American Journal of Medical Genetics 132A:210–214 (2005)

Research Letter

AGG Interspersion Patterns in the CGG Repeat of the *FMR1* Gene and Linked DXS548/FRAXAC1 Haplotypes in Brazilian Populations

To the Editor:

We have recently described in this journal the distribution of CGG repeats of the *FMR1* gene, the allelic frequencies at DXS548 and FRAXAC1 microsatellite loci, and DXS548/FRAXAC1 haplotype frequencies in African-derived, European-derived, and Amerindian populations from South America [Mingroni-Netto et al., 2002]. The CGG repeat sequences and AGG interspersion patterns were not available at that time.

A number of haplotype studies in affected males with the fragile X syndrome pointed to two main pathways for the origin of the fragile X mutation in European-derived populations. The two most frequent DXS548/FRAXAC1 haplotypes linked to the fragile X mutation have been found to be associated with CGG repeat structures in the general population, which would favor expansions: the haplotype 2-1 (204/158 bp) is associated to long repeats with uninterrupted CGG arrays, leading to predisposed chromosomes, which would increase in size gradually; the haplotype 6-4 (196/152 bp) appears associated to smaller repeats particularly prone to expansions because of the loss of AGG interruptions at the 3′ end of the repeat, also giving rise to long uninterspersed CGG arrays [Eichler et al., 1996]. In the city of São Paulo, Brazil, we observed that these two haplotypes were the most frequent among fragile X patients [Mingroni-Netto et al., 1999]. More recently, Crawford et al. [2000a] found that the two most prevalent haplotypes in African-American fragile X patients were 4-4 (200/152 bp) and 3-4 (202/152 bp). In normal African-American populations, the haplotype 4-4 was linked to CGG repeats devoid of AGG interruptions near the 5′ end of the repeat [Crawford et al., 2000b]. These findings suggested that the pathways leading to fragile X mutations in Africans might be distinct from those leading to mutations in Europeans.

To investigate the pathways leading to fragile X mutations in Africans, we expanded our previous sample of partially isolated communities of African ancestry (*quilombos*) investigating the populations of Abobral, Pedro Cubas, Galvão, and São Pedro in the Ribeira River Valley, state of São Paulo, South-Easthern Brazil (Fig. 1). Runaway slaves who reached the region between 1750 and 1850 founded these communities. A predominantly European-derived sample of 55 unrelated non-carrier brothers of fragile X males, collected in the city of São Paulo, was included in the study for comparison. All individuals were genotyped for DXS548 and FRAXAC1 alleles, as described by Mingroni-Netto et al. [2002]. Allelic frequencies at the two loci are shown in Table I. The frequencies of DXS548/FRAXAC1 haplotypes are in Table II.

For AGG interruption analysis, a subset of 32 African-Brazilian males was selected, including individuals from Pontal, Cajual, and Rio das Rãs, previously studied by Mingroni-Netto et al. [2002]. Repeats on haplotypes known to be in linkage disequilibrium with the fragile X mutation either in Europeans (2-1 and 6-4) or in African-Americans (4-4 and 3-4) were preferentially sequenced, as well as those associated to uncommon haplotypes in European-derived populations. The 55 alleles from São Paulo city population were all sequenced, regardless of the linked haplotypes.

In brief, the CGG repeat and surrounding DNA sequences were amplified from genomic DNA by *Pfx* polymerase (Invitrogen, Carlsbad, CA) with primers *c* and *f* [Fu et al., 1991]. In most samples, the forward and reverse strands were sequenced. Whenever sequencing of only one strand was obtained, the pattern was validated by re-sequencing. In the instances of rare or not previously reported patterns, sequencing was performed at least twice with each primer. After purification, approximately 60 ng of PCR products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Forster City, CA), in an ABI PRISM 310 Genetic Analyzer (Perkin-Elmer, Forster City, CA). CGG repeat interspersion patterns are presented in Figure 2.

