

# *FMR1* Linked Haplotype Analysis in a Mentally Retarded Male Population

## Research Article

Zanda Daneberga<sup>1,2\*</sup>, Natalija Pronina<sup>1,2</sup>,  
Baiba Lace<sup>1</sup>, Rita Lugovska<sup>1,2</sup>

<sup>1</sup>Medical Genetic Clinic, University Children's Hospital,  
Juglas Street 20, LV-1079, Riga, Latvia

<sup>2</sup>Department of Biology and Microbiology, Riga Stradins University,  
Dzirnciema Street 16, LV-1007, Riga, Latvia

Received 10 April 2011; Accepted 17 August 2011

**Abstract:** Fragile X syndrome is caused by dynamic mutation of *FMR1* gene CpG island CGG repeats. The underlying mutational mechanism is not fully understood. Different microsatellite markers and SNP have previously been reported as markers associated with *FMR1* CGG repeat instability. The aim of the present study was to identify specific haplotypes among Latvian FXS patients and the control group with respect to allelic stability. Eleven male FXS patients and 122 control male patients participated in the study. In total, 27 different DXS548-FRAXAC1-ATL1-FRAXAC2 haplotypes were found. The prevalent haplotype in the control group was 7-4-A-5+ (rel. frequency 0.327). The prevalent haplotype associated with the FXS group was 2-2-G-4 (rel. frequency 0.818;  $p < 0.0001$ ). Grey zone alleles with a long uninterrupted CGG tract at the 3' end were significantly associated with the 2-2-G-4 haplotype ( $p = 0.0022$ ). Our findings suggest that, for the Latvian population, the haplotype 2-2-G-4 is a marker of CGG tract instability. We conclude that a founder effect could not be an explanation for our findings on the basis of heterogeneity exhibited by the Latvian population and lack of studies throughout this geographical region. This data may provide evidence of different mutational pathways of expansion in the Baltic States region.

**Keywords:** Fragile X Syndrome • *FMR1* gene • DXS548 • FRAXAC1 • FRAXAC2 • ATL1 • Haplotype

© Versita Sp. z o.o.

## 1. Introduction

Fragile X syndrome (FXS; OMIM #300624; FRAXA, Xq27.3) is a well known common cause of X-linked mental retardation. The prevalence of FXS full mutation is 1/4000 - 6000 males and 1/8000 - 10000 females. The syndrome is caused by a dynamic mutation of *FMR1* gene CpG island CGG repeats [1-4]. Among individuals from the general population, the polymorphic CGG repeat ranges from 6 to 50 repeats and is usually interspersed every 9-10 repeats with an AGG [5,6]. Premutation alleles have a moderate expansion of the repeat (from 50 to ~200 units), they are unmethylated on an active X chromosome and do not affect *FMR1* expression. CGG repeat expansion over 200 is the basis for CpG island methylation, leading to silencing of the *FMR1* gene [7]. Intermediate or grey zone alleles

are poorly defined. Boundaries for the grey zone range vary among studies, from 34 or 35 CGG repeats for the lower boundary to 58/60 repeats for the upper boundary [7-10]. These alleles are often transmitted stably, but are more likely to exhibit unstable transmission with increasing size within this range.

The underlying mutational mechanism is not fully understood and remains a topic of debate. The gender of the parent carrying an expanded repeat (maternal imprinting), the number of repeats (dynamic mutation) and the absence of AGG interruptions in long tracts of CGG repeats have been described as the main factors related to this instability [5,9,11]. The microsatellite markers DXS548-FRAXAC1-FRAXAC2 and the ATL1 SNP have previously been reported as markers associated with *FMR1* CGG repeat instability [5,12-17].

Haplotypes linked to FXS are widely described

\* E-mail: zanda.daneberga@gmail.com

across Western European and Scandinavian populations; however, less is known regarding populations from Eastern Europe, including the Baltic States. The aim of the present study was to identify specific haplotypes among Latvian FXS patients and control mental retardation patients with a normal number of CGG repeats with respect to allelic stability.

## 2. Material and Methods

### 2.1. Patients

Eleven unrelated male FXS patients, confirmed by DNA analysis, participated in this study. The control group comprised of 122 unrelated male patients exhibiting mental retardation, referred by clinical geneticists for routine FRAXA screening, with a normal number of *FMR1* CGG repeats. The Latvian Central Committee of Medical Ethics and the Riga Stradins University Committee of Medical Ethics approved the study.

