



(GCG)₁₁ founder mutation in the PABPN1 gene of OPMD Uruguayan families

M. Rodríguez^a, C. Camejo^b, B. Bertoni^a, C. Braida^a, M.M. Rodríguez^a,
B. Brais^c, M. Medici^b, L. Roche^{a,*}

^aDepartamento de Genética, Facultad de Medicina, Universidad de la República, Avda. Gral. Flores 2125, 11800 Montevideo, Uruguay

^bInstituto de Neurología, Hospital de Clínicas, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay

^cCentre de Recherche du CHUM, Université de Montreal, Montreal, Que, Canada

Received 9 July 2004; received in revised form 20 October 2004; accepted 27 October 2004

Abstract

The dominant oculo-pharyngeal muscular dystrophy mutation consists of an expanded (GCN)_{12–17} in the coding region of the PolyA Binding Protein Nuclear 1 gene. A founder effect has been demonstrated in Canadian and Bukhara Jewish populations with relatively high prevalence of this disease. Since the oculo-pharyngeal muscular dystrophy prevalence was remarkably high in Southern Uruguay, a founder effect was hypothesized. To identify the ancestral haplotype we determined the (GCN) repeat number and the variants of four intragenic SNPs in Uruguayan OPMD families and a control sample. All families carrying the mutation (GCG)₁₁(GCA)₃(GCG) shared a common ancestral haplotype and the age of the mutation was estimated in 37–53 generations by a composite likelihood method. One family carrying the (GCG)₉(GCA)₃(GCG) allele had a different haplotype. The genealogical and molecular data suggested that the common ancestors were Canary Islands' settlers that arrived in Uruguay in the XIX century.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Haplotype; Linkage analysis; Repeat expansion; Founder effect; OPMD phenotype

1. Introduction

Autosomal dominant Oculo-Pharyngeal Muscular Dystrophy (OPMD) is an adult-onset myopathy with a worldwide distribution [1]. It usually presents during the fifth or sixth decade of life with progressive eyelid ptosis, dysphagia and proximal limb weakness. The autosomal dominant OPMD mutations have been identified as short (GCN)_{12–17} expansions of a (GCN)₁₀ repeat coding for the N terminus poly-alanine domain of the PolyA Binding Protein Nuclear 1 (PABPN1) [2]. Unique nuclear filament inclusions in skeletal muscle fibers are the OPMD pathological hallmark [3]. The mechanism by which protein aggregation in OPMD might relate to a toxic gain-of-function has so far remained elusive. Mutated PABPN1 proteins have been shown to be constituents of the nuclear

filament inclusions that also contain ubiquitin, proteasome subunits, HSP 40, HSP 70, SKIP, and abundant poly(A)-mRNA [4–6]. The recent description of a myopathy phenotype in transgenic mice expressing a 13-alanine stretch human PABPN1 will help the elucidation of OPMD molecular pathologic mechanisms [7].

A spectrum of mutations (GCG)_{8–13}(GCA)₃(GCG) was found in cohorts studied throughout the world [8–10]. A founder effect in populations with high prevalence of OPMD—Bukhara Jews (1/600) and French-Canadians (1/1000)—has been demonstrated by genealogical and haplotyping studies [11,12]. Both populations share the (GCG)₉(GCA)₃(GCG) allele, but the mutations originated independently as shown by haplotype analysis of linked markers [12].

OPMD Uruguayan patients represented 1.6% of the total attendance and 13% of the muscular dystrophies registered at the Neuromuscular Diseases Unit of the Neurology Institute of Uruguay in the 1961–1995 period [13].

* Corresponding author. Tel.: +598 2 924 9562; fax: 598 2 924 9563.

E-mail address: lroche@fmed.edu.uy (L. Roche).

In this study the mutations carried by the Uruguayan families were determined and a founder effect was demonstrated for the (GCG)₁₁(GCA)₃(GCG) mutation by haplotype analysis.