Haplotype 2-1 was not found among the four African-Brazilian populations from the state of São Paulo. In one sample, from Rio das Rãs, [Mingroni-Netto et al., 2002], this haplotype was linked to a 23 CGG repeat with one AGG interruption at the 5′ end (Fig. 2a). In the predominantly European-derived sample from São Paulo city, the 2-1 haplotype was associated to four alleles with high repeat numbers ranging from 39–48 repeats. The 39 repeat had three interruptions: $9+9+9+9$ (9 representing the number of CGG's, and the signal + one AGG interruption). The other alleles had two AGG interruptions: $9+9+22$, $9+9+23$, and $9+9+28$ (Fig. 2b).

Haplotype 6-4 was relatively frequent in the African-derived communities (8.3%–19.6%). It was less frequent in São Paulo city (1.8%, present study and 8.3%, Mingroni-Netto et al., 2002]. This is not a rare haplotype in populations hereto studied [Chiurazzi et al., 1996; Bonaventure et al., 1998; Jara

Claudia B. Angeli and Leonardo P. Capelli equally contributed to this study.

Grant sponsor: FAPESP; Grant sponsor: CNPq.

*Correspondence to: Regina C. Mingroni-Netto, Departamento de Biologia, Universidade de São Paulo, C.P. 11461, 05422-970, São Paulo, Brazil. E-mail: renetto@ib.usp.br

Received 11 February 2004; Accepted 28 July 2004

DOI 10.1002/ajmg.a.30376

Fig. 1. Localization of the African-derived populations from Cajual (CA) and Pontal (PO), in the state of Maranhão; Rio das Rãs (RR), in the state of Bahia; and from the Ribeira Valley complex (RV), in the State of São Paulo: AB, Abobral; PC, Pedro Cubas; GA, Galvão; SA, São Pedro. European-derived population from São Paulo city (SP), in the state of São Paulo.

et al., 1998; Peixoto et al., 1998; Mingroni-Netto et al., 1999; Patsalis et al., 1999; Pekarík et al., 1999; Poon et al., 1999; Limprasert et al., 2001], and only a fraction of linked CGG alleles shows AGG interspersion patterns that would predispose to large expansions [Hirst et al., 1993; Kunst and Warren, 1994; Eichler et al., 1996; Crawford et al., 2000a; Sullivan et al., 2002]. The eight CGG alleles from African-Brazilians linked to haplotype 6-4 neither had a long CGG array nor long

TABLE I. Distribution of DXS548 and FRAXAC1 Alleles in Normal X Chromosomes of African-Brazilians (Quilombo Communities) and Euro-Brazilians (São Paulo City)

	Qullombos								Euro-Brazilians	
	Abobral		Pedro Cubas		Galvão		São Pedro		São Paulo	
	N	%	N	%	N	%	N	%	N	%
DXS548 alleles										
1 (206 bp)	1	0.9	0		1	1.5	0		0	
2 (204 bp)	2	1.7	2	1.8	4	6.0	3	3.9	6	11.0
3 (202 bp)	6	5.3	2	1.8	0		0		0	
4 (200 bp)	7	6.1	10	9.1	9	13.2	9	11.8	0	
5 (198 bp)	25	21.9	20	18.2	0		5	6.6	2	3.6
6 (196 bp)	24	21.0	16	14.5	14	20.6	12	15.8	3	5.4
7 (194 bp)	49	43.0	59	53.6	39	57.3	45	59.2	43	78.2
8 (192 bp)	0		0		1	1.5	2	2.6	1	1.9
9 (190 bp)	0		1	0.9	0		0		0	
Total	114		110		68		76		55	
Expected heterozygosity	0.72 ± 0.026		0.692 ± 0.038		0.635 ± 0.064		0660 ± 0.059			
Observed heterozygosity	0.500 ± 0.107		0.459 ± 0.121		0.524 ± 0.151		0.370 ± 0.153		—	
(F)	0.307		0.336		0.175		0.439		—	
FRAXAC1 alleles										
1 (158 bp)	3	2.6	2	2.0	3	4.2	0		5	9.1
2 (156 bp)	4	3.5	4	4.0	15	20.8	22	28.9	5	9.1
3 (154 bp)	56	49.5	43	42.6	27	37.5	37	48.7	39	71.0
4 (152 bp)	40	35.4	35	34.6	27	37.5	16	21.0	5	9.1
5 (150 bp)	10	8.8	17	16.8	0		1	1.3	0	
6 (148 bp)	0		0		0		0		1	1.8
Total	113		101		72		76		55	
Expected heterozygosity	0.626 ± 0.036		0.683 ± 0.031		0.669 ± 0.034		0.640 ± 0.033			
Observed heterozygosity	0.738 ± 0.079		0.576 ± 0.113		0.739 ± 0.106		0.630 ± 0.117		—	
(F)	−0.179		0.157		−0.105		0.016		—	

