Genetic analysis of the rhodopsin gene identifies a mosaic dominant retinitis pigmentosa mutation in a healthy individual

Avigail Beryozkin[1], Gal Levy[1], Anat Blumenfeld[1], Segev Meyer[1], Prasanthi Namburi[1], Yair Morad[2], Libe Gradstein[3], Anand Swaroop[4], Eyal Banin[1*], Dror Sharon[1*]

[1] Departments of Ophthalmology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel.

[2] Department of Ophthalmology, Assaf Harofeh Medical Center, Zerifin 73000, Israel.

[3] Department of Ophthalmology, Clalit Health Services and Soroka University Medical Center, Beer-Sheva, Israel.

[4] Neurobiology-Neurodegeneration & Repair Laboratory, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, USA.

*- Shared contribution.

Addresses for correspondence:

Dror Sharon, Ph.D.; E-mail: dror.sharon1@gmail.com; Tel: (972) 2-6777112

Eyal Banin, M.D., Ph.D.; E-mail: banine@cc.huji.ac.il; Tel: (972) 2-6776585

At the Department of Ophthalmology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel

SUPPLEMENTARY MATERIALS AND METHODS

Clinical evaluation

FFERGs were recorded using monopolar corneal electrodes (Henkes-type, Medical Workshop B.V., Groningen, The Netherlands) and a computerized system (UTAS 3000; LKC, Gaithersburg, MD). Cone responses to 30-Hz flashes of white light were acquired under a background light of 21 cd/m^2. Scotopic responses including a rod response to a dim blue flash and a mixed cone-rod response to an ISCEV standard white flash were acquired following 30-45 minutes of dark adaptation. Between 2-4 sets of responses were recorded in each condition to verify repeatability. All ERG responses were filtered at 0.3 to 500 Hz, and signal averaging was used.

Whole exome sequencing (WES)

WES analysis was performed as follows: genomic DNA (3 μg) was fragmented by Covaris and subjected to whole exome capture using Agilent SureSelectXT Target Enrichment Kit (50 Mbp) for Illumina Multiplex Sequencing (Agilent Technologies, Santa Clara, CA), following manufacturer′s instructions. Captured libraries were amplified and converted to clusters using Illumina Cluster Station. Paired-end sequencing was performed on Illumina GAIIx. Approximately 3 GB of sequence was generated per individual, resulting in ∼90% coverage of targeted Consensus Coding Sequence project (CCDS) exonic bases, with an average depth of ∼80×. Sequence reads were aligned to the human genome reference (UCSC hg 19;http://genome.ucsc.edu/) using the Genomatix Mining Station (GMS) and variants were called and annotated using the DNAnexus software package (https://www.dnanexus.com/). Variant files were annotated using ANNOVAR