2. Patients and methods

2.1. Patients

Ninety-four individuals, affected patients and relatives, from 22 OPMD Uruguayan families were recruited. All patients were examined by the neurologists (C.C. and M.M.) of the Neuromuscular Diseases Unit of the Neurology Institute of the University Hospital (Hospital de Clínicas, Montevideo, Uruguay). Two control samples were also included for the haplotype analysis: 70 non-related individuals from the southern Uruguayan population (Montevideo and Canelones counties) and 16 unrelated foreign patients heterozygous for (GCN)₁₅/(GCN)₁₀ [Armenia (*N*=1), Canada (*N*=6), France (*N*=1), Japan (*N*=3) Norway (*N*=1), Spain (*N*=3) and South-Africa (*N*=1) (Table 1)].

2.2. Mutation detection

Genomic DNA was obtained from blood samples of the patients who signed an informed consent form. The number of (GCN) repeats was determined by PCR using the primers designed by Brais et al. [2]. The PCR reactions were performed in a total volume of 15 µl containing 80 ng of genomic DNA, 1.5 mg of BSA, 1 µM each primer, 250 µM of dTTP, 250 µM dCTP, 250 µM dATP, 125 µM

dGTP and 125 µM 7-deaza-dGTP (GIBCO), 7.5% DMSO, 1.5 mM MgCl₂, 1 U recombinant Taq DNA polymerase (Invitrogen) with the recommended buffer. Thermo cycling conditions were 95 °C for 5 min followed by 35 cycles of 95 °C for 15 s, 68 °C for 30 s, 72 °C for 30 s and a final elongation step at 72 °C for 7 min. Products were mixed with loading buffer (92% formamide, 0.02 M EDTA, 0.05% bromophenol blue, 0.05% xylene cyanole) and separated through a medium size (17×15 cm) 10% polyacrylamide denaturing gel (0.5 g/ml urea, 19:1 acrylamide/bis-acrylamide), power was set in order to keep temperature in the range of 48–50 °C. A tailor-made molecular weight marker was prepared mixing the PCR products obtained from amplification of pUC18 plasmid DNA with primers designed to obtain fragments of 260, 254, 251, 248 and 245 pb (corresponding to the expected size for 16, 14, 13, 12 and 11 (GCN) repeats, respectively). After electrophoresis gels were silver stained.

2.3. SNPs detection

Four PABPN1 intragenic candidate SNPs were analyzed by PCR-RFLP (Fig. 2): rs2239579 (C/T transition), rs1054084 (C/G transversion), rs14947 (C/T transition) were published in the web site of the National Center for Biotechnology Information (NCBI) SNPs databases and SNP2622 (C/T transition) was described by Blumen et al. [12]. A pair of primers were designed to amplify a PABPN1 gene segment that contained the candidate SNP (rs1054084 forward 5'AAAACAGAAGATGACCTTGATGGA3' and reverse 5'GGGAAGGTAACAAGCAGAACAGTT3', rs14947 forward 5'TCCCCTCCTGCCTGCTCCTGT3' and reverse 5'GGTAGTGCATTTCCCCCGGCGGG3', rs2239579 forward 5'CCTGGATGGGGAAAGTAAGC3' and reverse 5'GAGGCCCAAAGAGCGCGCAGC3', SNP2226 forward 5'CAATGACAATTCTTTCCAGG3' and reverse 5'TGGGATCACCTGAGTCAGAGATGA3'. Thermocycling conditions were the same for the four reactions: 95 °C for 5 min. followed by 35 cycles of 95 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s, and a final step of 72 °C for 5 min. The PCR reactions were performed in a final volume of 25 µl with 40 ng of genomic DNA, 2.5 µg of BSA, 100 µM of each primers, 240 µM of dTTP, 240 µM dCTP, 240 µM dATP, 240 µM dGTP and 1U Taq DNA recombinant polymerase with the recommended buffer (Invitrogen). Different MgCl₂ concentrations were used in each PCR reaction: 0.9 mM for rs1054084, 1.0 mM for rs14947, 1.5 mM rs2239579 and 2.0 mM for SNP2622. The products were digested with a specific restriction enzyme that cleaved the corresponding polymorphic site: *Hha* I (New England-Biolabs) for the restriction site of the C allele of the loci rs1054084 and rs2239579; *Nci* I (New England-Biolabs) for the restriction site of the C allele of the locus rs14947 and *ScrF* I (New England-Biolabs) for the restriction site of the C allele of the locus SNP2226. The PCR product that contains the rs1054084 polymorphism has

