PNAS
www.pnas.org

Supplementary Information for

Worldwide Carrier Frequency and Genetic Prevalence of Autosomal Recessive Inherited Retinal Diseases

Mor Hanany[a], Carlo Rivolta[b,c,d,1], Dror Sharon[a,1,2]

[a] Department of Ophthalmology, Hadassah Medical Center, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, 91120, Israel.

[b] Department of Genetics and Genome Biology, University of Leicester, Leicester, LE1 7RH, United Kingdom.

[c] Clinical Research Center, Institute of Molecular and Clinical Ophthalmology Basel (IOB), 4031 Basel, Switzerland.

[d] Department of Ophthalmology, University Hospital Basel, 4031 Basel, Switzerland.

[1]- D.S. and C.R. contributed equally to this work.

[2] - Corresponding author: Dror Sharon, PhD., Department of Ophthalmology, Hadassah-Hebrew University Medical Center, Jerusalem, Israel, 91120; Phone: +972 2 6777112; Fax: +972 2 6448917; E-mail: dror.sharon1@mail.huji.ac.il

www.pnas.org/cgi/doi/10.1073/pnas.1913179117

This PDF file includes:

Supplementary text
Figures S1 to S3

Other supplementary materials for this manuscript include the following:

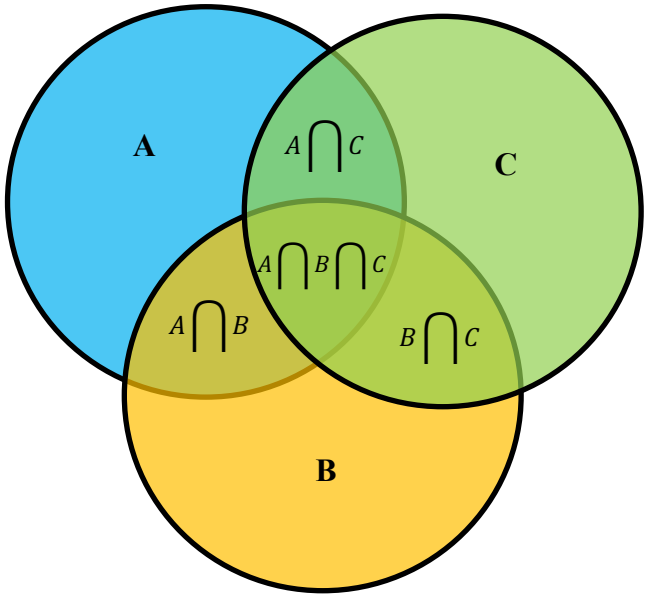
Datasets S1 to S5

Supplementary Information Text

Supplementary Note

Analysis of hypomorphic variants- Twenty-six variants in *ABCA4* (see Supplementary Table 3) (1–5), one possible hypomorphic variants in *BBS1*(6–8), and one in *NMNAT1*(9), were reported as hypomorphic alleles. The 26 *ABCA4* variants were reported to cause disease only when found *in trans* with a severe mutation, usually a truncating one. We therefore performed a separate calculation for the three above-mentioned genes (in which we excluded all combinations of hypomorphic and missense variants as well as homozygous hypomorphic genotypes) in order to assess more accurately disease prevalence and carrier frequency. When calculating genetic prevalence for the remaining genes, we included all possible combinations of mutations.

Fig. S1. Venn diagram



Venn diagram representing the inclusion-exclusion principle. In this example there are 3 groups, each represents the carriers of a specific mutation in a given gene. The overlapping regions represent the individuals who are carriers of more than one mutation (in this case carriers of mutations in two or three genes). In order to calculate the carrier frequency per each sub-population we summed all three groups |A| + |B| + |C| and then subtracted the overlapping areas ($|A \cap B| + |A \cap C| + |B \cap C|$). Since we subtracted the overlapping regions of all groups in the process, we added it again $|A \cap B \cap C|$. In the end, the calculation is –
|A| + |B| + |C| $-(|A \cap B| + |A \cap C| + |B \cap C|)$ + $|A \cap B \cap C|$

Fig. S2. The product calculation. A. Based on carrier frequency. B. Based on allele frequency.