according to the dbSNP database (built 137) with the following filtering steps: (1) All variants in 20 known adIRD genes were analyzed (Table S2); (2) Variant type: Missense, nonsense, splice-site, stop-loss, insertions and deletions in the coding region were included; (3) Variants that were found in repeat DNA segments were excluded; (4) Variants with minor allele frequency (MAF) greater than 0.1% in the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/) and in an in-house Israeli WES database were excluded; (5) Prediction of the possible effect of each variant was analyzed by 3 prediction online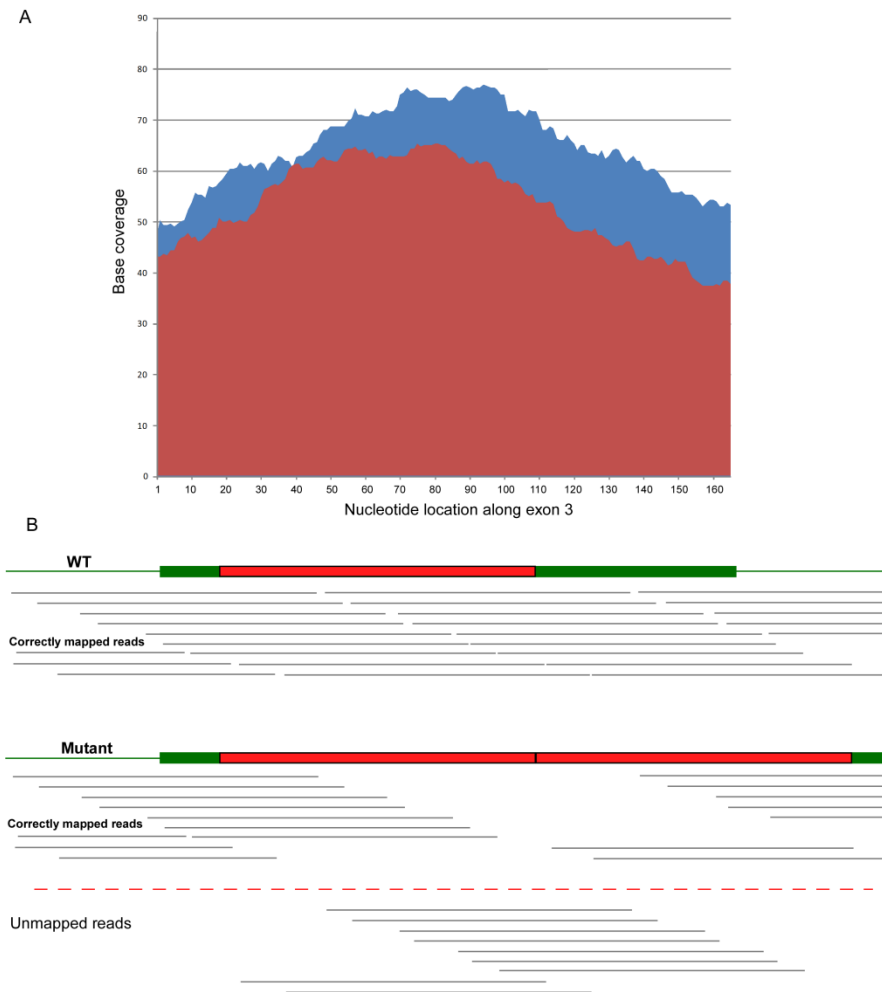 programs [SIFT (http://sift.jcvi.org/), MutationTaster (http://www.mutationtaster.org/) and PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/).

Protein structure analysis

The OCTOPUS online program (http://octopus.cbr.su.se/index.php), which is based on a hidden Markov model for predicting protein topology and localization of transmembrane domains was used to evaluate the mutation effect on protein topology, the Phyre[2] online program was used for predicting the 3-dimentional structure of a protein amino acid sequence (http://www.sbg.bio.ic.ac.uk/phyre2), and PDBsum (http://www.ebi.ac.uk/pdbsum/) was used to present 3D protein structure.

Supplementary table S1: Rhodopsin primers sequences

Exon #	Forward primer	Revers primer	TM
1	CAGCTGGAGCCCTGAGT	GGACAGGAGAAGGGAGAA	58 ºC
2	CTACCCTCTCCCTGTCTAGG	CAAGGTCAGTGCCTGGAACC	58 ºC
3	GTGAAGCCCCAGAAAGG	AGACCATGGCTCCTCCA	60 ºC
4	TCACCGTCAAGGAGGTA	CCAGTAACCCTGATGGAGCC	58 ºC
5	GGCGAACCTCACTAACG	ATGTGACTTCGTTCATTCTG	58 ºC

Supplementary Figure 1: WES analysis of RHO exon 3

A. coverage plot of *RHO* exon 3 in three affected individuals from family

MOL1076 (red) comparing to three controls (blue). The graph shows no

significant difference between the two groups. **B.** A scheme representing the

predicted alignment of sequence reads in a wt *RHO* exon 3 sequence (top) and an

exon including the c.548-638dup91bp duplication. The duplicated area is marked

by red color, intronic areas by thin green line and the remaining of exon 3 by a

thick green line.