Table 1

Genotypes of southern Uruguayan control sample, Uruguayan OPMD affected individuals, and individuals from seven countries carrying the (GCG)₁₁ mutation

	<i>N</i>	rs2239579			SNP2622		
		TT	TC	CC	TT	TC	CC
Uy control ^a	70	44	23	3	2	26	42
Uy OPMD (GCG) ₁₁ ^b	21		20	1		1	20
Uy OPMD (GCG) ₉	1	1					1
Canada	6	1	5		1	4	1
Spain	3	1	2				1
France	1		1				1
Norway	1	1					1
South Africa	1			1			1
Japan	3		2	1		2	1
Armenia	1	1					1

The Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium were tested for the Uruguayan control population.

Uy: Uruguayan.

^a Uruguayan control sample. HWE: rs2239579 *P*=0.709; SNP2622 *P*=1.000; Linkage disequilibrium *P*=0.0027±0.0016.

^b The eldest affected individual of each OPMD family.

an internal (non-polymorphic) restriction site for *Hha* I; for the other SNPs, the primers were modified in the 3' extreme in order to introduce a specific non-polymorphic control site for enzymatic digestion (italicized in the primer sequence). The digested products were resolved on a 1% agarose (ICN)/1.5% high-resolution agarose (SIGMA) gel containing ethidium bromide. The sequence of one PCR product from each reaction was determined in an automatic DNA sequencer (ABI PRISM 377, PERKIN ELMER).

2.4. Statistical analysis

Arlequin program version 2.0 was used to test Hardy-Weinberg equilibrium and linkage disequilibrium on the Uruguayan control sample [14]. The homogeneity test [15] and score test [16] were employed to detect possible associations between genotype of polymorphic loci (rs2239579 and SNP2226) and the (GCG)₁₁ mutation ($\alpha = 0.05$). Genotype data were obtained from the eldest affected patient of each Uruguayan OPMD family ($N = 21$) and controls ($N = 70$).

2.5. (GCG)₁₁ mutation age

The mutation age was estimated by the model described by Labuda [12,17,18]. A population growth rate of 0.006 was used for calculations, as it is considered the growth rate for the European population in the last 100,000 years (J. Reeve, personal communication).

3. Results

The determination of the triplet repeat number of the OPMD Uruguayan patients and their relatives showed that 56 patient genotypes were (GCN)₁₀/(GCN)₁₅ (46 clinically affected and 10 asymptomatic younger patients). One patient genotype was (GCN)₁₀/(GCN)₁₃ and 37 non-affected relatives were homozygous for the non-mutated allele (GCN)₁₀ (Fig. 1). Sequencing the PCR fragments from individuals carrying each mutation showed that the mutated alleles were (GCG)₁₁(GCA)₃(GCG) and (GCG)₉

(GCA)₃(GCG), so they corresponded to (GCG)₆ expansions and will be referred as (GCG)₁₁ and (GCG)₉, respectively.

The SNP haplotypes for the PABPN1 gene (Fig. 2) were determined in non-affected individuals from the southern Uruguayan population ($N = 70$), heterozygous (GCG)₆/(GCG)₁₁ patients from different countries ($N = 16$), and one affected patient from each OPMD Uruguayan family ($N = 22$). Loci rs1054084 and rs14947 were non-polymorphic in the three samples analyzed; all individuals were homozygous GG and TT, respectively. These SNPs were discarded for the subsequent analysis. The three genotype variants for the loci rs2239579 and SNP2622 were observed in the three samples; the genotypic frequencies are shown in Table 1. Both loci were in Hardy-Weinberg equilibrium ($P = 0.709$ and $P = 1.000$ for SNP2622 and rs2239579, respectively) and linkage disequilibrium ($P = 0.0027 \pm 0.0016$) in the control population. A statistically significant association between the (GCG)₁₁ mutation and the C allele of both polymorphisms was found by two different tests: (1) the score test revealed a P value of 0.000002 for rs2239579 and 0.019 for SNP2622, (2) the homogeneity test displayed a P value of 0.000002 for rs2239579 and 0.098 for SNP2622.