### 2.2 Analysis of the CGG repeat and AGG interspersed pattern Haplotype analysis

The CGG repeat number was determined by fluorescent PCR on an ABI Prism® 310 genetic analyser following the manufacturer's protocol (Applied Biosystems, USA). Twenty-six patients with grey zone alleles (35–50 repeats) were detected and subsequently selected for AGG interspersed pattern analysis by direct sequencing.

For CGG repeat pattern analysis, the flanking DNA sequences were amplified from genomic DNA using the PCR protocol described by Chong *et al.* [17] PCR products were concentrated and purified for the sequencing reaction by a Millipore Montage PCR filter device.

The sequencing reaction was performed using a BigDye® Terminator v3.1 kit (Applied Biosystems) according to the manufacturer's protocol. Due to the high G/C content of the template, 1 µl of DMSO and 0.5 µl of glycerol were added to the reaction. The primer used was 5'-GAC GGA GGC GCC GCT GCC AGG-3' [19]. The reaction was carried out on a PCR Eppendorf Mastercycler. Subsequent purification of the sequencing products was performed as recommended by the manufacturer. All sequencing reaction products were run on an ABI Prism® 310 genetic analyser and analysed by ABI DNA sequencing software.

### 2.3. Haplotype analysis

The ATL1 polymorphism (alleles A/G located 5613 bp upstream of the CGG repeat) was analysed by following the allele-specific oligonucleotide PCR protocol described by Dombrowski *et al.* [11].

Case-control haplotype analysis between the FXS patients and the control group was performed using the microsatellite markers DXS548, FRAXAC1 and FRAXAC2. The DXS548 microsatellite is located 189895 bp downstream of the CGG repeat, the FRAXAC1 microsatellite is located 7221 bp downstream of the CGG repeat and the FRAXAC2 microsatellite is located 12418 bp upstream of the CGG repeat. Multiplex PCR for DXS548 and FRAXAC2 was performed in a total reaction volume of 15 µl, containing 1× PCR reaction buffer (75 mM Tris-HCl (pH 8.8 at 25°C), 20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.01% Tween 20), 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 7% glycerol, 2 pmol of each primer DXS548A (5'-HEX-AGA GCT TCA CTA TGC AAT GGAATC-3') and DXS548B (5'-GTA CAT TAG AGT CAC CTG TGG TGC-3') [20], 1 pmol of each primer FRAXAC2A (5'-6-FAM-GAC TGC TCC GGAAGT TGAATC CTC A-3') and FRAXAC2B (5'-CTA GGT GAC AGA GTG AGA TCC TGT C-3') [20], and 0.15 U True start™ Hot Start *Taq* DNA polymerase. PCR was carried out by an initial step of 2 min at 95°C, then 10 cycles of 30 sec at 95°C, 1 min at 60°C and 1 min at 72°C. This was followed by a second round of amplification comprising 25 cycles of 30 sec at 95°C, 1 min at 55°C and 1 min at 72°C, and a final extension at 72°C for 7 min.

The FRAXAC1 microsatellite marker was amplified separately with 4.5 pmol of primer FRAXAC1A (5'-NED-GAT CTA ATC AAC ATC TAT AGA CTT TAT T-3') and FRAXAC1B (5'-AGA TTG CCC ACT GCA CTC CAA GCC T-3') [20] under the same PCR conditions.

Multiplex reaction product (0.5 µl) was mixed with an equal volume of GeneScan™ ROX 500™ size standard and 24 µl of deionised formamide. The lengths of product fragments were determined on an ABI Prism® 310 genetic analyser.

Genotyping results were analysed by GeneScan® Analysis software. The corresponding peak's length was calculated according to the calibration curve of GeneScan™ ROX 500™ size standard. Nomenclature for alleles was adjusted to the nomenclature recommended by Macpherson *et al.* [14].

Genotyping results were validated by direct sequencing of random alleles for each microsatellite marker. For each marker allele, the same PCR conditions as described above for genotyping were used, except the reverse primers A were not labelled with fluorescent dye. PCR products were concentrated and purified for the sequencing reaction by a Millipore Montage PCR filter device.

The sequencing reaction was performed using a BigDye® Terminator v3.1 kit according to the manufacturer's protocol. The reaction was carried out on a PCR

Eppendorf Mastercycler. Subsequent purification of the sequencing products was performed as recommended by the manufacturer. All sequencing reaction products were run on an ABI Prism® 310 genetic analyser and analysed by ABI DNA sequencing software.

