Increasing the Yield in Targeted Next-Generation Sequencing by Implicating CNV Analysis, Non-Coding Exons and the Overall Variant Load: The Example of Retinal Dystrophies

Tobias Eisenberger1, Christine Neuhaus1, Arif O. Khan2, Christian Deckert1, Markus N. Preising3, Christoph Friedburg3, Anika Bieg1, Martin Gliem4, Peter Charbel Issa4, Frank G. Holz4, Shahid M. Baig5, Yorck Hellenbroich6, Alberto Galvez7, Konrad Platzer6, Bernd Wollnik7,8,9, Nadja Laddach10, Saeed Reza Ghaffari11, Maryam Rafati12, Elke Botzenhart13, Sigrid Tinschert14,15, Doris Börger16, Axel Bohring17, Julia Schrem7,8, Stefani Körtge-Jung18, Chayim Schell-Apacik19, Khadijah Bakur20, Jumana Y. Al-Aama20, Teresa Neußann21, Peter Herkenrath22, Gudrun Nürnberg8,23, Peter Nürnberg8,23, John S. Davis24, Andreas Gal25, Carsten Bergmann1,26, Birgit Lorenz3, Hanno J. Bolz1,27

1 Bioscientia Center for Human Genetics, Ingelheim, Germany, 2 Department of Ophthalmology, King Khaled Eye Specialist Hospital, Riyadh, Saudi Arabia, 3 Department of Ophthalmology, Justus-Liebig-University Giessen, University Hospital Giessen and Marburg GmbH, Giessen Campus, Giessen, Germany, 4 Department of Ophthalmology, University of Bonn, Bonn, Germany, 5 Human Molecular Genetics Laboratory, Health Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan, 6 Institute of Human Genetics, University of Lübeck, Lübeck, Germany, 7 Institute of Human Genetics, University Hospital of Cologne, Cologne, Germany, 8 Center for Molecular Medicine Cologne (CMM), University of Cologne, Cologne, Germany, 9 Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany, 10 MRC Holland, Amsterdam, The Netherlands, 11 Comprehensive Genetic Center, Tehran University of Medical Sciences, Tehran, Iran, 12 Avicenna Biotechnology Research Institute, Tehran, Iran, 13 Pränatalzentrum Hamburg und Humangenetik, Hamburg, Germany, 14 Institute of Clinical Genetics, Technical University Dresden, Dresden, Germany, 15 Division of Human Genetics, Medical University Innsbruck, Innsbruck, Austria, 16 Humangenetik, Bremen, Germany, 17 Institute of Human Genetics, Westfälische Wilhelms-University, Münster, Germany, 18 Pränatal-Medizin und Genetik Düsseldorf, Düsseldorf, Germany, 19 Praxis für Humangenetik am DRK-Klinikum Westend, Berlin, Germany, 20 Princess Al Jawhara Alibrahim Center of Excellence in Research of Hereditary Disorders, King Abdulaziz University, Jeddah, Saudi Arabia, 21 Medizinisch Genetisches Zentrum, Munich, Germany, 22 Department of Pediatrics, University Hospital of Cologne, Cologne, Germany, 23 Cologne Center for Genomics and Center for Molecular Medicine, University of Cologne, Cologne, Germany, 24 Department of Ophthalmology, Zayed Military Hospital, Abu Dhabi, United Arab Emirates, 25 Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, 26 Center for Clinical Research, University Hospital of Freiburg, Freiburg, Germany

Abstract

Retinitis pigmentosa (RP) and Leber congenital amaurosis (LCA) are major causes of blindness. They result from mutations in many genes that has long hampered comprehensive genetic analysis. Recently, targeted next-generation sequencing (NGS) has proven useful to overcome this limitation. To uncover “hidden mutations” such as copy number variations (CNVs) and mutations in non-coding regions, we extended the use of NGS data by quantitative readout for the exons of 55 RP and LCA genes in 126 patients, and by including non-coding 5′ exons. We detected several causative CNVs which were key to the diagnosis in hitherto unsolved constellations, e.g. hemizygous point mutations in consanguineous families, and CNVs complemented apparently monoallelic recessive alleles. Mutations of non-coding exon 1 of EYS revealed its contribution to disease. In view of the high carrier frequency for retinal disease gene mutations in the general population, we considered the overall variant load in each patient to assess if a mutation was causative or reflected accidental carriership in patients with mutations in several genes or with single recessive alleles. For example, truncating mutations in RP1, a gene implicated in both recessive and dominant RP, were causative in biallelic constellations, unrelated to disease when heterozygous on a biallelic mutation background of another gene, or even non-pathogenic if close to the C-terminus. Patients with mutations in several loci were common, but without evidence for di- or oligogenic inheritance. Although the number of targeted genes was low compared to previous studies, the mutation detection rate was highest (70%) which likely results from completeness and depth of coverage, and quantitative data analysis. CNV analysis should routinely be applied in targeted NGS, and mutations in non-coding exons give reason to systematically include 5′-UTRs in disease gene or exome panels. Consideration of all variants is indispensable because even truncating mutations may be misleading.

Citation: Eisenberger T, Neuhaus C, Khan AO, Deckert C, Preising MN, et al. (2013) Increasing the Yield in Targeted Next-Generation Sequencing by Implicating CNV Analysis, Non-Coding Exons and the Overall Variant Load: The Example of Retinal Dystrophies. PLoS ONE 8(11): e78496. doi:10.1371/journal.pone.0078496

Editor: Tiansen Li, National Eye Institute, United States of America

Received July 15, 2013; Accepted September 12, 2013; Published November 12, 2013

Copyright: © 2013 Eisenberger et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors received no specific funding for this study.

Competing Interests: TE, CN, CD, AB, CB and HJB are employees of Bioscientia, which is part of a publicly traded diagnostic company. There are no patents, products in development or marketed products to declare. This does not alter the authors’ adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: hanno.bolz@bioscientia.de
Introduction

Retinal dystrophies result from degeneration of photoreceptor and retinal pigment epithelium cells. With a prevalence of ~1 in 3,000, they represent the major cause of hereditary blindness in developed countries [1]. Apart from the individual burden, retinal dystrophies significantly contribute to healthcare costs [2]. Retinal dystrophies are characterized by extensive genetic heterogeneity, with more than 60 genes currently known to underlie retinitis pigmentosa (RP), the most prevalent subtype that affects more than 1.5 million people worldwide [3,4]. Knowing the causative mutation is desirable for several reasons: It provides the basis for personalized genetic counseling and specification of the recurrence risk, and it may predict the natural clinical course (including the determination of a genetic syndrome). In clinically atypical presentations or ambiguous family history, the genotype may specify or even reverse the previous diagnosis or the assumed mode of inheritance. Regarding the progress of gene-replacement therapy approaches for several retinal dystrophies, the genetic diagnosis will be an essential prerequisite for gene-specific therapies [3,5]. However, apart from the c.2991+1655A>G mutation in CEP290 previously reported to be present in 20% of patients with Leber congenital amaurosis (LCA) and RPGR in male RP patients [6,7], there is no major mutation or disease gene for RP and LCA, and clear-cut genotype-phenotype correlations are largely lacking, which prevents efficient targeted Sanger sequencing. Because chip-based analysis for previously reported mutations detects only a fraction of the causative alleles [8], and gene-by-gene analysis by Sanger sequencing is too laborious and expensive, genetic testing has been the exception until recently. Now, next-generation sequencing (NGS) allows for simultaneous and efficient analysis of all known disease genes for a given trait.