TABLE II. DXS548/FRAXAC1 Haplotypes of African-Brazilian Males and Females Whose Haplotypes Could Be Determined (Homozygotes for One of the Microsatellites), and of Euro-Brazilian Males

	Quilombos								Euro-Brazilians São Paulo	
	Abobral		Pedro Cubas		Galvão		São Pedro		São Paulo	
DXS548/FRAXAC1 haplotypes (bp)	N	%	N	%	N	%	N	%	N	%
7-3 (194/154)	23	28.4	28	38.9	17	36.9	26	48.1	36	65.4
7-2 (194/156)	2	2.5	0		8	17.4	8	14.8	4	7.3
[a]2-1 (204/158)	0		0		0		0		4	7.3
2-4 (204/152)	0		1	1.4	0		1	1.8	2	3.6
6-3 (196/154)	6	7.4	3	4.2	0		3	5.5	2	3.6
[a]6-4 (196/152)	12	14.8	6	8.3	9	19.6	5	9.3	1	1.8
7-4 (194/152)	9	11.1	14	19.4	5	10.9	4	7.4	1	1.8
5-4 (198/152)	6	7.4	2	2.8	0		1	1.8	1	1.8
7-1 (194/158)	0		0		0		0		1	1.8
7-6 (194/148)	0		0		0		0		1	1.8
5-2 (198/156)	0		0		0		1	1.8	1	1.8
8-3 (192/154)	0		0		0		0		1	1.8
[b]4-4 (200/152)	3	3.7	3	4.2	6	13.0	2	3.7	0	
5-3 (198/154)	7	8.6	4	5.5	0		0		0	
3-3 (202/154)	3	3.7	0		0		0		0	
4-2 (200/156)	0		2	2.8	0		0		0	
4-5 (200/150)	0		2	2.8	0		0		0	
5-5 (198/150)	3	3.7	6	8.3	0		0		0	
4-3 (200/154)	2	2.5	0		0		0		0	
[b]3-4 (202/152)	2	2.5	0		0		0		0	
2-3 (204/154)	1	1.2	0		0		0		0	
6-1 (196/158)	1	1.2	0		0		0		0	
6-2 (196/156)	0		0		0		1	1.8	0	
2-2 (204/156)	0		0		0		1	1.8	0	
Others	1	1.2	1	1.4	1	2.2	1	1.8	0	
Total	81		72		46		54		55	

[a]Most frequent haplotypes described in fragile X chromosomes in European-derived populations.
[b]Most frequent haplotypes described in fragile X chromosomes in African-Americans.

uninterrupted tracts. However, two of them had asymmetrical interruptions (9 + 12 + 9 and 9 + 17 + 4). Eichler et al. [1996] suggested that asymmetrical arrays were more likely to undergo 3′ AGG loss, increasing the risk of further expansions. In the European-derived population from São Paulo city, this haplotype occurred once associated to an allele with 47 repeats, 9 + 9 + 27, also suggesting a higher risk of expansion.