Segregation of the SNP2622 and rs2239579 alleles was analyzed in 21 Uruguayan families with the (GCG)₁₁ mutation and one family with the (GCG)₉ mutation. Fig. 3 shows an example of a highly informative OPMD family. The genotype of the individuals III1 and III2 defined the mutated haplotype: rs2239579(C)-(GCG)₁₁-SNP2622(C). Subsequently, the gametic phases of the other individuals were established. For example, the haplotypes of individuals III5 and III6 were rs2239579(T)-(GCG)₆-SNP2622(T) and rs2239579(C)-(GCG)₁₁-SNP2622(C); the haplotypes of the asymptomatic IV3 young woman was rs2239579(T)-(GCG)₆-SNP2622(C) and rs2239579(C)-(GCG)₁₁-SNP2622(C). On the other hand, the haplotypes of the two non-affected members of the family, IV1 and IV2, were rs2239579(T)-(GCG)₆-SNP2622(T) and rs2239579(T)-(GCG)₆-SNP2622(C). Likewise, gametic phases could be accurately established in six other families where the rs2239579(C)-(GCG)₁₁-SNP2622(C) haplotype was segregating. The other 14 (GCG)₁₁ families allowed only partial information on gametic phases, but were concordant with the rs2239579(C)-(GCG)₁₁-SNP2622(C) haplotype. One family showed linkage between rs2239579(C) and (GCG)₁₁ and 13 families showed linkage between SNP2622(C) and (GCG)₁₁. The (GCG)₉ mutation found in a single family was linked to rs2239579(T) and SNP2622(C) alleles, since the patient studied exhibited the rs2239579(TT)-(GCG)_{6/9}-SNP2622(CC) genotype. The genotype data of the foreign patients (Table 1) revealed that their (GCG)₁₁ mutations have arisen in different chromosome environments. For example, the Armenian patient presents the rs2239579(T)-(GCG)₁₁-SNP2622(C) haplotype and a Canadian patient carries the rs2239579(T)-(GCG)₁₁-SNP2622(T) haplotype.

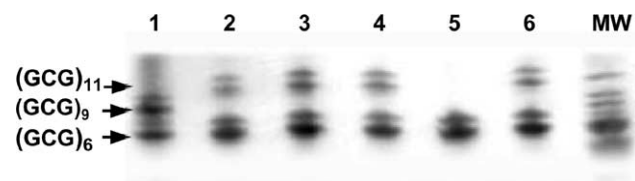


Fig. 1. Silver stained PAGE for quantifying triplet repeat number in PABPN1 gene. Lane 1: Patient carrying the mutation (GCG)₉(GCA)₃(GCG); Lanes 2–4 and 6: Patients heterozygous for the mutation (GCG)₁₁(GCA)₃(GCG); Lane 5: Non-affected individual, homozygous for the allele (GCG)₆(GCA)₃(GCG); MW: molecular weight marker (260, 254, 251, 248 and 245 pb).

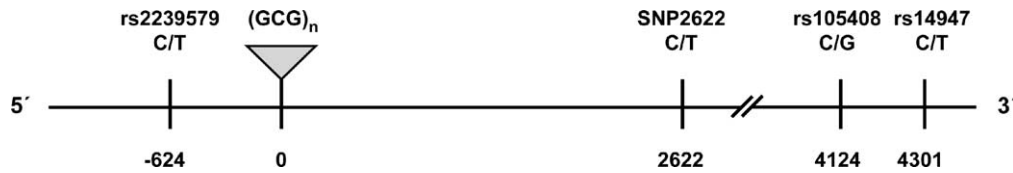


Fig. 2. Map with the (GCG)_n repeat region in PABPN1 gene (GenBank Accession Number AF026029) with relative positions in bp of the SNPs, each polymorphism was identified by the NCBI cluster ID, SNP2622 was described by Blumen et al. [12].

