SUPPLEMENTARY METHODS

Novel genetic causes for cerebral visual impairment

Daniëlle G.M. Bosch,[1,2,3,4] F. Nienke Boonstra,[2,4] Nicole de Leeuw,[1] Rolph Pfundt,[1] Willy M. Nillesen,[1] Joep de Ligt,[1,3,5] Christian Gilissen,[1,3] Shalini Jhangiani,[6] James R. Lupski,[6,7,8,9] Frans P.M. Cremers,[1,3] Bert B. A. de Vries.[1,4]*

[1]*Department of Human Genetics, Radboud university medical center, Nijmegen, the Netherlands*

[2]*Bartiméus, Institute for the Visually Impaired, Zeist, the Netherlands*

[3]*Radboud Institute for Molecular Life Sciences, Radboud university medical center, Nijmegen, the Netherlands*

[4]*Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud university medical center, Nijmegen, the Netherlands*

[5]*Hubrecht Institute-KNAW, University Medical Centre Utrecht, CancerGenomics.nl, Utrecht, the Netherlands*

[6]*Human Genome Sequencing Center, Baylor College of Medicine, Houston, USA*

[7]*Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, USA*

[8]*Texas Children's Hospital, Houston, USA*

[9]*Departments of Pediatrics, Baylor College of Medicine, Houston, USA*

SUPPLEMENTARY METHODS

Table of contents

SUPPLEMENTARY METHODS

Subjects

Twenty-five patients with CVI and a visual acuity ≤0.3 were included. The history was taken, and no potential risk factors for CVI, such as preterm birth, perinatal problems or hydrocephalus, were present. CVI was diagnosed when no other ocular diagnosis could explain the visual impairment. The diagnosis was made by a pediatric ophthalmologist after ophthalmological examination. This examination included visual acuity testing, crowding measurements, the examination of eye movements, fixation and oculomotor abnormalities, visual field measurements, slit lamp examination, and funduscopy. In patient with a lower cognitive level visual acuity was measured with forced preferential looking by using Teller acuity cards (TAC).[1] In patients with higher developmental level tests based on object recognition, such as the LH test or Landolt C test were used.[2] Visual fields were measured by using a confrontational method with white Stycar balls on a stick.[3]

In addition, the patients were clinically examined by a clinical geneticist and pathogenic chromosomal aberrations were excluded by array CGH. WES was performed in the patients and their parents, and except for patient 11, the parents were unrelated.

