La Prensa Medica Argentina

Research Article

Estimation of Allele Frequencies and Population Incidence of Wilson Disease in Brazil

Otto PA[1*], Deguti MM[2,3], Araújo TF[3], Barbosa ER[4], Bem RSD[5], Araújo FC[2] and Cançado ELR[2,3]

[1]*Department of Genetics and Evolutionary Biology, Institute of Biosciences, USP, Brazil*

[2]*Department of Gastroenterology, School of Medicine, University of Sao Paulo (USP), Brazil*

[3]*Laboratory of Gastroenterology and Hepatology, Institute of Tropical Medicine, USP, Brazil*

[4]*Department of Neurology, School of Medicine, USP, Brazil*

[5]*Service of Gastroenterology and Hepatology, Department of Internal Medicine, Federal University of Paraná, Brazil*

[*]**Corresponding author**: Paulo A. Otto, PO Box 11461, 05422-970 São Paulo, SP, Brazil, Tel: 55-11-3091-7591; Fax No.: 55-11-3091-7553; E-mail: otto@usp.br

Rec date: June 22, 2016 **Acc date**: Aug 11, 2016 **Pub date**: Aug 14, 2016

Abstract

The present paper deals with the estimation of the overall frequency of *ATP7B* alleles determining Wilson disease (WD) and the population frequency of the condition in Brazil. Genealogical, demographic, and molecular data from 83 WD probands, studied at three distinct WD referral centers, were used for obtaining a population estimate for the overall frequency of alleles that in homozygous or compound heterozygous state determine the frequency of the condition in Brazil; the method we used exploits the relatively high proportion of consanguinity among parents of affected individuals compared with the proportion in the general population. The value we obtained for the overall allele frequency was q = 0.006, with a 95% bootstrap confidence interval of 0.004 to 0.012. The corresponding value for the disease incidence or frequency was P (WD) = 0.000041.

Keywords: Wilson's disease; Gene frequency estimation; *ATP7B* alleles; Disease incidence (frequency); Consanguinity rates

Introduction

Wilson disease (WD) is an inborn error of copper metabolism determined by autosomal recessive mechanism. Affected individuals are homozygous or compound heterozygous as to mutations occurring in the gene *ATP7B* (located at chromosome 13q14.3), that encodes a P-type ATPase copper-transporting protein. Mutations on the *ATP7B* gene lead to a reduction in the conversion of apoceruloplasmin into holoceruloplasmin, an alpha-globulin which in normal conditions carries 95% of all copper present in the plasma. The clinical manifestations of WD result from an important impairment in the biliary excretion of the metal that chronically accumulates in different organs and tissues, particularly in the liver, cornea and the basal ganglia of the brain. Patients with WD exhibit psychiatric, neurological (mainly extra-pyramidal), and/or hepatic signs and symptoms with a variable age of onset (second to fourth decades of life). Neurological manifestations are rare among young patients, but when present, conspicuous dystonia is a very common sign. The Kayser–Fleischer ring that results from the deposit of copper in Descemet's membrane of the cornea is an important diagnostic physical sign found in virtually all cases with neurological manifestations. The differential diagnosis of WD should be considered in all young subjects presenting liver dysfunction signs and/or extra-pyramidal neurological manifestations, mainly when there are other affected subjects in the family or when there exists parental consanguinity. The real practical importance of the condition stems from the fact that it can be successfully managed and treated with chelating agents and zinc salts [1-3].

Like most hereditary conditions, WD is a rare condition, with a population frequency of 10^{-4} to 10^{-6}, and higher frequencies (of the order of 10^{-3} to 10^{-4}) reported exceptionally almost exclusively in very small isolated aggregates. There are at least three referral medical centers of WD in Brazil and the main *ATP7B* mutations of WD Brazilian patients belonging to these centers have been described [4-6]. The present paper deals with (1) the estimation of the overall population frequency of all *ATP7B* alleles that in homozygous or compound heterozygous condition determine WD and (2) the calculation of the disease incidence frequency.

Materials and Methods

For the present paper we used the clinical, genealogical and molecular data from banks of three Brazilian WD reference centers, totaling 324 affected individuals registered at the Services of Gastroenterology and Neurology of the Clinics Hospitals of USP at São Paulo (265 patients), USP at RibeirãoPreto (32 patients) and Universidade Federal of Paraná (UFPR) at Curitiba (27 patients). These data banks include all patients belonging to the samples used in the three molecular studies mentioned above. From the total data bank set of 324 individuals, a sub-sample of 83 patients who meet all of the following criteria has been rigorously selected: (1) have diagnosis of WD clinically well-defined according to the criteria defined by the literature [1,3]; (2) be the proband (the first individual in the family with the WD diagnosis or molecular study performed); (3) have the result of complete genotyping (for direct sequencing) in both alleles of the *ATP7B* gene locus; (4) have signed a participation consent prepared by an Ethics Committee.