NGS of 55 genes involved in RP and LCA (the term “LCA”) was applied for early-onset retinal dystrophies, including infant RP and infant cone-rod dystrophies, CRD; Additional Data File S1) in 126 patients. Causative mutations, including CNVs affecting one and multiple exons, were identified in the majority of patients and confirmed the extensive genetic heterogeneity. Our findings demonstrate the immense potential of NGS for diagnostics of retinal dystrophies and shed light on the genetic complexity of this disease group.

Results and Discussion

Performance of Two NGS Platforms in RD Gene Panel Analysis

Initially 79 samples were sequenced on the Roche 454 GS FLX platform, followed by 38 samples sequenced on the Illumina MiSeq system. With the Roche platform, 90% of the target exons were covered more than 15-fold, with an average coverage of 75-fold per sample. With the Illumina MiSeq instrument, the average coverage was significantly higher (250-fold) and more complete (15-fold for more than 99% of target sequences). 37% of the samples sequenced on the 454 platform were mutation-negative (29 of 79 samples), compared to only 10% sequenced on the MiSeq (7 of 38 samples). CNV analysis was only possible with high-coverage NGS as obtained with the MiSeq system.

High Mutation Detection Rate, Extensive Genetic Heterogeneity and Predominance of Novel Mutations

The overall mutation detection rate was 70% (88/126 patients). More specifically, causative mutations were detected in 38/53 patients (72%) with autosomal recessive (ar) and in 12/14 (86%) with autosomal dominant (ad) RP (Figure 1A,B). Three patients turned out to have X-linked RP based on the genetic findings. In LCA, causative mutations were identified in 35/56 patients (63%; Figure 1C). Although mutations in some genes (RP1 and EYS in arRP, and RPGRIP1, GCY2D and TULP1 in LCA) were more prevalent, mutations in many rare genes account for the majority of patients, confirming that these phenotypes are genetically highly heterogeneous and only comprehensively accessible by highly parallel sequencing of all known disease genes. CEP290, previously reported as the predominant LCA gene, was not a major contributor to this phenotype in our cohort, and its hot spot mutation, c.2991+1655A>G, was not found at all. This may partially be due to the ethnic background of LCA patients in our cohort with 43% of patients originating from the Arabian peninsula. In contrast to other large studies [9], USH2A mutations contributed only to a small proportion of arRP. Causative mutations were found in 28 different genes that encode proteins from diverse pathways and cellular compartments. Mutations in ciliary genes were most prevalent (Figure 1D), indicating the importance of the photoreceptor’s connecting cilium, its associated structures and functions (such as intraflagellar transport) for visual integrity. Of 98 different mutations, 67 were novel (68%) and would thus have been missed by approaches exclusively targeting known alleles such as genotyping microarrays. Below, we describe several families with peculiar findings that further expand our understanding of RD genetics beyond the mere identification of the causative mutations.

CNV Detection from High-coverage NGS Data

Virtualy any gene may be captured and subjected to NGS aimed not only at qualitative, but also quantitative readout. This utilization of NGS data enables CNV detection and can favourably complement MLPA (multiplex ligation-dependent probe amplification), where the application depends on the availability of commercial kits that currently cover only a fraction of known RD genes. We identified four alleles with pathogenic CNVs comprising one to multiple exons. Below, we describe exemplary constellations with CNVs contributing to retinal disease.

CNV and point mutation in a non-coding EYS exon contributing to arRP

Mutations in EYS account for 5–18% of arRP cases depending on the population [10,11]. It has been suggested that at least 15% of patients with monoallelic point mutations may carry midsize rearrangements as second mutant alleles [12]. In our study, EYS mutations were found in 9.4% of arRP patients (five families). One patient was compound heterozygous for a truncating mutation in the coding region and a deletion of non-coding exon 1 at least (Figure 2A,C; Figure S2A in File S1). 5’ non-coding gene sequences, especially first exons, usually contain the promoter and are thus important for gene regulation and vulnerable to mutations [13]. In a recent example, a recurrent de novo mutation creating an aberrant initiation codon of the IFITM5 gene was found to cause a genetic subtype of osteogenesis imperfecta [14,15]. Promoter site prediction programs TSSG [16] and NNPP [17] predict the EYS transcription start site at the beginning of exon 1 and the TATA box upstream. The potential disease-causing effect of exon 1 mutations in EYS is supported by two siblings of a second family with a putative splice site mutation of exon 1 in trans to a truncating mutation in a coding exon (Figure 2B,C). We therefore propose that loss and aberrant splicing of EYS exon 1 should impair transcription of the mutant gene copy and result in a null allele. Our findings illustrate the potential benefit of including 5’-UTRs in NGS of disease gene or
even exome panels. Evaluation of non-coding regulatory regions may identify the “missing hit” in heterozygous carriers of recessive mutations.

Hemizygosity of a CRX mutation in a consanguineous LCA family. In a consanguineous Turkish LCA family with two affected siblings (Figure 3A), homozygosity mapping by genome-wide linkage analysis had initially failed to identify an unambiguous chromosomal candidate region, and the combined maximum parametric LOD score of 2.4 was not obtained (Figure 3B). NGS of a sample from the index patient identified an apparently homozygous CRX mutation in exon 4 that abrogates the natural translation termination codon (c.899A>G), predicting an elongated protein with 118 unrelated residues (p.*300Trpext*118). Subsequent quantitative analysis revealed a heterozygous deletion of exon 4 in trans to the no-stop mutation which was thereby recognized as hemizygous (Figure 3C; Figure S2B in File S1). Both mutations cosegregated with LCA in the family. Interestingly, CRX mutations have mostly been observed in autosomal dominant LCA and CRD [18,19]. Congenital retinal degeneration in a patient with homozygosity for a missense allele, p.Arg90Trp, suggested that CRX may also be a recessive LCA gene [20]. The lack of retinal degeneration in both parents of the index patient and LCA in her brother who also carried both mutations strongly indicate that both CRX mutations identified here represent recessive loss-of-function alleles, confirming the previous assumption that recessive LCA may result from biallelic CRX mutations.

CNVs are common in PRPF31, an adRP gene to be considered in “simplex” RP. Mutations of PRPF31 account for about 5–10% of adRP cases (RP11) [4,21]. RP11 families often display incomplete penetrance, and dominant inheritance may not be obvious from the family history. In five patients, we identified heterogeneous PRPF31 mutations, including deletions of multiple (patient 116) or even all coding exons (patient 113) (see Figure S1 and Figure S2C in File S1). By Sanger sequencing and subsequent MLPA in seven patients with pedigrees suggesting incomplete penetrance, we identified point mutations in two patients, and three had multiple exon to whole-gene deletions (these patients were not part of this study), compatible with a previous study suggesting that the RP11 locus is prone to genomic rearrangements [22]. Patients 22, 23 and 116 had a provisional diagnosis of sporadic and thus recessive RP which was revised after the genetic findings – resulting in significantly higher recurrence risks of up to 50% for the patients’ offspring to be communicated in genetic counseling. Evaluation of PRPF31, including CNV analysis, is therefore advisable in all RP patients independent of the assumed inheritance mode.