### 2.3. Statistics

The statistical calculations of AMOVA for the haplotype analysis and level of heterozygosity for all polymorphisms were done using Arlequin 3.5 software [21]. Comparison of obtained data was performed by Fisher's exact test of 2×2 contingency tables and chi-square (<http://www.graphpad.com/quickcalcs/index.cfm>). The differences were considered statistically significant at  $p < 0.05$ . In case-control analysis Bonferroni correction was applied for multiple testing.

## 3. Results

### 3.1. ATL1 SNP

The control group CGG repeat alleles were analysed with respect to ATL1 alleles. The control group was divided into two subgroups according to CGG allele repeat number – normal size alleles and grey zone alleles (Table 1). The relative frequency (RF) of allele A in the normal subgroup was 0.563, for allele G it was 0.438. In the grey zone subgroup, allele A had a RF of 0.308, for allele G the RF was 0.692. Allele G differed significantly between normal size alleles and grey zone alleles ( $p = 0.0271$ ).

**Table 1.** Frequencies of the ATL1 SNP in CGG Alleles.

|              | Normal CGG alleles<br>(16–34 repeats) |       | Grey zone CGG alleles<br>(35–50 repeats) |        |
|--------------|---------------------------------------|-------|------------------------------------------|--------|
|              | n                                     | RF    | n                                        | RF     |
| <b>ATL1</b>  |                                       |       |                                          |        |
| <b>A</b>     | 54                                    | 0.563 | 8                                        | 0.308  |
| <b>G</b>     | 42                                    | 0.438 | 18                                       | 0.692* |
| <b>Total</b> | 96                                    | 1.000 | 26                                       | 1.000  |

*n* = number of chromosomes; *RF* = relative frequency; \*  $p < 0.05$

A statistically significant association between individual CGG alleles and the ATL1 SNP was found for allele 29 and G ( $p = 0.001$ ); 30 CGG repeats and A ( $p < 0.0001$ ) and allele 31 and A ( $p = 0.0013$ ).

All FXS group chromosomes were found to be exclusively associated with ATL1 allele G, and this association was statistically significant ( $p = 0.0008$ ).

### 3.2. Repeat structures of grey zone alleles

Twenty-six grey zone alleles were sequenced in order to characterise the AGG interruption within the CGG repeat. Twelve chromosomes exhibited a CGG interspersed pattern with three AGG, 12 chromosomes with two AGG, one chromosome with one AGG and one pure CGG tract (Table 2). For all chromosomes, the loss of AGG was detected at the 3' end of the sequence.

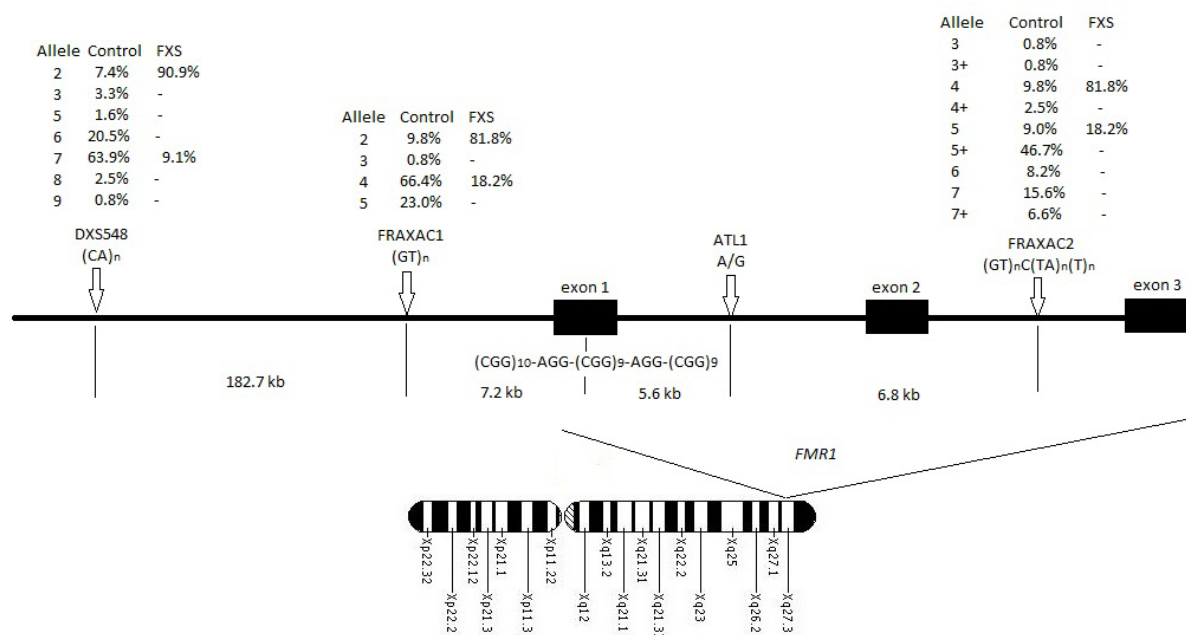
**Table 2.** AGG Interspersion Pattern and Linked DXS548-FRAXAC1-ATL1-FRAXAC2 Haplotype Frequencies of Grey Zone CGG Alleles.