Genealogical information gathered in the interviews revealed that several (GCG)₁₁ OPMD families came from a few small towns of Canelones County close to Montevideo, (14/21). Nine of these families were descendants of Canary Island immigrants. Moreover, several families shared some of the family names. On the other hand, the only family with the (GCG)₉ mutation is original from Cerro Largo County, in the north-eastern part of the country by the Brazilian border and this patient reported that her ancestors came from Brazil.

The haplotype analysis, together with the statistically significant association and the family history data suggest that all the Uruguayan (GCG)₁₁ OPMD families share a common ancestor. The (GCG)₁₁ mutation age was

estimated through a composite likelihood method as 37 generations using marker rs2239579 and 53 generations using marker SNP2622.

4. Discussion

Previously, a founder effect had been demonstrated in two populations with high OPMD incidence. Exhaustive genealogical reconstruction showed that three French sisters introduced the (GCG)₉ mutation found in the French-Canadian population in 1648 [1], while haplotype analysis of the Bukhara Jew OPMD families demonstrated a (GCG)₉

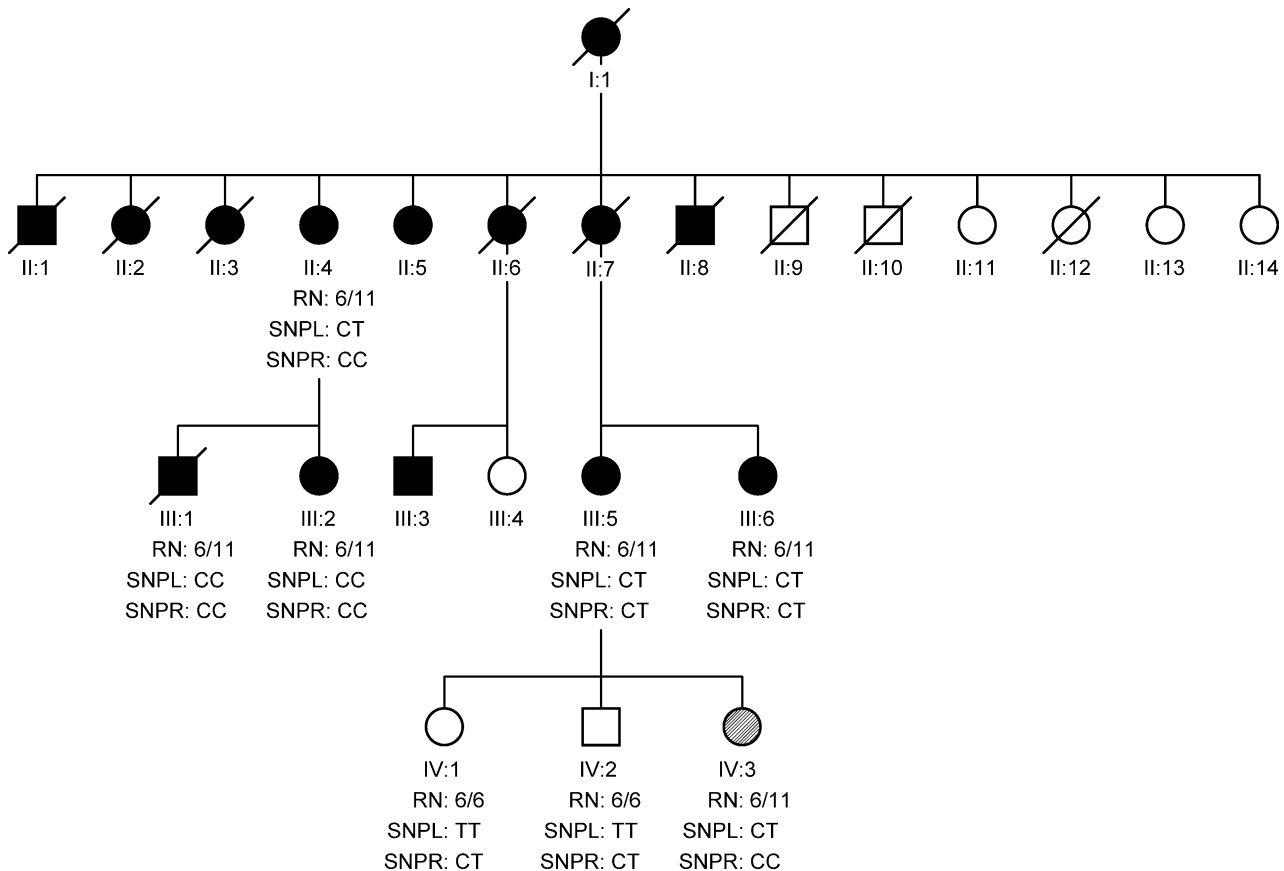


Fig. 3. Genealogical tree of one Uruguayan OPMD family. All affected individuals (filled symbols) are heterozygous for the mutation (GCG)₁₁. The individual IV3 is an asymptomatic young woman (striped symbol). Genotypes of the analyzed individuals are shown below the corresponding symbols. RN: (GCG) repeat number, L, (left) rs2239579 genotype; R, (right) SNP2622 genotype.

mutation founder effect dated between the XIII and XIV centuries [12].

In Uruguayan families early clinical and genealogical studies of OPMD also suggested a founder effect in this population; the incidence of OPMD was unusually high and many of the families manifested to be descendants of Canary Islands immigrants [13]. However, this was not investigated further. In this study we found that all the OPMD Uruguayan families analyzed had the (GCG)₁₁ mutation, excepting one family carrying the (GCG)₉ mutation. The association and family segregation analysis supported the hypothesis of a founder effect with the existence of an ancestral haplotype rs2239579(C)-(GCG)₁₁-SNP2622(C).