Oligogenic Heterozygosity: Accidental Carriership, Potential Modifiers and Non-pathogenic Truncating Mutations

Given the multitude of genes implicated in RP and LCA, it is not surprising that NGS, providing a “full picture” of the mutational load, identifies constellations with mutations in several genes. In view of a recent study of genome sequences from 46 control individuals from various regions of the world indicating that one in 4–5 individuals from the general population may be a
carrier of null mutations in a gene for inherited retinal degeneration [23], constellations with mutations in multiple loci need to be anticipated in a comprehensive NGS approach. In our study, many patients with causative biallelic mutations carried singular heterozygous missense variants in other RD genes (Table 1). These additional alleles were frequently indicated as likely protein-damaging by the prediction programs applied herein, and their contribution to disease severity as modifiers or in an oligo/digenic setting cannot be excluded. Digenic inheritance has been reported for non-syndromic RP due to double heterozygosity for recessive mutations in RDS and ROM1, both encoding interacting structural components of rod outer segments [24], and for deafblindness with mutations in genes encoding interacting proteins (GPR98 and PDZD7) of the Usher protein interactome [25]. However, a final proof of causative oligogenic constellations is often impossible because it usually requires segregation analysis and precise phenotyping in extended families, determination of the variants' prevalence in large cohorts or simulation in animal models as previously reported for AHI1 and PDZD7 [25,26]. Although oligogenic inheritance cannot be excluded in some families, there was no clear evidence for digenic disease or a modifying effect in any patient from our cohort.

However, we identified patients with causative biallelic mutations in recessive RP genes who additionally carried heterozygous truncating mutations in secondary loci. RP1, the most prevalent arRP gene in our cohort, was frequently found together with mutations in other RD genes. The observed constellations resulted in different deductions regarding the pathogenicity of the respective RP1 allele:

Pathogenic RP1 truncations with causality in the family. RP1 mutations are mostly truncating and may cause adRP [27] or arRP [28]. Of note, no RP1 mutations were observed in our adRP patients, but RP1 was the most prevalent arRP gene, with clearly causative biallelic mutations in several cases (11.3%; Table 1). Patient 25 was compound-heterozygous for two truncating RP1 alleles, c.597C>A (p.Tyr199*) and

c.4045C>T (p.Arg1349*)
Figure 3. Hemizygosity of a CRX mutation in a recessive consanguineous LCA family. A. Compound-heterozygosity for a potentially protein-extending no-stop mutation (c.899A>G/p.(*300Trpext*118); here designated as Ext) abrogating the natural termination codon in exon 4 and...
Monoallelic Mutations in Genes Underlying Recessive Retinal Dystrophies

Monoallelic mutations in recessive disease genes represent a challenge for interpretation regarding their causality in the patient, especially if there are no biallelic mutations in another gene for the trait that would qualify such mutations as incidental findings (i.e. carrier status unrelated to the disease in the individual). While single non-synonymous variants in recessive disease genes may often represent rare non-pathogenic variants (see Table 1, “Additional Alleles”, and Table S2 in File S1), the nature of the alteration strongly suggests loss of function for two monoallelic mutations in recessive RP genes in our cohort: The same large in-frame deletion-insertion mutation in TULP1 (p.Aspl24132delinsAla) was identified in two independent simplex RP patients (patients 33 and 82), and a SAG nonsense mutation was found in patient 29. Patient 33 in addition carried two CRX missense variants, p.Arg41Gln and p.Tyr142Cys, that we consider likely benign (both are listed as disease-causing in HGMD, but also in dbSNP, and p.Tyr142Cys was found in several patients with disease-causing mutations in other genes). The sporadic occurrence of RP suggests autosomal recessive inheritance in all three patients, making a dominant-negative effect of the TULP1 and the SAG mutation unlikely. The three DNA samples with monoallelic mutations in TULP1 and SAG mutations were initially sequenced on the GS FLX system; subsequent analysis on the MiSeq platform did not identify additional mutant alleles, in particular no CNVs.

Patients 33, 82 and 29 are therefore either accidental carriers of the TULP1 and SAG mutations with the causative mutation in another arRP gene not known at the time of study design, or the “missing alleles” escaped detection by exonic sequencing because they are deep intronic (as exemplified by the LCA mutation c.2991+1655A>G,chr2p29g or the only known RP23 mutation in the OFD1 gene [34]), or because they localize in regulatory non-coding regions (as shown for EYS exon 1 in this study).

Patients without Mutations – possible Explanations

As discussed above, mutations in known retinal dystrophy genes may escape detection because of their localization – about 13% of disease-causing mutations localize outside coding exonic sequences [35]. Non-coding exons were not systematically included in our study; the identification of mutations in non-coding exon 1 of EYS suggests that such exons should be included in upcoming disease gene panels. Mutation-negative cases in our study will in part be due to mutations in RP and LCA genes that were identified after the design of our gene panel (e.g. NMNAT1, DHDDS, ZNF513, FAM161A, KCNJ13, IMP62, IQCB1, CLRN1, MAK, C8orf37, PRPF6, OFD1). For example, subsequent exome sequencing for patient 15 identified a homozygous nonsense mutation in IMP62
| Patient | Gene | Allele 1 | Allele 2 | Additional Allele | Gender | Age (years) | Phenotype | Inheritance (family history) | Consang. Origin |
|---------|------|----------|----------|------------------|--------|-------------|-----------|----------------------------|----------------|
| Autosomal recessive retinitis pigmentosa, arRP | | | | | | | | | |
| 2      | ABC4 | c.768G>T | p.V256V | – | [1,2,3] c.5603A>T p.Asn1868Ile rs180466 | m | 60 | RP | a/n | no Cau |
| 30     | ABC4 | c.1622T>C | p.Leu541Pro/ | rs1801466 | [4,5] c.3113C>T p.Asn1868Ile rs180466 | f | 17 | RP | a/n | no Ger |
| 31     | ABC4 | c.768G>T | p.V256V | – | [1,2,3] c.6089G>A p.Arg2030Gln rs61750641 | f | 8 | RP | a/n | no Ger |
| 53     | ABC4 | c.1A>G | p.Met1Val | – | [10] c.3113C>T p.Asn1868Ile rs1801466 | f | 44 | RP | a/n | yes Ger |
| 45     | CNAG1 | c.1035C>T | p.Arg346Trp | – | [12] c.1166C>T p.Ser389Phe rs62625014 | m | 43 | RP | a/n | no Ger |
| 74     | CNAG1 | c.2195A>G | p.Glu732Gly | – | [a] c.1233A>G p.Glu411Arg rs142938748 | m | 50 | RP | a/n | no Ger |
| 5      | CRB1 | c.1494T>C | p.Ser498Pro | – | [a] c.1547T>C p.Cys470Arg | m | n.d. | RP | a/n | yes Pak |
| 7      | CRB1 | c.2042G>A | p.Gly741Arg | – | [a] | f | 28 | RP | a/n | no Pol |
| 48     | CRB1 | c.2367T>A | p.Asn789Lys | – | [a] | f | 37 | RP | a/n | no Ger |
| 8      | EYS | c.604T>C | p.Cys202Arg | – | [a] | m | 52 | RP | a/n | no Au |
| 34     | EYS | c.7055+1G>T | spliced | – | [a] | f | 40 | RP | a/n | yes Syr |
| 57     | EYS | c.4040C>T | p.Cys1342Arg | – | [a] | f | 32 | RP | a/n | no Ger |
| 88     | EYS | c.675delA | p.Thr225His | – | [a] | m | 48 | RP | a/n | no Ger |
| 9      | MERTK | c.345C>G | p.Cys115Phe | – | [a] | m | 23 | RP | a/n | no Ger |
| 26     | MERTK | c.1786G>A | p.Gly596Arg | – | [a] | f | 36 | RP | a/n | n.d. n.d. |
| 76     | MERTK | c.1450G>A | p.Gly484Ser | – | [a] | m | 30 | RP | a/n | no Italy |