| DXS548       | FRAXAC1 | ATL1 | FRAXAC2 | CGG | AGG       | n         | RF           |
|--------------|---------|------|---------|-----|-----------|-----------|--------------|
| 2            | 2       | G    | 4       | 38  | 9+9+18    | 8         | 0.308**      |
|              |         |      |         | 39  | 9+29      |           |              |
|              |         |      |         | 40  | 9+9+20    |           |              |
|              |         |      |         | 40  | 9+9+20    |           |              |
|              |         |      |         | 41  | 9+9+21    |           |              |
|              |         |      |         | 45  | 9+9+25    |           |              |
|              |         |      |         | 47  | 9+9+27    |           |              |
|              |         |      |         | 50  | 9+9+30    |           |              |
| 7            | 4       | A    | 5+      | 39  | 10+9+9+8  | 6         | 0.231        |
|              |         |      |         | 41  | 10+9+9+10 |           |              |
|              |         |      |         | 41  | 10+9+10+9 |           |              |
|              |         |      |         | 41  | 10+9+10+9 |           |              |
|              |         |      |         | 42  | 10+9+21   |           |              |
|              |         |      |         | 43  | 10+9+22   |           |              |
| 6            | 5       | G    | 7+      | 37  | 9+10+6+9  | 2         | 0.077        |
|              |         |      |         | 38  | Pure      |           |              |
| 7            | 4       | G    | 6+      | 39  | 9+9+9+9   | 2         | 0.077        |
|              |         |      |         | 39  | 9+9+9+9   |           |              |
| 7            | 4       | G    | 5       | 39  | 9+9+9+9   | 2         | 0.077        |
|              |         |      |         | 39  | 9+9+9+9   |           |              |
| 6            | 4       | G    | 5       | 41  | 9+9+21    | 2         | 0.077        |
|              |         |      |         | 42  | 9+9+22    |           |              |
| 6            | 5       | A    | 7       | 37  | 9+10+6+9  | 1         | 0.038        |
| 6            | 4       | A    | 5+      | 35  | 10+6+8+8  | 1         | 0.038        |
| 3            | 2       | G    | 4       | 38  | 9+9+18    | 1         | 0.038        |
| 7            | 4       | G    | 5+      | 36  | 10+9+5+9  | 1         | 0.038        |
| <b>Total</b> |         |      |         |     |           | <b>26</b> | <b>1.000</b> |

*RF* = relative frequency; *AGG* = pattern of CGG tract, the digit corresponds to the CGG repeat number and '+' denotes the AGG interspersed position; \*\*  $p < 0.01$ .

### 3.3. DXS548-FRAXAC1-ATL1-FRAXAC2 haplotypes

Several microsatellite markers and the ATL1 SNP were analysed in the FXS patient group and the control group. Seven different microsatellite alleles were detected with respect to the DXS548 locus, four alleles were identified regarding the FRAXAC1 locus and nine alleles were found for the FRAXAC2 locus. The frequencies of the described alleles in the control and FXS group are shown in Figure 1.



**Figure 1.** The location, nomenclature and frequencies of *FMR1* linked STR- and SNP- polymorphisms, tested in this study.

The level of heterozygosity for all polymorphisms was examined under finite island model and compared (Table 3).

**Table 3.** Expected Heterozygosity and Observed Heterozygosity among the Control and FXS Groups.

| Locus    | Control $H_E$ | FXS $H_E$ | Mean  | S.D.  | $H_T$ | $H_O$ | $X^2$  | p-value |
|----------|---------------|-----------|-------|-------|-------|-------|--------|---------|
| ATL1 SNP | 0.504         | 0.000     | 0.252 | 0.356 | 0.501 | 0.696 | 20.332 | <0.0001 |
| FRAXAC1  | 0.501         | 0.327     | 0.414 | 0.123 | 0.545 | 0.863 | 53.654 | <0.0001 |
| FRAXAC2  | 0.734         | 0.327     | 0.530 | 0.288 | 0.757 | 1.055 | 62.591 | <0.0001 |
| DXS548   | 0.546         | 0.182     | 0.364 | 0.258 | 0.594 | 1.009 | 94.310 | <0.0001 |

$H_E$  = expected heterozygosity;  $H_T$  = total heterozygosity;  $H_O$  = observed heterozygosity; S.D. = standard deviation;  $X^2$  = chi-square.