In order to estimate the frequency q of the WD gene (that is, the global frequency of all mutant *ATP7B* alleles that in homozygous or compound heterozygous state determine WD), we used the formula q=f/(f+X-1). This formula was obtained from the equations for the frequency of consanguineous marriages among the parents of affected individuals, C=(cq+cfp)/(q+cfp) and for the relative risk (indicating how many times more the disease affects inbred than non-inbred individuals), $X=(q^2+pqf)/q^2=1+pf/q$.

In these formulas, p=1-q is the frequency of the normal (dominant) allele, q^2+pqf and q^2 are respectively the expected frequencies of affected individuals in the offspring of consanguineous and unrelated parents, c is the frequency of consanguineous marriages in the general population, and f is the average coefficient of kinship of the parents of the set of patients who are inbred. These formulas are known in the specialized literature and are briefly discussed in some advanced textbooks on population genetics, such as Cavalli-Sforza et al. [7] on section 7.5 at page 382 and Crow et al. [8] on section 3.5 at page 75 (where f is called "coefficient of consanguinity of the parents").

The derivation of the formula q=f/(f+X-1) is simple, because it is based on intuitive algebraic argumentation: since, C=(cp+cfp)/(q+cfp) it comes out that, 1-C=(q-c)/(q+cfp) so that C/(1-C)=(cq+cfp)/(q-cq) and C(1-c)/(1-C)c = (q+fp)/q = 1+fp/q = X. Therefore an indirect estimate of the relative risk X in the formula q=f/(f+X-1) can be obtained from the equation X=C(1-c)/(1-C)c

In order to evaluate a proper value of q, we have also therefore to provide estimates for the values of f, C and c, by means of data extracted from our sample of WD affected subjects and from the general Brazilian population. Since the q value is estimated indirectly and subject therefore to a relatively large degree of uncertainty, it is important to provide also its corresponding 95% confidence interval. This was obtained averaging the output values of five bootstrap computer simulations of size 10,000 each. Finally, the incidence (frequency) of WD was calculated by means of the formula $P(WD)=q^2+q(1-q)f'$, where q is defined in the previous paragraphs and f' is now the average coefficient of kinship of the whole Brazilian population from which the WD patients were drawn.

Results

Estimate of c: out of a total of 83 probands, each of them belonging to a different family, there is non-ambiguous information about their geographic origin in 80 of them: 21 from the northern or eastern regions (frequency of consanguineous marriages of the order of 0.04), 21 from the southern region, including the state of São Paulo (frequency of consanguineous marriages of the order of 0.01), and 27 from the northeastern region (frequency of consanguineous marriages of the order of 0.08). The average frequency of consanguineous marriages in the general population is therefore calculated as c= (27x0.08+32x0.01+21x0.04)/80=0.042, a rate that is not significantly different from that of 0.046 estimated for the Brazilian population as a whole [9], from which paper we used the above regional frequencies.

Estimate of C: twenty-four probands (out of the total of 83) were born to consanguineous parents. The value of C is therefore taken directly from C=24/83=0.289.

Estimate of f: there exists exact information on the type of parental consanguinity in 16 out of the 24 probands: in 12 cases they are the product of first cousin marriages, in three of first cousins once removed and in one of second cousins (inbreeding coefficients of 1/16, 1/32 and 1/64 respectively). Assuming conservatively a similar distribution of parental consanguinity among the eight individuals who were unable to provide reliable information about their ancestry, the estimate of the average inbreeding coefficient of affected individuals that are the offspring of consanguineous couples is therefore obtained by weighing the figures above, from which we obtain f=(12/16+3/32+1/64)/16 = 0.054

Estimate of X: the value of the relative risk X=C(1-c)/(1-C)c is obtained by putting C=0.289 and c=0.042; we then obtain X=9.395.

Estimate of q: finally, the value of q is calculated using the estimated values of f and X: q=f/(f+X-1)=0.006. Its 95% bootstrap confidence interval is 95% c.i. (q)=(0.004-0.012).

