Journal of Genetics, Vol. 97, No. 1, March 2018, pp. 307–309
https://doi.org/10.1007/s12041-017-0871-y

RESEARCH NOTE

Genetic disease in India and the West compared: provisional analysis of population dynamics

NICHOLAS MITCHISON[1]* and TIMOTHY MITCHISON[2]

[1] *University College London, WC1E 6BT, UK*
[2] *Harvard Medical School, Boston, MA 02115, USA*
*For correspondence. E-mail: n.mitchison@ucl.ac.uk.

Received 18 August 2016; revised 8 February 2017; accepted 10 February 2017; published online 11 January 2018

Abstract. The Indian Genetic Disease Database (IGDD) and Online Mendelian Inheritance in Man (OMIM) survey human populations that have different climate histories. Comparison of the two shows an outstanding difference in the relative frequency of recessive disease genes. Several of the diseases mediated at least in part by recessive gene mutations in India are not so mediated in the Western populations covered by OMIM, or are so mediated to a lesser extent. This we attribute to climate history, where population fall leading to inbreeding in the last ice age appears to have reduced the frequency of recessive disease genes in the Western world. This 'ice age benefit' hypothesis is further confirmed, partially by Kolmogorov–Smirnov analysis.

Keywords. The Indian Genetic Disease Database; Online Mendelian Disease in Man; recessive dominant; population; ice age.

Introduction

The Indian Genetic Disease Database (IGDD) allows comparison of the prevalence of recessive (AR) versus dominant (AD) genetic disease in India compared with that in the West. The relative prevalence is of interest in relation to population dynamics: dominant disease mutations get eliminated rapidly, unlike recessives that can persist indefinitely as heterozygotes. The ratio of recessive to dominant mutation thus reflects historical changes in population size. Human populations of the northern hemisphere appear to have lost recessive disease alleles during the population shrinkage of the last ice age (~20,000 years ago) to an extent that is still evident (Mitchison *et al.* 2011; Erickson and Mitchison 2014). No such population bottlenecks affected southern Asia: a coalescence analysis of mitochondrial DNA revealed steady population growth over a period running from 50,000 years ago, in a study that included comprehensive survey of previous work (Atkinson *et al.* 2008).

A crucial test of this hypothesis is therefore whether the southern Asia populations, unaffected by the last ice age, have a larger proportion of recessive disease mutations. IGDD provides an arena in which to test this prediction, as it allows the proportion of recessive disease genes obtained from IGDD data to be compared with that found in Online Mendelian Inheritance in Man (OMIM) for the European/North American population.

Our findings

Inheritance data were collected from IGDD for the 38 diseases that so far have at least 10 patients among the 107 diseases so far listed. They are presented here in table 1, arranged side by side with the Euro-American data from OMIM for the corresponding diseases. The OMIM gene frequencies were checked in NCBI Gene Reviews, and where necessary also in original publications, and are jointly cited here as OMIM+ data. Care was taken to exclude the South Asian data occasionally cited in OMIM, such as recessive inheritance of cogenital cataract in Iran and of spinocerebellor ataxia (SCA) in Pakistan and Saudi Arabia. We exclude as exceptional the massive data for beta-thalassaemia. We did not attempt to find relevant data for consanguinity, which might well influence the comparison made here.

A usual test case in the population genetic studies is the well-documented disease retinitis pigmentosa which

Table 1. Autosomal recessive inheritance found in our collection of 38 diseases listed in IGDD (for Indian) and OMIM (Euro-American) data.

Genetic disease	%AR	
	OMIM+	IGDD
Achondroplasia	0	0
Androgen insensitivity syndrome	0	0
Aniridia	0	0
Becker muscular dystrophy	0	0
Chronic pancreatitis	0	19.1
Congenital cataract	0	13.8
Congenital hereditary endothelial dystrophy	100	57.1
Cystic fibrosis	100	34.4
Duchenne muscular dystrophy	0	0
Familial hypercholesterolaemia	0	7.8
Gilbert syndrome	0	62.8
Glanzmann thrombasthenia	92	87.7
Haemophilia A	0	2.8
Haemophilia B	0	0.8
Haim Munk syndrome	100	100
Huntington disease	0	0
Hypertrophic cardiomyopathy	0	32.9
Isolated growth hormone deficiency type I	0	31.4
Isolated growth hormone deficiency type 2	100	24.1
Lafora disease	100	94.4
Lattice corneal dystrophy	0	100
Macular corneal dystrophy	100	11.6
Megaloblastic anaemia	100	100
Myotonic dystrophy	0	0
Nonsyndromic hearing loss	77	60.2
Oculocutaneous albinism type 1	100	81
Oculocutaneous albinism type 2	100	75
Parkinson disease	85	5
Porphyria	0	0
Primary angle closure glaucoma	0	0
Primary congenital glaucoma	0	49.4
Pure gonadal dystrophy	4	16.7
Retinitis pigmentosa	55	57.9
Spinocerebellar ataxia 1	0	0
Spinocerebellar ataxia 2	0	9
Spinocerebellar ataxia 12	0	9.1
Venous thrombosis	0	9.1
Wilson disease	100	100
Average	32.83	28.6025

in a world-wide survey has 50–60% AR inheritance (Hartong *et al.* 2006). This value was indeed found again in the IGDD.