Haplotype 4-4 was rare in the majority of African-derived populations (3.7%–4.2%), but it was found in 6 out of 46 chromosomes (13%) investigated in the community of Galvão. The high frequency in this small inbred population, but not in the others, is indicative of a founder effect. The three CGG alleles linked to 4-4 haplotype showed (a) a long uninterrupted 25 CGG sequence (from Galvão), (b) a sequence with only one AGG near the 5′ end (9 + 21, from Abobral), or (c) a long uninterrupted array at the 5′ end of the repeat (20 + 9, from Abobral). These findings are in accordance with the observation of Crawford et al. [2000b] that long uninterrupted sequences and sequences devoid of 5′ AGG interspersions are associated with this 4-4 haplotype, leading to the hypothesis that this haplotype in African-Americans is linked to CGG repeats favoring expansions. Haplotype 4-4 was not detected in Euro-Brazilians from São Paulo city.

The other high risk haplotype in African-Americans highlighted by Crawford et al. [2000a], 3-4, was also rare among African-Brazilians, and appeared in only one population (2.5% in Abobral). The only sequenced allele on this haplotype had a structure (10 + 9 + 10) that does not seem prone to expansion.

We observed additional CGG alleles with peculiar interruption patterns associated with other haplotypes. On the haplotype 7-6 (194/148 pb), one allele from an African-Brazilian had 39 repeats and three AGG interruptions (9 + 9 + 9 + 9). Indeed Crawford et al. [2000b] found that, among African-Americans, this haplotype appeared in association with long CGG repeats with three or only one AGG interruption near the 5′ end. On the other hand, the allele linked to the 7-6 haplotype in our European-derived sample had a common pattern of 29 repeats with two AGG interruptions (9 + 9 + 9). In this group, another rare AGG interspersion pattern (10 + 7 + 1 + 9) was found on the haplotype 7-3 (194/154 pb). This pattern was reported by Larsen et al. [2001] in Norway. An African-Brazilian allele associated to haplotype 5-3 had a distinctive interruption pattern (9 ++ 9 ++ 9) not reported so far.

Our results, thus, confirm that haplotype 2-1 is linked to long CGG repeats in European-derived populations. In individuals of African ancestry the haplotype 4-4 appeared in association with uninterrupted CGG sequences that would be at risk of expanding as reported by Crawford et al. [2000b]. The rarity of 3-4 haplotype among African-Brazilians did not allow investigating the nature of its reported association with fragile X syndrome in African-Americans.