A

RDH12 mutation		c.184C>T	c.193C>T	c.250C>T	c.524C>T	c.698T>A	c.701G>A	.806_810delCCCTG	c.844T>G	c.869T>G	c.910T>C
		1	2	3	4	5	6	7	8	9	10
	CF of Mutations	0.0003331390	0.0001306848	0.0001306848	0.0000832016	0.0002501251	0.0001667361	0.0000843313	0.0000843811	0.0019141145	0.0001306848
c.184C>T	0.0003331390	1.11E-07	4.35E-08	4.35E-08	2.77E-08	8.33E-08	5.55E-08	2.81E-08	2.81E-08	6.38E-07	4.35E-08
c.193C>T	0.0001306848	4.35E-08	1.71E-08	1.71E-08	1.09E-08	3.27E-08	2.18E-08	1.10E-08	1.10E-08	2.50E-07	1.71E-08
c.250C>T	0.0001306848	4.35E-08	1.71E-08	1.71E-08	1.09E-08	3.27E-08	2.18E-08	1.10E-08	1.10E-08	2.50E-07	1.71E-08
c.524C>T	0.0000832016	2.77E-08	1.09E-08	1.09E-08	6.92E-09	2.08E-08	1.39E-08	7.02E-09	7.02E-09	1.59E-07	1.09E-08
c.698T>A	0.0002501251	8.33E-08	3.27E-08	3.27E-08	2.08E-08	6.26E-08	4.17E-08	2.11E-08	2.11E-08	4.79E-07	3.27E-08
c.701G>A	0.0001667361	5.55E-08	2.18E-08	2.18E-08	1.39E-08	4.17E-08	2.78E-08	1.41E-08	1.41E-08	3.19E-07	2.18E-08
c.806_810delCCCTG	0.0000843313	2.81E-08	1.10E-08	1.10E-08	7.02E-09	2.11E-08	1.41E-08	7.11E-09	7.12E-09	1.61E-07	1.10E-08
c.844T>G	0.0000843811	2.81E-08	1.10E-08	1.10E-08	7.02E-09	2.11E-08	1.41E-08	7.12E-09	7.12E-09	1.62E-07	1.10E-08
c.869T>G	0.0019141145	6.38E-07	2.50E-07	2.50E-07	1.59E-07	4.79E-07	3.19E-07	1.61E-07	1.62E-07	3.66E-06	2.50E-07
c.910T>C	0.0001306848	4.35E-08	1.71E-08	1.71E-08	1.09E-08	3.27E-08	2.18E-08	1.10E-08	1.10E-08	2.50E-07	1.71E-08

B

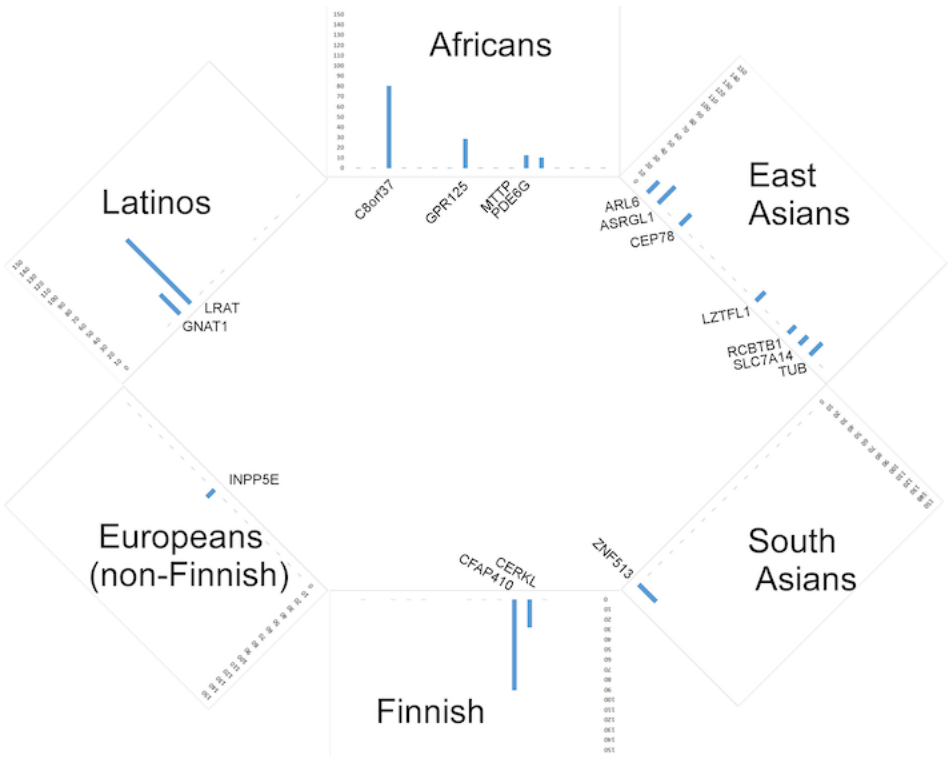
RDH12 mutation		c.184C>T	c.193C>T	c.250C>T	c.524C>T	c.698T>A	c.701G>A	.806_810delCCCTG	c.844T>G	c.869T>G	c.910T>C
	AF of Mutations	3.33E-04	0.000130685	1.31E-04	8.32E-05	0.000250125	1.67E-04	8.43E-05	8.44E-05	1.91E-03	1.31E-04
c.184C>T	3.33E-04	1.11E-07	4.35E-08	4.35E-08	2.77E-08	8.33E-08	5.55E-08	2.81E-08	2.81E-08	6.38E-07	4.35E-08
c.193C>T	0.000130685	4.35E-08	1.71E-08	1.71E-08	1.09E-08	3.27E-08	2.18E-08	1.10E-08	1.10E-08	2.50E-07	1.71E-08
c.250C>T	1.31E-04	4.35E-08	1.71E-08	1.71E-08	1.09E-08	3.27E-08	2.18E-08	1.10E-08	1.10E-08	2.50E-07	1.71E-08
c.524C>T	8.32E-05	2.77E-08	1.09E-08	1.09E-08	6.92E-09	2.08E-08	1.39E-08	7.02E-09	7.02E-09	1.59E-07	1.09E-08
c.698T>A	0.000250125	8.33E-08	3.27E-08	3.27E-08	2.08E-08	6.26E-08	4.17E-08	2.11E-08	2.11E-08	4.79E-07	3.27E-08
c.701G>A	1.67E-04	5.55E-08	2.18E-08	2.18E-08	1.39E-08	4.17E-08	2.78E-08	1.41E-08	1.41E-08	3.19E-07	2.18E-08
c.806_810delCCCTG	8.43E-05	2.81E-08	1.10E-08	1.10E-08	7.02E-09	2.11E-08	1.41E-08	7.11E-09	7.12E-09	1.61E-07	1.10E-08
c.844T>G	8.44E-05	2.81E-08	1.10E-08	1.10E-08	7.02E-09	2.11E-08	1.41E-08	7.12E-09	7.12E-09	1.62E-07	1.10E-08
c.869T>G	1.91E-03	6.38E-07	2.50E-07	2.50E-07	1.59E-07	4.79E-07	3.19E-07	1.61E-07	1.62E-07	3.66E-06	2.50E-07
c.910T>C	1.31E-04	4.35E-08	1.71E-08	1.71E-08	1.09E-08	3.27E-08	2.18E-08	1.10E-08	1.10E-08	2.50E-07	1.71E-08