Table 1 shows the frequencies of WD mutant alleles of gene *ATP7B* among affected individuals, estimated by direct counting from the genotypes of 83 WD probands and the values extrapolated for the Brazilian general population, adjusted to the estimated overall value of q = 0.006. Since the average coefficient of kinship of the parents of 24 inbred affected individuals was f = 0.054 and the rest of the probands (59) originated from unrelated couples, the average coefficient for the parents of the whole set of our WD probands is calculated as (24x0.0054+59x0)/83=0.016 This value is about 18 times larger than the estimated value for the whole Brazilian population, f'=0.0009 [9,10]. Using the latter value in the formula $P(WD)=q^2+q(1-q)f'$ we obtain the expected frequency (population incidence) of WD in the Brazilian population, $P(WD)=0.000041=41\times10^{-6}$ or 4 affected individuals per 100,000 persons of both sexes.

alleles	frequencies among affected subjects	adjusted population frequencies
p.L708P	18/166 = 0.1084	0.0007
c.2304dupC	4/166 = 0.0241	~0.0001
p.A1135Qfs	59/166 = 0.3554	0.0021
c.2296dupA	2/166 = 0.0120	~0.0001
p.R1319X	2/166 = 0.0120	~0.0001
p.V949G	1/166 = 0.0060	<0.0001
p.S932X	4/166 = 0.0241	~0.0001
p.M645R	3/166 = 0.0181	~0.0001
p.H1069Q	13/166 = 0.0783	0.0005
c.2438_2440delinsAT	1/166 = 0.0060	<0.0001
p.L1088X	1/166 = 0.0060	<0.0001
p.R778G	1/166 = 0.0060	<0.0001
p.P984R	1/166 = 0.0060	<0.0001
S921N	1/166 = 0.0060	<0.0001
p.P1273L	7/166 = 0.0422	0.0003
p.K1238SfsX20	2/166 = 0.0120	~0.0001
p.A736D	1/166 = 0.0060	<0.0001
p.R778L	1/166 = 0.0060	<0.0001
p.N1270S	3/166 = 0.0181	~0.0001
c.1672-73delGA	1/166 = 0.0060	<0.0001
p.K1258X	1/166 = 0.0060	<0.0001
IVS13-13G/C	1/166 = 0.0060	<0.0001
p.N41S	1/166 = 0.0060	<0.0001
p.G1061E	2/166 = 0.0120	~0.0001
p.T1232P	3/166 = 0.0181	~0.0001
p.D1271G	1/166 = 0.0060	<0.0001
p.D765N	2/166 = 0.0120	~0.0001
p.K1238Sfs	1/166 = 0.0060	<0.0001
p.I1230V	1/166 = 0.0060	<0.0001
p.A1294Qfs	1/166 = 0.0060	<0.0001

p.R1041W	1/166 = 0.0060	<0.0001
c.3713_14delAA	3/166 = 0.0181	~0.0001
p.L1373R	1/166 = 0.0060	<0.0001
IVS 3 +1 G/C	2/166 = 0.0120	0.0001
p.L1275S	1/166 = 0.0060	<0.0001
c.1436delC	1/166 = 0.0060	<0.0001
p.F1094L	1/166 = 0.0060	<0.0001
undetected alleles	16/166 = 0.0964	0.0006
total	166/166 = 1.0000	0.0060

Table 1: Frequencies of WD mutant alleles of gene *ATP7B* among affected individuals, estimated by direct counting from the genotypes found in the set of 83 WD probands. The last column lists the corresponding values extrapolated for the Brazilian general population, adjusted to the overall value of q=0.006.

Discussion

Table 2 lists figures of disease incidence and allele and heterozygous frequencies obtained in recent surveys from the literature, together with our results, which are in the usual range determined for most populations.

Exceptionally large and unusual frequencies, confirmed by molecular analysis, were already known to exist in two European-derived population aggregates, those of the islands of Sardiana [11] in southern Italy and Gran Canaria in the archipelago of Spanish Canary Islands [12]; these extraordinary frequency rates result from the combined effects of random genetic drift (responsible for the increase in WD gene frequency) and increased chance of consanguineous mating (a factor favoring the occurrence of homozygous affected individuals) that usually take place in small population aggregates.

The work of O'Brien et al. [13] was based solely on data from hospital registries and demographic census of Ireland. The molecular data obtained by Gialluisi et al. [11] were used to obtain an estimate of q based on the relative frequencies of homozygous and compound heterozygous affected individuals, a method proposed independently by Petukhova et al. [14] and ten Kate et al. [15] that is specially reliable in homogeneous population aggregates without significant levels of stratification, which is not the case of Brazilian population samples like ours.