Table 1. Causative mutations and putatively pathogenic variants identified in this study.
| Patient Gene | Allele 1  | Protein | Cds | db SNP | Ref | Allele 2  | Protein | db SNP | Ref | Additional Allele | Cds | Protein | db SNP | Ref | Gender | Age (years) | Phenotype | Inheritance (family history) | Consang. Origin |
|--------------|----------|---------|-----|--------|-----|----------|---------|--------|-----|-------------------|-----|---------|--------|-----|--------|-------------|-----------|----------------------|-----------------|
| PDE6B        | c.669T>\(A\) | p.Tyr223\* | –   | [a]    | c.669T>\(A\) | p.Tyr223\* | –   | [a]    | RBP3 | f                 | 30  | RP      | a/s    | yes  | Iran   |
| PDE6B        | c.1699C>T  | p.Gln667* | –   | [a]    | c.1699C>T  | p.Gln667* | –   | [a]    | ABCA4 | f                 | 36  | RP      | a/s    | yes  | Ger    |
| PDE6B        | c.2193+1G>A splicing | –   | [20] |        | c.2193+1G>A splicing | –   | [20] |        | ABCA4 | m                 | 47  | RP      | a/s    | yes  | Ger    |
| PDE6B        | c.2193+1G>A splicing | –   | [20] |        | c.2193+1G>A splicing | –   | [20] |        | ABCA4 | f                 | 36  | RP      | a/s    | yes  | Ger    |
| PROM1        | c.642T>G   | p.Tyr214* | –   | [a]    | c.1209_1229del | p.Gln403_Ser | –   | [a]    | RP1   | f                 | 26  | RP      | a/s    | no   | Ger    |
| RDH12        | c.226G>C   | p.Gly76Arg | –   | [23]   | c.866T>G   | p.Val290Glu | rs61740289 | [a]   | ABCA4 | f                 | 34  | RP      | a/s    | yes  | DRC    |
| RDH12        | c.226G>C   | p.Gly76Arg | –   | [23]   | c.866T>G   | p.Val290Glu | rs61740289 | [a]   | ABCA4 | f                 | 34  | RP      | a/s    | yes  | DRC    |
| RP1          | c.597C>A   | p.Tyr199* | –   | [a]    | c.3157delT | p.Tyr1053Thr | –   | [26]  | CDH23 | f                 | 38  | RP      | a/s    | no   | Ger    |
| RP1          | c.597C>A   | p.Tyr199* | –   | [a]    | c.3157delT | p.Tyr1053Thr | –   | [26]  | CDH23 | f                 | 38  | RP      | a/s    | no   | Ger    |
| RP1          | c.4242_4243del | p.His1414Gln | –   | [a]    | c.4474G>T  | p.Glu1492* | –   | [a]    | f                 | 47  | RP      | a/s    | yes  | Iran   |
| RP1          | c.1012C>T  | p.Arg338* | –   | [a]    | c.1012C>T  | p.Arg338* | –   | [a]    | ABCA4 | f                 | 37  | RP      | a/s    | yes  | Iran   |
| RP1          | c.5278_5287del | p.Gln1760Cys | –   | [a]    | c.5278_5287del | p.Gln1760Cys | –   | [a]    | f                 | 12  | RP      | a/s    | yes  | Tur    |
| RP1          | c.5278_5287del | p.Gln1760Cys | –   | [a]    | c.5278_5287del | p.Gln1760Cys | –   | [a]    | f                 | 12  | RP      | a/s    | yes  | Tur    |
| RP1          | c.607G>A   | p.Gly203Arg | –   | [a]    | c.607G>A   | p.Gly203Arg | –   | [a]    | m                 | n.d. | RP      | a/s    | yes  | Iran   |
| RP1          | c.3843delT | p.Pro1282Leu | –   | [a]    | c.3843delT | p.Pro1282Leu | –   | [a]    | f                 | 37  | RP      | a/s    | yes  | Tur    |
| SAG          | c.573C>T   | p.Asn191* | –   | [28]   | ?             | ?             | ?             | ?             | m                 | 16  | RP      | a/s    | no   | Ger    |
| TULP1        | c.371_394del | p.Asp124_ | –   | [a]    | ?             | ?             | ?             | ?             | m                 | 52  | RP      | a/s    | no   | Ger    |
| TULP1        | c.371_394del | p.Asp124_ | –   | [a]    | ?             | ?             | ?             | ?             | m                 | 52  | RP      | a/s    | no   | Ger    |
| TULP1        | c.371_394del | p.Asp124_ | –   | [a]    | ?             | ?             | ?             | ?             | m                 | 52  | RP      | a/s    | no   | Ger    |
## Table 1. Cont.

| Patient Gene | Allele 1 | Allele 2 | Additional Allele | Gender | Age (years) | Phenotype | Inheritance (family history) | Consang. Origin | Cds | Protein | dbSNP | Ref | Cds | Protein | dbSNP | Ref |
|--------------|---------|---------|------------------|--------|-------------|-----------|----------------------------|----------------|-----|---------|-------|-----|-----|---------|-------|-----|-----|
| Autosomal recessive retinitis pigmentosa, arRP |         |         |                  |        |             |           |                            |                |     |         |       |     |     |         |       |     |     |
| 17 | TULP1 | c.1047T>G | p.Asn349Lys | 132delinsAa | (a) | c.1047T>G | p.Asn349Lys | PDE6A | f | n.d. | RP | a/s | yes | Iran |
| 38 | TULP1 | c.1192C>T | p.Arg331Trp | c.31T>G | (a) | c.31T>G | p.Arg331Trp | ABCA4 | f | n.d. | RP | a/s | yes | Iran |
| 13 | USH2A | c.10421A>G | p.Tyr3474Cys | 37 | (a) | c.10421A>G | p.Tyr3474Cys | ABCA4 | f | 50 | RP | a/s | no | Ger |
| 18 | USH2A | c.6925T>C | p.Cys2309Arg | 18 | (a) | c.6925T>C | p.Cys2309Arg | ABCA4 | f | 50 | RP | a/s | no | Ger |
| 68 | USH2A | c.1256G>A | p.Cys419Phe | 113 | (a) | c.1256G>A | p.Cys419Phe | ABCA4 | f | 44 | RP | a/s | no | Ger |
| Autosomal dominant RP, adRP |         |         |                  |        |             |           |                            |                |     |         |       |     |     |         |       |     |     |
| 22 | PRPF31 | c.1048C>T | p.Gln350* | 22 | (a) | c.1048C>T | p.Gln350* | AIPL1 | f | 38 | RP | a/s | no | E-Eur |
| 23 | PRPF31 | c.10561T>C | p.Trp3521Arg | 32,33 | (a) | c.10561T>C | p.Trp3521Arg | AIPL1 | f | 36 | RP | a/s | no | Ger |
| 43 | PRPF31 | c.217A>G | p.Lys73* | 43 | (a) | c.217A>G | p.Lys73* | ABCA4 | f | 14 | RP | a/s | No | Cau |
| 113 | PRPF31 | Deletion of | haplo- | 37,38 | (a) | Deletion of | haplo- | ABCA4 | f | 14 | RP | a/s | No | Cau |
| 116 | PRPF31 | Deletion of | haplo- | 37,38 | (a) | Deletion of | haplo- | ABCA4 | f | 14 | RP | a/s | No | Cau |