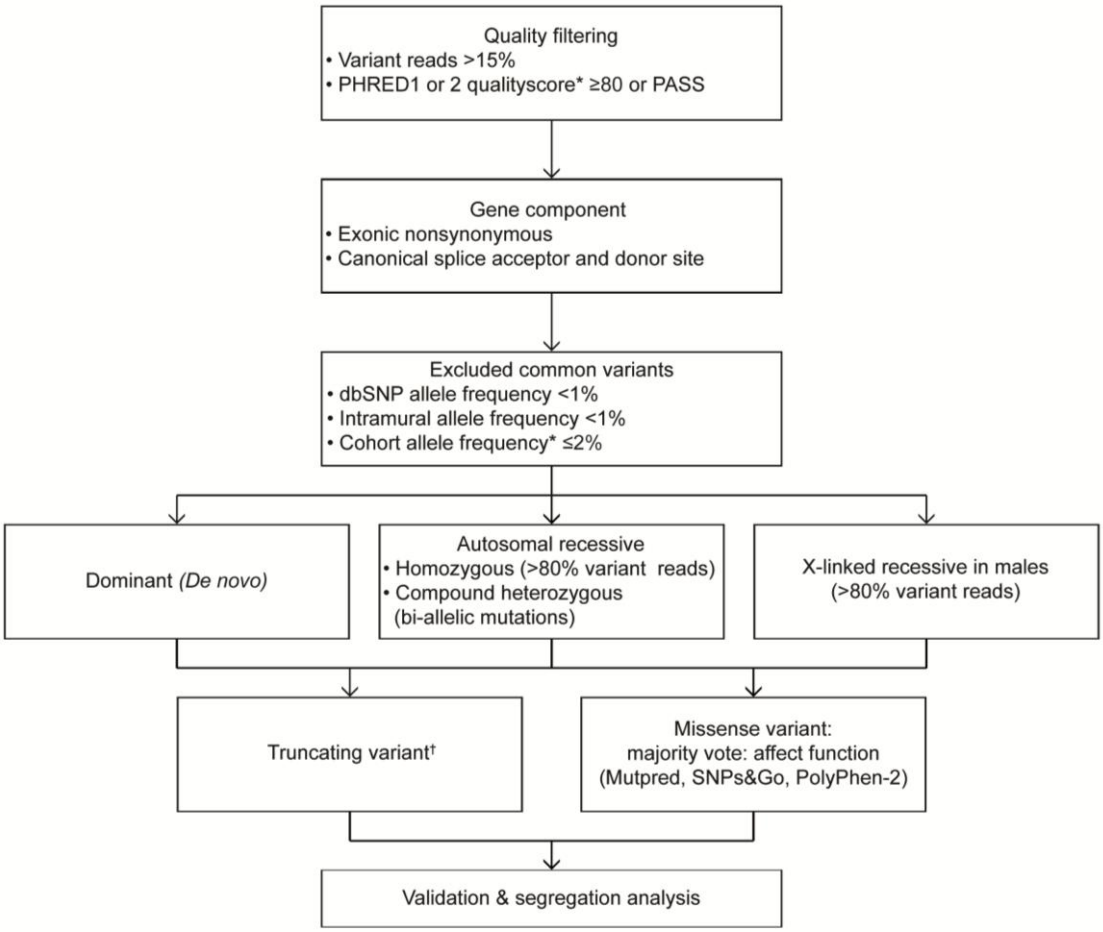
Whole exome sequencing

In all 25 patients WES was performed using the trio approach (patient-parents).[4] In 11 trios (patients 1-11) WES was performed on an Illumina HiSeq platform with the Agilent SureSelect All Exon V4 reagent for target enrichment (Agilent Technologies, Inc., Santa Clara, CA, USA). Alignment and variants and indels were called with Burrows-Wheeler Aligner, BWA, and Genome Analysis Toolkit, GATK.[5,6] WES was executed in 10 trios (patients 12-21) at the Baylor-Hopkins Center for Mendelian Genomics on a Illumina Hiseq 2000 platform (Illumina, San Diego, CA, USA) and in three trios (patients 22, 23 and 24) on a Solid 5500 XL platform (Life Technologies, Carlsbad, CA, USA) of which the methods have been reported previously.[7,8] In one trio (patient 25) the WES was also performed on a Solid 5500 XL platform, but the Agilent XL SureSelect All Exon 50Mb reagent was used for target enrichment (Agilent Technologies, Inc., Santa Clara, CA, USA) and the

variants and indels were called with Lifescope v2.1 (Life Technologies, Carlsbad, CA, USA). In all 25 patients and their parents the sequence reads were mapped and aligned to the USCS genome Browser GRCh37/hg19 Human Genome Reference Assembly.

Variant prioritization and validation

After quality filtering (variant reads >15%) local and global variants (≥1% allele occurrence in intramural database, dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) or Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (http://evs.gs.washington.edu/EVS/) were excluded (Supplementary Figure S1). For the 10 patients for which WES was performed at the Baylor-Hopkins Center for Mendelian Genomics, the data were also filtered for cohort allele frequency ≤2% (cohort consisted of patients with various disorders and their healthy parents, n=200) and an additional quality score was used (PHRED ≥80 or passed). The exonic nonsynonymous and canonical splice site variants were selected for further analysis. A *de novo* analysis was performed for all trios. Furthermore, the results were analyzed for homozygous variants (>80% variant reads), compound heterozygous (two or more variants present in one gene) and hemizygous variants in males (X-chromosome, >80% variant reads). When the autosomal recessive variants were present on one allele, the variants were excluded (based on the raw data (BAM-file) of a patient and its parents). Truncating variants, consisting of frameshift, nonsense or splice site variants (assessed using Interactive Biosoftware Alamut version 2.3 rev2), and missense variants predicted to affect function (majority vote of Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/), MutPred (http://mutpred.mutdb.org/) and SNPs&GO (http://snps.path.uab.edu/snps-and-go/pages/method.html)) were validated by Sanger sequencing in patient and parents.[9-11] For compound heterozygous variants at least one of the variants should be predicted to be protein truncating, or, when concerning missense variants, both variants should be predicted to affect function. For the maternally inherited variants on the X-chromosome in males and the autosomal recessive variants further segregation analysis in the family was undertaken, whenever possible.



Supplementary Figure S1: Flow chart of variant prioritization

*Only used for the 10 patients in whom WES was performed at the Baylor-Hopkins Center for Mendelian Genomics. †Frameshift, nonsense or splice site variant.