All other studies listed on Table 2 obtained frequency rates of WD by means of indirect estimation methods that combine the (total or partial) molecular typing of large numbers of newborn babies or normal random samples of individuals to demographic parameters extracted from census data.

So all incidence rates (including the estimated allele and heterozygous frequencies) shown in Table 2, just like ours, are subject to relatively large levels of variation, with confidence intervals spanning over a large range of values [16-18].

Country or region	Disease incidence $\times 10^6$	Population	Overall Allele frequency $\times 10^4$	Heterozygous frequency $\times 10^4$	Most frequent allele
Canary Islands	385		196	392	p.Leu708Pro
Sardinia	366		191	374	del-441/-427
Brazil	41		60	120	p.A1135Qfs
Korea	32		57	113	p.Arg778Leu
Ireland	26		51	102	-
U.S.A	18		43	86	p.His1069Gln
Venezuela	11		33	66	p.Ala1135Gln

Acknowledgements

The useful comments, suggestions and corrections from two anonymous referees from the editorial board of La Prensa Medica are deeply acknowledged; many of them (some literally) were included in the revised version of this paper. We thank also Drs. Ana De Lourdes Candolo Martinelli (Faculdade de Medicina, USP, Ribeirão Preto), Gilda Porta (Instituto da Criança, Hospital das Clínicas, Faculdade de Medicina, USP), Dominique Muzzillo and HelioTeive (Universidade Federal do Paraná), Alexandre A. Machado and Andreia S Evangelista (Faculdade de Medicina, USP) for their contribution with data collection; Maria Cristina Nakhle, Clarice Pires Abrantes-Lemos, Michele Gomes Gouvêa and João Renato Rebello Pinho for laboratory assistance. We also thank Professor Flair Carrilho for supporting the Wilson´s Disease Reference Center at the Hospital das Clínicas, Faculdade de Medicina, USP). This work was partly supported by grants from FAPESP, CAPES and CNPq.

References

1. Ferenci P (2005) Wilson's Disease. Clin Gastroenterol Hepatol 3: 726-733.
2. Das SK, Ray K (2006) Wilson's disease: an update. Nat Clin Pract Neurol 2: 482-493.
3. Roberts EA, Schilsky ML (2008) American Association for Study of Liver Diseases. Diagnosis and treatment of Wilson disease: an update. Hepatology 47: 2089-2111.

4. Bem RS, Raskin S, Muzzillo DA (2013) Wilson's disease in Southern Brazil: genotype-phenotype correlation and description of two novel mutations in ATP7B gene. Arq Neuropsiquiatr 71: 503-507.
5. Deguti MM, Genschel J, Cancado EL (2004) Wilson disease: novel mutations in the ATP7B gene and clinical correlation in Brazilian patients. Hum Mutat 23: 398-406.
6. Machado AA, Deguti MM, Genschel J (2008) Neurological manifestations and ATP7B mutations in Wilson's disease. Parkinsonism Relat Disord 14: 246-249.
7. Cavalli-Sforza LL, F Bodmer (1971) The genetics of human populations. Freeman, San Francisco.
8. Crow JF, Kimura M (1970) An introduction to theoretical population genetics. Harper & Row, New York.
9. Freire-Maia N (1957) Inbreeding in Brazil. Am J Hum Genet 9: 284-298.
10. Freire-Maia N (1990) Genetic effects in Brazilian populations due to consanguineous marriages. Am J Med Genet 35: 115-117.
11. Gialluisi A, Incollu S, Pippucci T (2013) The homozygosity index (HI) approach reveals high allele frequency for Wilson disease in the Sardinian population. Eur J Hum Genet 21: 1308-1311.
12. Garcia-Villarreal L, Daniels S, Shaw SH (2000) High prevalence of the very rare Wilson disease gene mutation Leu708Pro in the Island of Gran Canaria (Canary Islands, Spain): a genetic and clinical study. Hepatol 32: 1329-1336.
13. Brien M, Reilly M, Sweeney B (2014) Epidemiology of Wilson's disease in Ireland. Mov Disord 29: 1567-1568.
14. Petukhova L, Shimomura Y, Wajid M, Gorrochurn P et al. (2009) The effect of inbreeding on the distribution of compound heterozygotes: a lesson from Lipase H mutations in autosomal recessive woolly hair/hypotrichosis. Hum Hered 68: 117-130.
15. Ten Kate LP, Teeuw M, Henneman L, Cornel MC (2010) Autosomal recessive disease in children of consanguineous parents: inferences from the proportion of compound heterozygous. J Community Genet 1:37-40.