IGDD conveniently records the relative AR and AD frequencies directly, whereas the OMIM frequencies need to be searched out (apart from the zero frequencies, which are provided directly and unequivocally). Additional South Asian data can be expected in the future from the Iranian Human Mutation Gene Bank, which so far lists only few diseases.

In making the comparison, the averages shown in table 2 are clearly inappropriate for data that are not normally distributed and have so many zeros. Accordingly our first step was to consider the 25 diseases listed in OMIM (Euro-American) data as having zero AR causal mutations. Of these diseases, 13 with a total of 760 patients have AR causal mutations listed in the IGDD data, a finding that supports our hypothesis that the ice age purge hardly affected Indian recessive disease alleles. The finding is all the more telling as OMIM lists zero frequencies unequivocally, without need here for further information.

A particularly telling finding is that the *SPINK1* gene, for chronic pancreatitis, has AD but no AR entries in OMIM, but 70 of 331 entries for the disease in IGDD are AR. And as further expected from our hypothesis, no instances were found of the reverse: zero-AR IGDD entries paired with positive AR entries in OMIM.

The %AR data collected here are not normally distributed, so to better visualize the two datasets we plotted them as cumulative histograms in the

Table 2. Twelve diseases listed in OMIM as exclusively dominant but which in IGDD have a total of 810 AR cases listed here

AR cases listed in IGDD with zero AD	
Chronic pancreatitis	41
Congenital cataract	26
Famelial hypercholesterolaemia	50
Gilbert syndrome	129
Haemophilia A	178
Haemophilia B	128
Hypertrophic cardiomyopathy	24
Isolated growth hormone deficiency type1	12
Lattice corneal dystrophy	37
Primary congenital glaucoma	154
Spinocerebellar taxia	12
Venous thrombosis	19

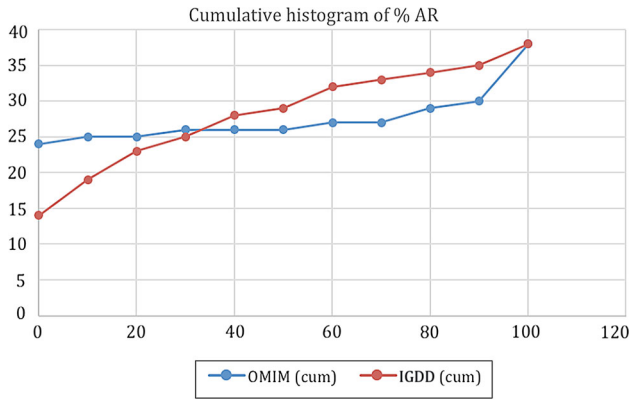

Figure 1. Kolmogorov–Smirnov analysis of OMIM and IGDD %AR data.

Kolmogorov–Smirnov test (czaiontz@gmail.com, info@real-statistics.com). The K–S statistic observed is 0.25, and thus below 0.304 (the 5% level of the K–S test). However, as shown in figure 1, over most of the range (40–90) the Indian data have higher values than the Euro-American data, a finding that weakly supports the ice age purge hypothesis. We thank Dr Gerald Clarke for help with this analysis.

Corresponding editor: Rajiva Raman

Summary

Data collected here show that the proportion of recessive disease is higher in India that in Euro-American, and thus support the hypothesis that the population bottleneck of the last ice age has diminished the prevalence of this form of inherited disease in the West. The range of IGDD data is expected to expand in the future, and thus allow further test of this hypothesis.

References

Atkinson Q. D., Gray R. D. and Drummond A. J. 2008 mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Mol. Biol. Evol*. **25**, 468–474.

Erickson R. P. and Mitchison N. A. 2014 The low frequency of recessive disease: insights from ENU mutagenesis, severity of disease phenotype, GWAS associations, and demography: an analytical review. *J. Appl. Genet*. **55**, 319–327.

Hartong D. T., Berson E. L. and Dryja T. 2006 Retinitis pigmentosa. *Lancet* **368**, 1795–1809.

Mitchison N. A., Bhattacharya S. and Tuddenham E. G. 2011 Human congenital diseases with mixed modes of inheritance have a shortage of recessive disease. A demographic scenario? *Ann. Hum. Genet*. **75**, 688–693.