Increasing the Yield of Targeted NGS
### Table 1. Cont.

| Patient Gene | Allele 1 | Allele 2 | Additional Allele | Gender | Age (years) | Phenotype | Inheritance (family history) | Consang. Origin |
|--------------|---------|---------|-------------------|--------|-------------|-----------|-------------------------------|---------------- |
| Cds Protein  | db SNP  | Ref     | Cds Protein       | db SNP | Ref         |           |                               |                 |
| **Autosomal recessive retinitis pigmentosa, arRP** | | | | | | | | |
| RHO          | c.937+1G>T | splice | – | [46,47] | | | | |
| TOPORS       | c.2554_2557del | p.Glu853Gln | – | [a] | | | | |
| TOPORS       | c.2590_2553del | p.Arg810Glu | – | [a] | | | | |
| RP2          | c.226G>T | p.Asp76Tyr | – | [a] | | | | |
| RPGR         | c.1853_1856dup | p.Glu621Lys | – | [a] | | | | |
| AIPL1        | c.834G>A | p.Trp278* | m | 2 | LCA | ar | possible Ger | |
| CEP290       | c.3640dupG | p.Glu1214Gly | – | [a] | | | | |
| CEP290       | c.257G>T | p.Glu860* | – | [a] | | | | |

### Increasing the Yield of Targeted NGS

PLOS ONE | www.plosone.org 10 November 2013 | Volume 8 | Issue 11 | e78496
Table 1. Cont.

| Patient Gene | Allele 1 | Allele 2 | Additional Allele | Gender | Age (years) | Phenotype | Inheritance (family history) | Consang. | Origin |
|--------------|---------|---------|------------------|--------|------------|----------|-----------------------------|---------|--------|
| CRB1         | c.2842+2T>A | splice | - | (a) c.2842+2T>A | splice | - | [a] | m | 20 | LCA | no | Tur |
| CRX         | dilation of exon NMID? | 4 | - | (a) c.894A>G | - | [a] | RPI: | f | 9 | LCA | ar | yes | Tur |
| GUCY2D       | c.2080C>T | p.Gln694* | - | [a] | | | | | |
| RCBD         | c.1180T>C | p.Cys394Arg | - | [a] | m | 6 | LCA | ar | yes | KSA |
| RDH12        | deletion of exon NMD? | 4 | - | [a] c.894A>G | - | [a] | RPI: | f | 18 | LCA | ar | yes | Tur |
| RDH5         | c.602C>T | p.Ser201Phe | - | [a] | | | | | |
| RDH5         | c.602C>T | p.Ser201Phe | - | [a] | | | | | |
| LRAT         | c.763C>T | p.Arg255* | - | rs151017794 | c.763C>T | p.Arg255* | - | [a] | | m | 5 | LCA | ar | no | Ger |
| LRAT         | c.763C>T | p.Arg255* | - | rs151017794 | c.763C>T | p.Arg255* | - | [a] | | m | 5 | LCA | ar | no | Mor |
| LRAT         | c.763C>T | p.Arg255* | - | rs151017794 | c.763C>T | p.Arg255* | - | [a] | | m | 5 | LCA | ar | no | KSA |
| LRAT         | c.763C>T | p.Arg255* | - | rs151017794 | c.763C>T | p.Arg255* | - | [a] | | m | 5 | LCA | ar | no | KSA |
| PDE6B        | c.704G>A | p.Arg235His | - | [a] | | | | | |
| PDE6B        | c.704G>A | p.Arg235His | - | [a] | | | | | |
| PDE6B        | c.704G>A | p.Arg235His | - | [a] | | | | | |
| PDE6B        | c.704G>A | p.Arg235His | - | [a] | | | | | |
| PDE6B        | c.704G>A | p.Arg235His | - | [a] | | | | | |

*Table 1: Autosomal recessive retinitis pigmentosa, arRP*
| Patient Gene | Allele 1 | Allele 2 | Additional Allele | Gender | Age (years) | Phenotype | Inheritance (family history) | Cons.ang. Origin |
|--------------|---------|---------|------------------|--------|------------|-----------|-----------------------------|------------------|
| RDH12 | c.133A>G | p.Thr45Ala | – | (56) | c.133A>G | p.Thr45Ala | – | (56) | GUCl2D | m | 12 | LCA | ar | no | Ger |
| | | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
Mutations in the X-linked RP genes *RP2* and *RPGR* have been reported to account for 8.5% of cases with RP of apparently autosomal dominant transmission and for 15% of males with simplex retinal degenerative disease [7,36]. While enrichment and NGS of *RP2* is uncomplicated, the mutational hot spot exon of *RPGR*, ORF15, is not accessible by our NGS approach due to its highly repetitive sequence. Because about 2/3 of *RPGR* mutations reside in ORF15 [37], its inaccessibility causes a diagnostic gap. Thus, male patients (but also females) without mutations in the genes investigated herein may carry mutations in ORF15. The four classes of *RP1* truncating mutations [31] are displayed. Class I, NMD-sensitive truncations; class II, NMD-insensitive truncating mutations representing the majority of pathogenic truncation mutations in *RP1* (dominant negative pathomechanism); class III, NMD-insensitive truncation mutations representing loss-of-function arRP mutations; class IV, NMD-insensitive, non-pathogenic truncations located 3’ of p.1816. CP, “critical position”: 65-residue region between p.1751 and p.1816 containing a yet undefined protein residue before which truncation causes disease.

doi:10.1371/journal.pone.0078496.g004

Comparison with Other NGS Studies on RD

This is the largest NGS study on retinal dystrophies to date. Compared to other NGS studies on this disease group, we
obtained a significantly higher diagnostic yield – which is remarkable because the number of analyzed disease genes (55) in this study was much smaller than in similar studies [9,30–43] (Table 2). This may in part be due to different enrichment and sequencing methods, factors that both influence depth and completeness of coverage and accuracy (for example, NGS with the 454 GS FLX platform results in a higher error rate in homopolymer stretches). High and extensive coverage, as obtained in this study, allow for systematic analysis for CNVs and reduce the risk of mutations escaping detection because of their localization in regions with low coverage. Finally, direct comparison of studies is difficult because of differences in cohort size and composition regarding phenotypes, clinical characterization and traits.