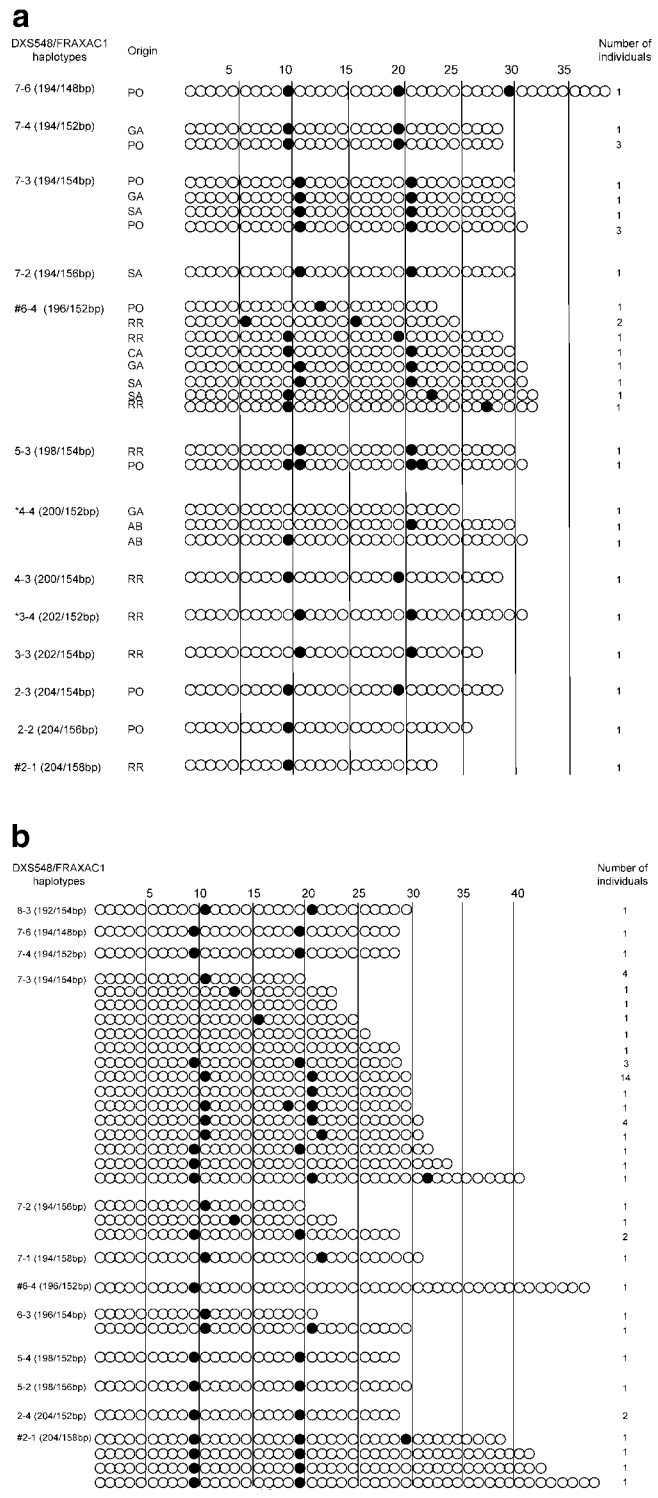

Fig. 2. CGG repeat interspersion patterns and linked haplotypes DXS548/FRAXAC1 observed in (**a**) 32 African-Brazilians, (GA) Galvão; (SA) São Pedro; (RR) Rio das Rãs; (PO) Pontal populations. African-Brazilian alleles on haplotypes reported to be in linkage disequilibrium with fragile X mutations were preferentially sequenced. **b**: 55 individuals from the mainly European-derived population from São Paulo. ○ CGG trinucleotide; ● AGG trinucleotide; #most frequent haplotypes in fragile X chromosomes in European-derived populations. *Most frequent haplotypes in fragile X chromosomes in African-Americans.

ACKNOWLEDGMENTS

We are thankful to the individuals from quilombos and fragile X families, and to João Pedro Vicente, MD, for clinical evaluation of individuals in Ribeira River Valley. We also thank Fundação Cultural Palmares and Ministério da Saúde do Brasil.

REFERENCES

Bonaventure G, Torrado M, Barreiro C, Chertkoff L. 1998. Fragile X founder effects in Argentina. Am J Med Genet 49:200–204.

Chiurazzi P, Genuardi M, Kozak L, Giovannucci-Uzielli ML, Bussani C, Dagna-Bricarelli F, Grasso M, Perroni L, Sebastio G, Sperandeo MP, Oostra BA, Neri G. 1996. Fragile X founder chromosomes in Italy: A few initial events and possible explanation for their heterogeneity. Am J Med Genet 64:209–215.

Crawford DC, Schwartz CE, Meadows KL, Newman JL, Taft LF, Gunter C, Brown WT, Carpenter NJ, Howard-Peebles PN, Monaghan KG, Nolin SL, Reiss AL, Feldman GL, Rohlfs EM, Warren ST, Sherman SL. 2000a. Survey of the fragile X syndrome CGG repeat and the short-tandem-repeat and single-nucleotide-polymorphism haplotypes in an African American population. Am J Hum Genet 66:480–493.

Crawford DC, Zhang F, Wilson B, Warren ST, Sherman SL. 2000b. Fragile X CGG repeat structures among African-Americans: Identification of a novel factor responsible for repeat instability. Hum Mol Genet 9:1759–1769.

Eichler EE, Macpherson JN, Murray A, Jacobs PA, Chakravarti A, Nelson DL. 1996. Haplotype and interspersion analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin on the fragile X syndrome. Hum Mol Genet 5:319–330.

Fu Y-K, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richard S, Verkerk AJMH, Holden JJA, Fenwick RG Jr, Warren ST, Oostra BA, Nelson DL, Caskey CT. 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman Paradox. Cell 67:1047–1058.