In this example, one specific gene (*RDH12*) in a specific population (Africans) is depicted. All mutations are shown in the first column and first row and the mutation carrier frequency (panel A, used for calculations by the "inclusion-exclusion" method) or allele frequency (panel B, used for calculations by the "independence probability" method) are in red.

The green cells in panel A represent the chances of two individuals who are heterozygous for different mutations to meet and the blue cells represent the chances of two individuals who are heterozygous for the same mutation to meet. Summing all cells (green and blue) and dividing the result by four (chances of an affected offspring if two parents are carriers) is equal to the genetic prevalence. The CF values for all remaining genes are listed in Table S3.

The green cells in panel B represent the genetic prevalence of compound heterozygous patients for each pair of mutations and the blue cells represent the same values for homozygotes. Again, the sum of all values of the matrix corresponds to the genetic prevalence for *RDH12* mutations in Africans.

Fig. S3. Genes that are commonly mutated in each sub-population and their corresponding prevalence.



Each of the six panels represents a different sub-population, with mutated genes represented on the X-axis and the ratio between the genetic prevalence per million in the studied sub-population and the remaining sub-populations on the Y-axis.

Dataset S1 (separate file).

Table S1: A list of 180 IRD-causing genes analyzed in this study
Table S2: Distribution of mutations, CF and GP for each IRD gene in the different sub-populations
Table S3: A list of all 3,934 likely pathogenic variants
Table S4: Phenotypes determined for each IRD gene
Table S5: Homozygous individuals in gnomAD

Supplementary References:

1. Collison F.T. et al. Clinical characterization of Stargardt disease patients with the p . N1868I ABCA4 mutation. *Retina*. 2018:1-15.
2. Cornelis S.S. et al. In Silico Functional Meta-Analysis of 5,962 ABCA4 Variants in 3,928 Retinal Dystrophy Cases. *Hum Mutat*. 2017;38(4):400-408.
3. Zernant J. et al. Extremely hypomorphic and severe deep intronic variants in the ABCA4 locus result in varying Stargardt disease phenotypes. *Cold Spring Harb Mol Case Stud*. 2018;4(4):1-11.
4. Kaway C.S., Adams K.M. A Novel ABCA4 Mutation Associated with a Late-Onset Stargardt Disease Phenotype : A Hypomorphic Allele ? *Case Rep Ophthalmol*. 2017;4120:180-184.
5. Zernant J. et al. Frequent hypomorphic alleles account for a significant fraction of ABCA4 disease and distinguish it from age-related macular degeneration. *J Med Genet*. 2017;54(6):404-412.
6. Castro-Sánchez S. et al. Exploring genotype-phenotype relationships in Bardet-Biedl syndrome families. *J Med Genet*. 2015;52(8):503 LP - 513.
7. Estrada-Cuzcano A. et al. BBS1 mutations in a wide spectrum of phenotypes ranging from nonsyndromic retinitis pigmentosa to Bardet-Biedl syndrome. *Arch Ophthalmol*. 2012;130(11):1425-1432.
8. Forsythe E. et al. Genetic predictors of cardiovascular morbidity in Bardet – Biedl syndrome. 2015:343-349.
9. Eblimit A. et al. NMNAT1 E257K Variant, Associated with Leber Congenital Amaurosis (LCA9), Causes a Mild Retinal Degeneration Phenotype Aiden. 2019:32-43.