In conclusion, the identification of mutations in 28 RD genes in our cohort, with most alterations previously undescribed, clearly demonstrates that this disease group is accessible only by massively parallel multi-gene sequencing. Although our NGS study was rather conservative and confined to only 55 genes, we detected the causative mutations in the majority from a large cohort of RP and LCA patients. Regular updating of such panels and inclusion of genes for related disorders (e.g. cone-rod dystrophies) is needed to maximize the mutation detection rate. CNV detection from high-coverage NGS data was a major benefit from switching to a high-capacity NGS platform. Therefore, we currently favor NGS of an RD gene panel over exome sequencing where RD gene coverage is reduced due to distribution of reads across some 20,000 genes. Both, oligogenic heterozygosity and monoallelic constellations were observed and may require segregation analysis and careful evaluation of clinical data. Importantly, NGS readout should implicate the overall variant load in order to avoid interpretation pitfalls – as exemplified by the identification of RP1 truncations unrelated to disease in certain constellations. “Missing alleles” in seemingly accidental carriers of recessive RD gene mutations were partly large CNVs and mutations affecting non-coding 5’ exons, demonstrating that both UTR inclusion and quantitative analysis should be part of a comprehensive NGS approach. Because such mutations may also be deep intronic variants with impact on splicing, genominc sequencing, where necessary followed by RNA analysis, may complement primary exonic sequencing in the future. Careful consideration of all variants led to revision of the assumed mode of inheritance, e.g. in case of PRPF81 mutations in simplex RP patients.

As indicated by several exceptional findings in our study, scientific gain of knowledge will strongly benefit from the recent advent of NGS in routine diagnostics and the “hyproducts” of such unprecedented large-scale analyses – not only for RD, but for many other genetically heterogeneous conditions.

Materials and Methods

Ethics Statement

All samples in this study were obtained with written informed consent accompanying the patients’ samples. All clinical investigations have been conducted according to the principles expressed in the Declaration of Helsinki. The study was approved by the institutional review board of the Ethics Committee of the University Hospital of Cologne.

Patients and DNA Samples

A total of 126 patients (53 with arRP, 14 with adRP, 3 with X-RP and 56 with LCA) were included in this study. Genomic DNA was isolated from EDTA blood following standard protocols. The diagnoses of all patients were established by medical history, family history and detailed clinical evaluation of vision. Ophthalmological examination included stereoscopic funduscopy, standard ERG, perimetry, measurement of dark adaptation, and determination of best-corrected visual acuity in most patients.

NimbleGen SeqCap EZ Choice Library Design

Genomic coordinates of coding and non-coding exons in all isoforms were identified in the RefSeq database (hg19) using the University of California Santa Cruz (USCS) table browser [44]. All coding exons (31 arRP genes, 413 exons; 23 adRP genes, 248 exons; 16 LCA genes, 215 exons) of 55 known genes (as of end of year 2010; Table S1 in File S1) including 35 pb of flanking 5’ and 3’ intronic sequence were targeted by a custom SeqCap EZ Choice library (NimbleGen, Madison, Wisconsin, USA). In total, 732 regions were targeted comprising 213 kb of target sequence. The final design covered about 99% of the requested target regions. Because of its highly repetitive sequence which precludes efficient enrichment and sequencing, RPGR exon ORF15 was excluded from panel design. Because the USH2A gene was not included in the panel design, all coding exons of the gene were analyzed either by conventional Sanger sequencing or by a complementary USH2A1-including NGS gene panel in arRP patients without mutations in the RP genes covered by our panel.

Sequence Capture and Next-generation Sequencing (NGS)

Samples from 79 patients were subjected to NGS on the Roche GS FLX platform (454 Life Sciences, Branford, CT; average output 400–500 Mb). In the second part of the study, 38 samples were sequenced on the Illumina MiSeq system (Illumina, San Diego, CA; average output 1.5–5 Gb) with only the latter allowing for CNV detection due to high and uniform coverage. Samples from nine patients were analysed on both systems (two samples with no mutations, four samples with monoallelic mutations in 454 sequencing and three samples with confirmed mutations from 454 sequencing). Between eight (GS FLX) and 20 (MiSeq) samples were pooled and sequenced in a multiplexing procedure. Multiple DNAs were enriched using the NimbleGen SeqCap EZ choice sequence capture approach and sequenced by Roche 454 GS FLX pyrosequencing or by Illumina MiSeq sequencing-by-synthesis technology according to the manufacturers protocols. In brief, 0.5–1 μg of genomic DNA per sample was sheared using the Covaris S2 AFA system (Covaris Inc., Woburn, MA, USA) and ligated to barcoded adaptors for multiplexing. Pre-capture amplified samples were pooled and hybridized to the customized in-solution capture library for 72 hours, subsequently eluted and post-capture amplified by ligation-mediated (LM-) PCR. This amplified enriched DNA was used as input for emulsion PCR (emPCR) and subsequent massively parallel sequencing on one full PTP of a Roche 454 GS FLX platform or as input for direct cluster generation and sequencing on the Illumina MiSeq system (2×150 bp paired-end reads). Uncovered regions of LCA genes (n = 16) in negative samples from LCA patients designated as having “LCA” were sequenced by conventional Sanger sequencing for completeness, whereas in RP samples, gaps of uncovered exons of arRP genes (n = 31) samples were only eliminated by Sanger sequencing in search of a second mutation in an incompletely covered arRP gene.

Read Mapping and Variant Analysis

Demultiplexed reads from the GS FLX platform or paired end reads (2×150 bp) from the Illumina MiSeq instrument were mapped against the hg19 human reference genome using SMALT
### Table 2. Comparison of this study with previous NGS studies on retinal dystrophies.

| Study                  | Cohort size (no. of patients) | Genes (n) | Platform/System | Enrichment | CNV detection | Phenotypes | Detection rate |
|------------------------|-------------------------------|-----------|-----------------|-------------|----------------|------------|----------------|
| This study             | 126                           | 55        | GS FLX (Roche)  | Nimblegen (in solution) | yes            | arRP       | 72%            |
|                        |                               |           | MiSeq (Illumina)|             |                | adRP       | 86%            |
|                        |                               |           | Nimblegen (array) |             |                | LCA       | 63%            |
|                        |                               |           | PCR             |             |                | all        | 70%?           |
| Glockle et al. [43]    | 170                           | 74        | SOLiD 4, −5500 x I (Life Tech) | SureSelect | no             | arRP, sporadic RP | 60%           |
|                        |                               |           | SureSelect      |             |                | adRP       | 41%            |
|                        |                               |           | PCR             |             |                | LCA (4 cases) | not given      |
| Neveling et al. [9]    | 100                           | 111       | GS FLX (Roche)  | Nimblegen (array) | no            | RP         | 36%            |
| O’Sullivan et al. [41] | 50                            | 105       | SOLiD 4 (Life Tech) | SureSelect | no            | RP, CRD    | 50-55%         |
| Shanks et al. [42]     | 36*                           | 73        | GS FLX (Roche)  | Nimblegen (array) | no            | RP         | 25%            |
| Chen et al. [40]       | 25                            | 179 (189) | GAII (Illumina) | Nimblegen | no             | RP (19), STGD (2), FA (1), Usher (1), BBS (1), undefined (1) | 56% |
| Bowne et al. [39]      | 21                            | 46        | GS FLX (Roche)  | GAII (Illumina) | PCR | no | RP (dominant) | 64% |
| Audo et al. [38]       | 17**                          | 177 (254) | GAII (Illumina) | SureSelect (array) | no | RP, CSNB, early CD, M. Best, STGD | 57% |