Gene classification

For the validated variants in genes that were previously indicated in the Online Mendelian Inheritance in Man (OMIM) database (www.omim.org) the inheritance pattern and the variant type were compared. Furthermore, a phenotypic comparison of the patient and the reported individuals in literature was performed by a clinical geneticist and assessment was made whether the phenotype showed similarities. If the phenotype was distinct the gene was classified as unlikely to be causative. When the phenotype showed similarities and the patients reported did not have CVI, the gene was classified as a candidate gene for CVI. Genes that were previously reported to be involved in the

pathogenesis of CVI were classified as "known CVI-associated gene". The remaining (non-OMIM disease related) validated variants were classified based on the previously reported method by De Ligt et al. and Gilissen et al.[12,13] In brief, the variants were scored for their functional relevance on four items. First, it was assessed whether the variant was disruptive or whether the missense variants involved a conserved nucleotide (phyloP >3.5). Subsequently, the list of 525 genes known to be associated with ID from Gilissen et al. was loaded into ToppGene to select Gene Ontology (GO) and Mouse Phenotype (MP) terms with a significant enrichment (FDR<0.05 by Benjamini Hochberg method) (https://toppgene.cchmc.org/, January 2015).[13,14] When there was overlap between the enriched terms and the GO- or MP-terms for the genes identified in the present study the variant scored positive. Finally, it was assessed whether the gene was expressed during brain development in the Human Brain Transcriptome (http://hbatlas.org/).[15] A gene was considered to be expressed when the Log2 intensity was ≥6 for at least three periods of the developing brain (<20 years of age). When the detected variant and/or the gene scored positive for at least two items (conservation/disruptive, brain expression, GO- and/or MP term) the aberrant gene was classified as a possible candidate gene for CVI. Otherwise the aberrant gene was classified as unlikely to be causative for CVI.

Stringent criteria

For the *de novo* and X-linked missense variants, the variants with an allele frequency ≥0.1% in intramural database, dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) or Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (http://evs.gs.washington.edu/EVS/) were excluded. This frequency of 0.1% is, with a wide margin, based on the estimated incidence of 0.0225% of CVI in children in the Netherlands.[16] In addition, for loss of function variants, the ExAC browser (Exome Aggregation Consortium (ExAC), Cambridge, MA (http://exac.broadinstitute.org) [(June, 2015 accessed]) was used to obtain all the loss of function variants in the identified gene. Genes with a truncating allele frequency ≥0.1% (*de novo* or X-linked variants) or ≥1.0% (autosomal recessive variants) were excluded.

References

1. Watson T, Orel-Bixler D, Haegerstrom-Portnoy G: VEP vernier, VEP grating, and behavioral grating acuity in patients with cortical visual impairment. *Optom Vis Sci* 2009; **86:** 774-780.

2. Hyvarinen L, Nasanen R, Laurinen P: New visual acuity test for pre-school children. *Acta Ophthalmol (Copenh)* 1980; **58:** 507-511.

3. Sheridan MD: The STYCAR graded-balls vision test. *Dev Med Child Neurol* 1973; **15:** 423-432.

4. Vissers LE, de Ligt J, Gilissen C *et al*: A de novo paradigm for mental retardation. *Nat Genet* 2010; **42:** 1109-1112.

5. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25:** 1754-1760.

6. 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:** 1297-1303.

7. Lupski JR, Gonzaga-Jauregui C, Yang Y *et al*: Exome sequencing resolves apparent incidental findings and reveals further complexity of SH3TC2 variant alleles causing Charcot-Marie-Tooth neuropathy. *Genome Med* 2013; **5:** 57.

8. Bosch DG, Boonstra FN, Gonzaga-Jauregui C *et al*: *NR2F1* mutations cause optic atrophy with intellectual disability. *Am J Hum Genet* 2014; **94:** 303-309.

9. Adzhubei IA, Schmidt S, Peshkin L *et al*: A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7:** 248-249.

10. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R: Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat* 2009; **30:** 1237-1244.

11. Li B, Krishnan VG, Mort ME *et al*: Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 2009; **25:** 2744-2750.

12. de Ligt J, Willemsen MH, van Bon BW *et al*: Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012; **367:** 1921-1929.

13. Gilissen C, Hehir-Kwa JY, Thung DT *et al*: Genome sequencing identifies major causes of severe intellectual disability. *Nat New Biol* 2014; **511:** 344-347.

14. Chen J, Bardes EE, Aronow BJ, Jegga AG: ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009; **37:** W305-311.

15. Kang HJ, Kawasawa YI, Cheng F *et al*: Spatio-temporal transcriptome of the human brain. *Nat New Biol* 2011; **478:** 483-489.

16. Boonstra N, Limburg H, Tijmes N, van Genderen M, Schuil J, van Nispen R: Changes in causes of low vision between 1988 and 2009 in a Dutch population of children. *Acta Ophthalmol* 2012; **90:** 277-286.