Hirst MC, Knight SJ, Christodoulou Z, Grewal PK, Fryns JP, Davies KE. 1993. Origins of the fragile X syndrome mutation. J Med Genet 30:647–650.

Jara L, Aspillaga M, Avendaño I, Obreque V, Blanco R, Valenzuela CY. 1998. Distribution of (CGG)n repeats and *FMR1* associated microsatellite alleles in a normal Chilean population. Am J Med Genet 75:277–282.

Kunst CB, Warren ST. 1994. Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77:853–861.

Larsen LA, Vuust J, Nystad M, Evseeva I, Van Ghelue M, Tranebjaerg L. 2001. Analysis of FMR1 (CGG)(n) alleles and DXS548-FRAXAC1 haplotypes in three European circumpolar populations: Traces of genetic relationship with Asia. Eur J Hum Genet 9:724–727.

Limprasert P, Saechan V, Ruangdaraganon N, Sura T, Vasiknanote P, Jaruratanasirikul S, Brown WT. 2001. Haplotype analysis at the FRAXA locus in Thai subjects. Am J Med Genet 98:224–229.

Mingroni-Netto RC, Costa SS, Angeli CB, Vianna-Morgante AM. 1999. DXS548/FRAXAC1 haplotypes in fragile X chromosomes in the Brazilian population. Am J Med Genet 84:204–207.

Mingroni-Netto RC, Angeli CB, Auricchio MT, Leal-Mesquita ER, Ribeiro-dos-Santos AK, Ferrari I, Hutz MH, Salzano FM, Hill K, Hurtado AM, Vianna-Morgante AM. 2002. Distribution of CGG repeats and FRAXAC1/DXS548 alleles in South American populations. Am J Med Genet 111:243–252.

Patsalis PC, Sismani C, Hettinger JA, Boumba I, Georgiou I, Stylianidou G, Anastasiadou V, Koukoulli R, Pagoulatos G, Syrrou M. 1999. Molecular screening of fragile X (FRAXA) and FRAXE mental retardation syndromes in the Hellenic population of Greece and Cyprus: Incidence, genetic variation, and stability. Am J Med Genet 84:184–190.

Peixoto A, Santos MR, Seruca R, Amorim A, Castedo S. 1998. Analysis of *FMR1* and flanking microsatellite markers in normal and fragile X chromosomes in Portugal: Evidence for a "protector" haplotype. Eur J Hum Genet 6:518–522.

Pekarík V, Blazková M, Kozák L. 1999. Haplotype analysis of the fragile X syndrome gene *FMR1* in the Czech Republic. Am J Med Genet 84:214–216.

Poon PM, Pang CP, Chen QL, Zhong N, Lai KY, Lau CH, Wong CK, Brown WT. 1999. FRAXAC1 and DXS548 polymorphisms in the Chinese population. Am J Med Genet 84:208–213.

Sullivan AK, Crawford DC, Scott EH, Leslie ML, Sherman SL. 2002. Paternally transmitted FMR1 alleles are less stable than maternally transmitted alleles in the common and intermediate size range. Am J Hum Genet 70:1532–1544.

Claudia B. Angeli
Leonardo P. Capelli
Maria Teresa B.M. Auricchio
Angela M. Vianna-Morgante
Regina C. Mingroni-Netto*
Centro de Estudos do Genoma Humano
Departamento de Biologia
Instituto de Biociências
Universidade de São Paulo
São Paulo, Brazil

Emygdia R. Leal-Mesquita
Departamento de Biologia
Universidade Federal do Maranhão
Maranhão, Brazil

Ândrea K.C. Ribeiro-dos-Santos
Centro de Ciências Biológicas
Universidade Federal do Pará
Pará, Brazil

Iris Ferrari
Silviene F. Oliveira
Maria de Nazaré Klautau-Guimarães
Departamento de Genética e Morfologia
Instituto de Ciências Biológicas
Universidade de Brasília
Brasília, Brazil