The Uruguayan population grew exponentially during the XVIII and XIX centuries, as a consequence of the immigration from Europe. Most settlers from the Canary Islands (16–19% of the European immigration) became established in Canelones County, at the south-central part of the country [19]. Accordingly, 9/21 of the (GCG)₁₁ OPMD families studied had Canary Islands' ancestors and 14/21 families originated in a few small towns of Canelones County. In addition, we also identified a single family carrying a (GCG)₉ mutation inside a different haplotype rs2239579(T)-(GCG)₉-SNP2622(C). It is unlikely that the (GCG)₉ mutation arose as a (GCG)₂ deletion of the (GCG)₁₁ allele, because two independent mutation events should have occurred in a short period of time: the deletion and the C/T substitution in the rs2239579 locus. Moreover, this family came from the northeastern area of the country and its origin could be traced to Brazil. We conclude therefore that the (GCG)₉ and the (GCG)₁₁ mutation had independent origins.

Our data from non Uruguayan OPMD patients show that the (GCG)₁₁ mutation occurred in different haplotypes (Table 1), in agreement with the hypothesis that this expansion has been a recurrent event in human population [2,12]. Two models have been proposed as possible causes of the OPMD expansion: unequal crossing-over and DNA slippage [20]. Both mutations sequenced involved the expansion of the (GCG) tract; no amplification in (GCA) triplet was observed. Therefore, no information favoring one or the other model was obtained.

The age of the (GCG)₁₁ mutation found in OPMD Uruguayan families was estimated as 37–53 generations, using the two polymorphisms employed in the haplotype analysis [17]. A more accurate estimation of the mutation age could be achieved using markers not too close to the mutation -that allow the occurrence of recombination events-, ideally SNPs located 1–10 cm of the mutation [21].

We hypothesize that the (GCG)₁₁ ancestral mutation of most OPMD Uruguayan families may have originated between the X and XIV centuries in the Old World and it was brought by a family of Canary Islands settlers during the immigration waves of the XVIII–XIX centuries. The number of OPMD Uruguayan patients who have been

clinically and genetically examined has contributed to the large-scale international cooperative effort aimed at addressing the main clinical, pathological and genetical issues on OPMD. In particular, the availability of an important number of patients with the (GCG)₁₁ mutation will allow statistical correlation between the size of the (GCN)_n triplet repeat mutation and the severity of the OPMD phenotype.