*Positive controls not included.
**Additional samples from the same families not included. Gene numbers in brackets include additionally screened candidate genes that are not yet proven retinal disease genes. BBS, Bardet-Biedl syndrome; CRD, cone-rod dystrophy; CD, cone dystrophy; CSNB, congenital stationary night blindness; FA, fundus albipunctatus; STGD, Morbus Stargardt; USH, Usher syndrome.

doi:10.1371/journal.pone.0078496.t002
Validation and Segregation Analysis

Sequence variants of interest identified by high-throughput sequencing were verified by Sanger sequencing following PCR amplification of the respective coding exons and adjacent intronic sequences by standard protocols. Purified PCR fragments were sequenced using Big Dye Terminator Cycle sequencing and analyzed on an 3500 Genetic Analyzer sequencer (Applied Biosystems, Foster City, CA, USA). Where applicable, DNA from affected and unaffected family members was analyzed for segregation analysis of putatively causative sequence variants. The resulting sequence data were compared to the reference sequence of the RefSeq database [62].

Copy Number Variation Analysis

Very high coverage was reproducibly achievable by sequencing with the Illumina MiSeq system and enabled copy number variation (CNV) analysis for most of the analyzed genes. Potential copy number alterations (CNA) were initially identified by VarScan [63] on mapped reads. Thereby, coverage of every target region of the sample of interest was internally normalized and compared versus normalized control data of other samples of the same run (VarScan copy number mode and standard settings). Potential CNVs were reported, if the CNV was detected against at least 75% of the control patients. CNVs were annotated using reGene from UCSC (ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/reGene.txt.gz). Potential CNVs were visualized and recalculated with the CNV mode of SeqNext using standard settings and the analysis mode “all vs. all.” Thereby, the normalized relative coverage of every target ROI (region of interest) of a patient sample (relative product coverage, RPC P) was calculated against the normalized average relative target coverage of several control samples (RPC C) to obtain the ratio relative coverage (ratio RPC). Deletions were reported if the ratio RPC fell below 75%, CNVs that fulfilled these criteria were validated by multiplex ligation dependent probe amplification (MLPA) for the affected gene. For the EYS gene, the SALSA MLPA probemix P-329-A1 EYS, for the CRX gene the SALSA MLPA probemix P221-B1 LCA and for the PRPF31 gene the SALSA MLPA KIT P235-B1 Retinitis Pigmentosa was used (MRC-Holland, Amsterdam, The Netherlands). Only CNVs that could be confirmed by MLPA were considered real. MLPA results were visualized with the MLPA module of the SeqPilot software (JSI Medical Systems). The ratio RPA (relative peak area) was calculated as the RPA of the patient versus controls.

Exclusion of the CEP290 Hot Spot Mutation in LCA Patients

For exclusion of the common c.2991+1655A>G mutation in the CEP290 gene mutation in all LCA patients prior to NGS analysis, the region of interest in intron 26 was amplified by PCR. Genotyping for the presence of the mutation was performed by pyrosequencing using QIAGEN Pyro Gold chemistry according to the manufacturers instructions and subsequent analysis on a PSQ 96MA system (QIAGEN, Hilden, Germany).

Linkage Analysis

In the family of patient 110 afflicted with LCA, we performed genome-wide homozygosity mapping using the Affymetrix Gen-chip Human Mapping 10K Array, version 2.0 (Affymetrix, Santa Clara, CA, USA). GRR [64] and PedCheck [65] were used to verify relationships and to identify Mendelian errors. Nonparametric linkage analysis was done with MERLIN [66], Parametric linkage and haplotype analysis was performed using the program ALLEGRO [67]. All data handling was performed using the graphical user interface ALOHOMORA [68]. Graphic output of haplotypes was generated with haploPainter [69].
the patient versus controls. Deletions are indicated if the ratio RPA falls below 75%.

Table S1. Genes analyzed in this study. A. arRP, adRP, and LCA genes that were captured and subjected to NGS in this study. B. Functional categorization of genes with causative mutations. C. Table S2. Additional variants classified as "likely pathogenic". Classification as pathogenic by at least three out of five bioinformatic prediction programs and a minor allele frequency below 3% in unresolved patients. Although a contribution of these variants to the phenotype cannot be excluded, they were not considered causative. In many cases, they represented monoallelic variants in recessive genes which would not sufficiently explain the phenotype. References S1. References for Table 1 and Table S2.

(ZIP)