Supplementary Figure 2: Bioinformatic analyses of the wt and predicted

mutated rhodopsin proteins.
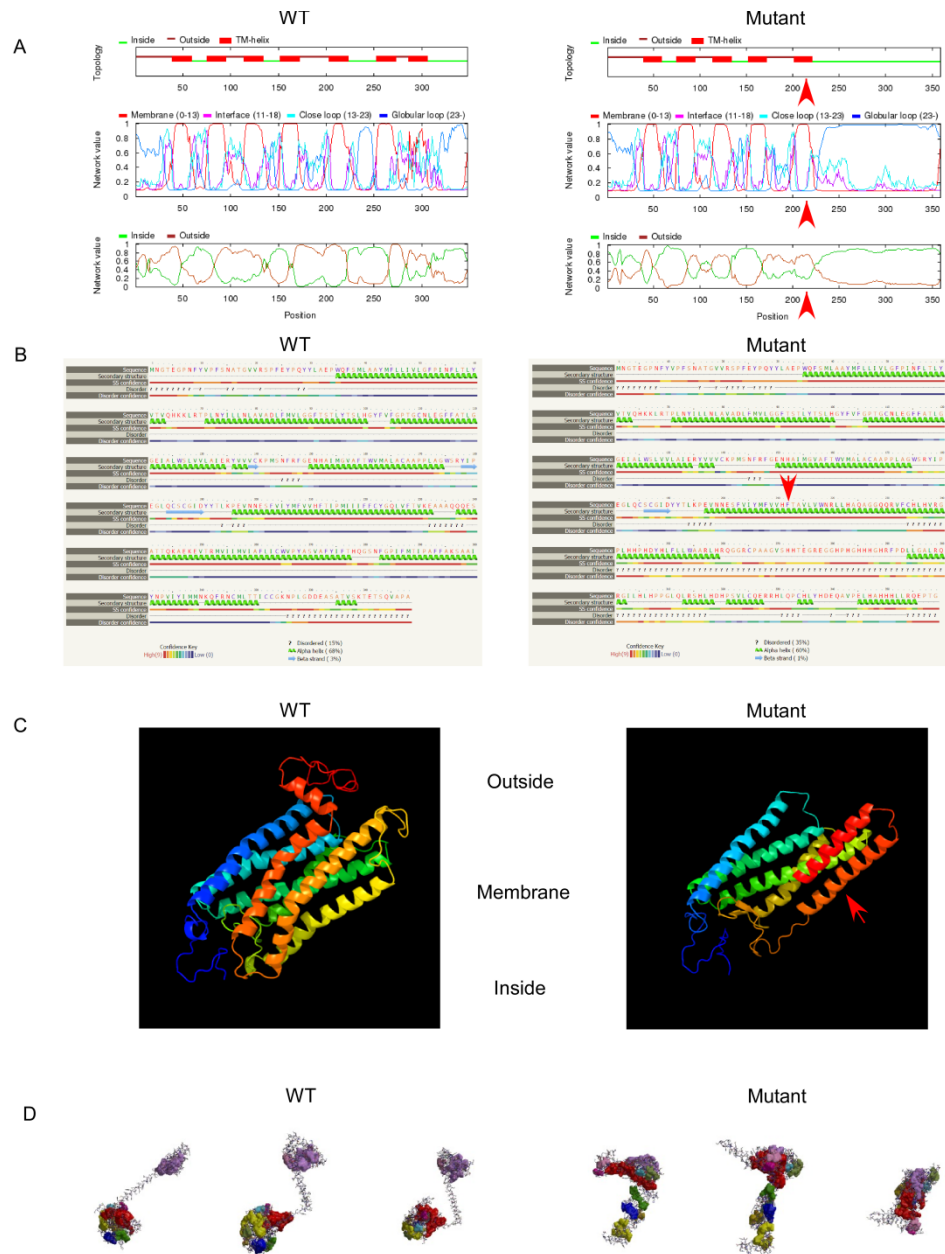
Supplementary Figure 2: Bioinformatic analyses of the wt and predicted mutated rhodopsin proteins.

A. An OCTOPUS analysis of the WT rhodopsin protein (left side in panels A-D) and mutant protein (right side in panels A-D) showing the membrane-associated structures. In all panels, the mutation site (a 91-bp duplication causing a frameshift) is indicated by an arrow. **B.** A similar analysis using the Phyre^{2} program. Please note that the last two transmembrane domains are missing leading to a disordered protein structure. **C.** A Phyre^{2} analysis of the wt (left) and mutated (right) tertiary protein structures at the same orientation showing the large effect of the duplication on the whole protein structure. **D.** A PDBsum cleft analysis showing modeling of the protein surface as viewed from three different angles. The mutant protein (three images on the right side) shows a disrupted structure even in regions that are located upstream to the mutation site.