C>T	p.Thr222Ile	rs6081270	1 (heterozygous)	1 (heterozygous)
		c.715delA	p.Ile239fsX	-	0	1 (heterozygous)
		c.716T>A	p.Ile239Lys	-	2 (heterozygous)	2 (heterozygous)
		c.719delA	p.Lys240LysfsX39	-	0	1 (heterozygous)
		c.720delA	p.Lys240LysfsX39	-	2 (heterozygous)	2 (heterozygous)
		c.721insA	p.Glu241ArgfsX25	-	2 (heterozygous)	2 (heterozygous)

Candidate interval	Gene (refSeq)	Nucleotide Change	AA Change	refSNP ID	Family Members	
					Affected (N=2)	Unaffected (N=2)
		c.726_730dupGAAAA	p.Arg244ArgfsX37	-	2 (1 heterozygous)	2 (heterozygous)
		c.726_730dupGAAAA [3]	p.Arg244ArgfsX25	-	1 (heterozygous)	2 (heterozygous)
		c.943G>C	p.Val315Leu	rs6132093	2 (1 heterozygous)	2 (heterozygous)
		c.1124T>C	p.Val375Ala	rs6081268	2 (1 heterozygous)	2 (heterozygous)
		c.1156C>T	p.Pro386Ser	rs6075382	2 (1 heterozygous)	2 (heterozygous)
	LOC100132358 (NC_000020)	No Changes				
		No Changes				
	HSPC072 (AF161557)	No Changes				
		No Changes				
	LOC100128496 (XM_001725705)	c.138C>G	p.Leu46Leu	rs56743271	2 (heterozygous)	1 (heterozygous)
	C20orf79 (NM_178483)	c.255C>T	p.Thr85Thr	rs1053834	2 (heterozygous)	0
		c.295C>T	p.Pro99Ser	rs1053839	2 (heterozygous)	0
	LOC100130062 (XM_001718559)	No Changes				
		No Changes				
	SLC24A3 (NM_020689)	c.52_54delCGC	p.Arg18del	-	2 (heterozygous)	2 (heterozygous)
		c.163G>A	p.Val55Ile	rs1569767	1 (heterozygous)	2 (1 heterozygous)
		c.369G>A	p.Ala123Ala	rs3790261	0	1 (homozygous)
		c.480G>A	p.Ser160Ser	rs17293432	1 (heterozygous)	1 (heterozygous)
		c.639G>A	p.Leu213Leu	rs3790278	2 (1 heterozygous)	2 (1 heterozygous)
		c.654T>C	p.Ile218Ile	rs3790279	2 (1 heterozygous)	2 (1 heterozygous)
	LOC100130264 (XM_001717962)	c.472T>C	p.Tyr458His	rs58486151	1 (heterozygous)	1 (homozygous)
CHED1	LOC100130408 (XM_001724726)	c.87C>T	p.Ala29Ala	rs910983	0	2 (heterozygous)

Candidate interval	Gene (refSeq)	Nucleotide Change	AA Change	refSNP ID	Family Members	
					Affected (N=2)	Unaffected (N=2)
		c.504A>C	p.Thr168Thr	rs708983	2 (1 heterozygous)	2 (1 heterozygous)
	<i>RPL12L3</i> (NG_000945)	No Changes				

* SNORD17 is a non-coding microRNA, so no protein is formed

Gamete competition analyses for SNPs rs6111803, RPSnovel, rs1053834 and rs1053839 using Mendel 8.0. The table includes the p-value for the association of each SNP to PPCD. Also listed for each SNP is data for the least likely allele to be transmitted to affecteds and the most likely allele. Tau is the relative likelihood of this transmission.

SNP	P-VALUE	LEAST TRANSMITTED			MOST TRANSMITTED		
		TAU	FREQ	ALLELE	TAU	FREQ	ALLELE
rs6111803	8.6×10^{-7}	1.00	0.9435	1	428.79	0.0565	2
RPSnovel	2.5×10^{-7}	1.00	0.8881	1	452.23	0.1119	2
rs1053834	2.5×10^{-7}	1.00	0.8881	1	452.23	0.1119	2
rs1053839	2.5×10^{-7}	1.00	0.8881	1	452.23	0.1119	2