References

1. Wright AF, Chakarova CF, Abd El-Aziz MM, Bhattacharya S (2010) Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait. Nat Rev Genet 11: 274–284.
2. Frick KD, Roebuck MC, Feldstein JJ, McCarthy GA, Grover LJ (2012) Health services utilization and cost of retinitis pigmentosa. Arch Ophthalmol 130: 629–634.
3. den Hollander AI, Wu W, Cook J, Coats C, et al. (2012) Health-related quality of life of patients with autosomal recessive retinitis pigmentosa. J Clin Invest 120: 3042–3053.
4. Hartong DT, Berson EL, Dryja TP (2006) Retinitis pigmentosa. Lancet 368: 1795–1809.
5. Sheldahl VC, Stone EM (2011) Genomics and the eye. N Engl J Med 364: 1932–1942.
6. den Hollander AI, Koenekoop RK, Vingerling JR, Lutjens E, et al. (2006) Mutations in the CEP290 (NPHP5) gene are a frequent cause of Leber congenital amaurosis. Am J Hum Genet 79: 536–561.
7. Brannham K, Othman M, Brumm M, Karoukis AJ, Atmaca-Sonmez P, et al. (2012) Mutations in RPGR and RP2 account for 15% of males with simple retinal degenerative disease. Invest Ophthalmol Vis Sci 53: 8252–8257.
8. Yager S, Leroy BP, De Barre E, de Ravel TJ, Zonneveld MN, et al. (2006) Microarray-based mutation detection and phenotypic characterization of patients with Leber congenital amaurosis. Invest Ophthalmol Vis Sci 47: 1167–1173.
9. Neuringer K, Collin RW, Gilsen C, van Huet RA, Visser L, et al. (2012) Next-generation genetic testing for retinitis pigmentosa. Hum Mutat 33: 963–972.
10. Littin KV, van den Born LJ, Koenekoop RK, Collin RW, Zonneveld MN, et al. (2010) Mutations in the EYS gene account for approximately 5% of autosomal recessive retinitis pigmentsos and cause a fairly homogeneous phenotype. Ophthalmololgy 117: 2062–2063, 2063 e2061–2067.
11. Hosono K, Ishigami C, Takahashi M, Park DH, Hiran Y, et al. (2012) Two novel mutations in the EYS gene are possible major causes of autosomal recessive retinitis pigmentosa in the Japanese population. PLoS One 7: e31036.
12. Peruja JI, Barragán M, Borrego S, Amo I, González-De Díez Pizo M, et al. (2011) Copy-number variations in patients with Leber congenital amaurosis. Invest Ophthalmol Vis Sci 52: 5602–5611.
13. Davuluri RV, Grosse I, Zhang MQ (2001) Computational identification of promoters and first exons in the human genome. Nat Genet 29: 412–417.
14. Senler O, Garbes I, Kuepp K, Susan D, Zimmermann K, et al. (2012) A mutation in the 5'-UTR of ITIM5 creates an in-frame start codon and cause autosomal-dominant retinitis pigmentosa and cause a fairly-homogenous phenotype. Am J Hum Genet 91: 349–357.
15. Cho TJ, Lee KE, Lee SK, Song SJ, Kim KJ, et al. (2012) A single recurrent mutation in the 5'-UTR of ITIM5 causes autosomal-dominant cataract type V with hyperplastic cilia. Hum Mutat 91: 349–348.
16. Selvev V, Salam A (1997) The Gene-Finder computer tools for analysis of human and model organism genome sequences. Proc Int Conf Inett Syst Mol Biol 5: 294–302.
17. Reese MG (2001) Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Comput Chem 26: 51–56.
18. Freund CL, Gregory-Evans CY, Furukawa T, Papatosta N, Looser J, et al. (1997) Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) as essential for the maintenance of the photoreceptor. Cell 91: 543–555.
19. Freund CL, Wang QL, Chen S, Maskat BI, Wiles CD, et al. (1998) De novo mutations in the CRX homebox gene associated with Leber congenital amaurosis. Nat Genet 18: 311–312.
20. Swaroop A, Wang QL, Wu W, Cook J, Coats C, et al. (1999) Leber congenital amaurosis is caused by a homozygous mutation (c.2600W) in the human orthologue of the rod photoreceptor transtition factor CRX: direct evidence for the involvement of CRX in the development of photoreceptor function. Hum Mol Genet 8: 299–305.
40. Chen X, Zhao K, Sheng X, Li Y, Gao X, et al. (2013) Targeted Sequencing of 179 Genes Associated with HRDs and 10 Candidate Genes Identifies Novel and Recurrent Mutations in Chinese Patients With Various Retinal Diseases. Invest Ophthalmol Vis Sci.

41. O’Sullivan J, Mullane G, Bhaskar S, Dickerson J, Hall G, et al. (2012) A paradigm shift in the delivery of services for diagnosis of inherited retinal disease. J Med Genet 49: 322–326.

42. Shanks ME, Downes SM, Copley RR, Lise S, Broxholme J, et al. (2013) Next-generation sequencing (NGS) as a diagnostic tool for retinal degeneration reveals a much higher detection rate in early-onset disease. Eur J Hum Genet 21: 274–280.

43. Glockle N, Kohl S, Mohr J, Scheurenbrand T, Sprecher A, et al. (2013) Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies. Eur J Hum Genet.

44. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, et al. (2013) The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res 41: D64–69.

45. Li H, Handsaker B, Wyseker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

46. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. (2010) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

47. Liu X, Jian X, Boerwinkle E (2011) dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum Mutat 32: 894–899.

48. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31: 3812–3814.

49. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7: 575–576.

50. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. (2010) A method and server for predicting damaging missense mutations. Nat Mutat 32: 894–899.

51. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31: 3812–3814.

52. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7: 575–576.

53. Ferrer-Costa C, Gelpi JL, Zamakola I, Parraga I, de la Cruz X, et al. (2005) PMUT: a web-based tool for the annotation of pathological mutations on proteins. Bioinformatics 21: 3176–3178.

54. Pollard KS, Huibsz MJ, Roseribloom KR, Siepel A (2010) Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20: 110–121.

55. Davydov EV, Goode DL, Sirotz M, Cooper GM, Sidow A, et al. (2010) Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol 6: e1001025.

56. Reese MG, Eeckman FH, Kulp D, Haasler D (1997) Improved splice site detection in Genie. J Comput Biol 4: 311–323.

57. Brunak S, Engelbrecht J, Knausen S (1993) Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol 220: 49–65.

58. Hebsgaard SM, Korning PG, Toftstrup N, Engelbrecht J, Pouze P, et al. (1996) Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. Nucleic Acids Res 24: 3439–3452.

59. Rogaiz IB, Milaneu L (1997) Analysis of donor splice sites in different eukaryotic organisms. J Mol Evol 45: 50–59.

60. Cartegni L, Wang J, Zhu Z, Zhang MQ, Kramer AR (2003) ESEFinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31: 3560–3571.

61. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, et al. (2012) The Human Gene Mutation Database (HGMD) and its exploitation in the field of personalized genomics and molecular evolution. Curr Protoc Bioinformatics Chapter 1: Unit 13.

62. Pruitt KD, Tatusova T, Brown GR, Maglott DR (2012) NCBI Reference Sequence (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res 40: D130–135.

63. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, et al. (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22: 568–576.

64. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) GRR: graphical representation of relationship errors. Bioinformatics 17: 742–743.

65. O’Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63: 259–266.

66. Davydov EV, Goode DL, Sirotz M, Cooper GM, Sidow A, et al. (2010) Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol 6: e1001025.

67. Gudbjartsson DF, Jonasson K, Frigge ML, Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. Nat Genet 30: 97–101.

68. Rogaiz IB, Milaneu L (1997) Analysis of donor splice sites in different eukaryotic organisms. J Mol Evol 45: 50–59.

69. Thiele H, Nurnberg P (2005) HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics 21: 1730–1732.

70. Pollard KS, Huibsz MJ, Roseribloom KR, Siepel A (2010) Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 20: 110–121.

71. Hebsgaard SM, Korning PG, Toftstrup N, Engelbrecht J, Pouze P, et al. (1996) Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. Nucleic Acids Res 24: 3439–3452.

72. Cartegni L, Wang J, Zhu Z, Zhang MQ, Kramer AR (2003) ESEFinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31: 3560–3571.

73. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, et al. (2012) The Human Gene Mutation Database (HGMD) and its exploitation in the field of personalized genomics and molecular evolution. Curr Protoc Bioinformatics Chapter 1: Unit 13.

74. Pruitt KD, Tatusova T, Brown GR, Maglott DR (2012) NCBI Reference Sequence (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res 40: D130–135.

75. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, et al. (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22: 568–576.

76. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) GRR: graphical representation of relationship errors. Bioinformatics 17: 742–743.

77. O’Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63: 259–266.

78. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin: rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30: 97–101.

79. Guilbjartsson DF, Jonasson K, Frigge ML, Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. Nat Genet 30: 12–15.

80. Ruscendorf F, Nurnberg P (2005) HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics 21: 1730–1732.