Acknowledgements

We express our gratitude to all the families who took part in this study. This work was supported by Glaxo-Smith-Klein Laboratory (Uruguay) and Fundación Manuel Pérez, Facultad de Medicina, Universidad de la República (Uruguay).

References

- [1] Brais B, Rouleau GA, Bouchard JP, Fardeau M, Tome FM. Oculopharyngeal muscular dystrophy. *Semin Neurol* 1999;19:59–66.
- [2] Brais B, Bouchard JP, Xie YG, et al. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 1998;18:164–7.
- [3] Tome FM, Chateau D, Helbling-Leclerc A, Fardeau M. Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul Disord* 1997;7(Suppl. 1):S63–S9.
- [4] Calado A, Tome FM, Brais B, et al. Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum Mol Genet* 2000;9:2321–8.
- [5] Fan X, Rouleau GA. Progress in understanding the pathogenesis of oculopharyngeal muscular dystrophy. *Can J Neurol Sci* 2003;30:8–14.
- [6] Brais B. Oculopharyngeal muscular dystrophy: a late-onset poly-alanine disease. *Cytogenet Genome Res* 2003;100:252–60.
- [7] Hino H, Araki K, Uyama E, et al. Myopathy phenotype in transgenic mice expressing mutated PABPN1 as a model of oculopharyngeal muscular dystrophy. *Hum Mol Genet* 2004;13:181–90.
- [8] Mirabella M, Silvestri G, de Rosa G, et al. GCG genetic expansions in Italian patients with oculopharyngeal muscular dystrophy. *Neurology* 2000;54:608–14.
- [9] Nagashima T, Kato H, Kase M, et al. Oculopharyngeal muscular dystrophy in a Japanese family with a short GCG expansion (GCG)₁₁ in PABP2 gene. *Neuromuscul Disord* 2000;10:173–7.
- [10] Hill ME, Creed GA, McMullan TF, et al. Oculopharyngeal muscular dystrophy: phenotypic and genotypic studies in a UK population. *Brain* 2001;124:522–6.
- [11] Bouchard JP. Andre Barbeau and the oculopharyngeal muscular dystrophy in French Canada and North America. *Neuromuscul Disord* 1997;7(Suppl. 1):S5–11.
- [12] Blumen SC, Korczyn AD, Lavoie H, et al. Oculopharyngeal MD among Bukhara Jews is due to a founder (GCG)₉ mutation in the PABP2 gene. *Neurology* 2000;55:1267–70.
- [13] Medici M, Pizzarossa C, Skuk D, Yorio D, Emmanuelli G, Mesa R. Oculopharyngeal muscular dystrophy in Uruguay. *Neuromuscul Disord* 1997;7(Suppl. 1):S50–S2.
- [14] Schneider S, Roessli D, Excoffier L, Arlequin A. A software for population genetics data analysis version 2.000. Switzerland: Genetics and Biometry Laboratory, University Geneva; 2000.
- [15] Weir B. Disequilibrium. In: *Genetics data analysis II. Methods for discrete population genetic data*. Massachusetts, USA: Sinauer Associates; 1996 p. 99–138.

- [16] Clayton D. Population association. In: Balding DJ et al, editor. Handbook of statistical genetics. New York: Wiley; 2001. p. 519–40.
- [17] Labuda D, Zietkiewicz E, Labuda M. The genetic clock and the age of the founder effect in growing populations: a lesson from French Canadians and Ashkenazim. *Am J Hum Genet* 1997;61:768–71.
- [18] Slatkin M, Rannala B. Estimating allele age. *Annu Rev Genomics Hum Genet* 2000;1:225–49.
- [19] Barreto MS. Investigations in human biodiversity (in Spanish), España 2000.
- [20] Nakamoto M, Nakano S, Kawashima S, et al. Unequal crossing-over in unique PABP2 mutations in Japanese patients: a possible cause of oculopharyngeal muscular dystrophy. *Arch Neurol* 2002;59(3):474–7.
- [21] Rannala B, Bertorelle G. Using linked markers to infer the age of a mutation. *Hum Mutat* 2001;18:87–100.