The identified markers were combined to form haplotypes. In total, 27 different haplotypes were detected – 26 in the control group, three in the FXS group. Only one haplotype from the FXS group was unique (Table 4). The most common haplotype among the control group chromosomes was 7-4-A-5+ (RF = 0.327;  $p$  = 0.0336). Among FXS patients, the prevalent haplotype was 2-2-G-4 (RF = 0.818;  $p$  < 0.0001).

Statistical analysis of the data revealed that the molecular variation between the groups was 27.04% and within the groups was 72.96%. The fixation index  $F_{st}$ , calculated based on haplotype frequencies between the control and FXS group, was 0.27042.

A haplotype analysis based on the CGG tract pat-

tern in grey zone alleles was performed (Table 2). The most common haplotypes in this subgroup were 2-2-G-4 (RF = 0.308) and 7-4-A-5+ (RF = 0.231). All the alleles with 2-2-G-4 haplotype had a long ( $\geq 18$  CGG repeats) uninterrupted sequence at the 3' end ( $p$  = 0.0022). Six alleles out of 18 with other detected haplotypes had the same feature.

## 4. Discussion

Our study is the first study in the Baltic States region regarding *FMR1* linked haplotypes. In the present study, we characterised the microsatellite markers DXS548, FRAXAC1 and FRAXAC2, the ATL1 SNP and the corresponding haplotypes in a mentally retarded male population from Latvia with normal and expanded *FMR1* gene CGG repeats.

Previous studies have suggested linkage of CGG tract instability with the G allele of ATL1, specific microsatellite marker haplotypes and a CGG tract AGG interspersed pattern exhibiting a long uninterrupted CGG repeat at the 3' end [5,11,22-25]. Our results revealed a statistically significant prevalence of the G allele of ATL1 among grey zone alleles and full mutation alleles as an indicator of instability.

The AMOVA data suggested that the differences between detected haplotypes within the control and FXS group were significant. In our Latvian population, 7-4-A-5+ was determined as the prevalent haplotype for nor-

**Table 4.** Frequencies of DXS548-FRAXAC1-ATL1-FRAXAC2 Haplotypes in the Control and FXS Group.

| Haplotype    |         |      |         | Control group |              |              | FXS group |              |              |
|--------------|---------|------|---------|---------------|--------------|--------------|-----------|--------------|--------------|
| DXS548       | FRAXAC1 | ATL1 | FRAXAC2 | n             | RF           | SD           | n         | RF           | SD           |
| 7            | 4       | A    | 5+      | 40            | 0.327        | 0.043        | 0         | -            | -            |
| 7            | 5       | G    | 7       | 11            | 0.090        | 0.026        | 0         | -            | -            |
| 2            | 2       | G    | 4       | 9             | 0.074        | 0.024        | 9         | 0.818***     | 0.122        |
| 7            | 4       | G    | 5       | 8             | 0.066        | 0.023        | 1         | 0.091        | 0.091        |
| 7            | 4       | G    | 6+      | 8             | 0.066        | 0.023        | 0         | -            | -            |
| 6            | 4       | A    | 5+      | 7             | 0.057        | 0.021        | 0         | -            | -            |
| 6            | 5       | G    | 7       | 6             | 0.049        | 0.020        | 0         | -            | -            |
| 6            | 5       | G    | 7+      | 6             | 0.049        | 0.020        | 0         | -            | -            |
| 8            | 4       | A    | 5+      | 3             | 0.025        | 0.014        | 0         | -            | -            |
| 3            | 2       | G    | 4       | 3             | 0.025        | 0.014        | 0         | -            | -            |
| 5            | 4       | A    | 5+      | 2             | 0.016        | 0.012        | 0         | -            | -            |
| 7            | 5       | G    | 7+      | 2             | 0.016        | 0.012        | 0         | -            | -            |
| 7            | 4       | G    | 5+      | 2             | 0.016        | 0.012        | 0         | -            | -            |
| 7            | 4       | A    | 4+      | 2             | 0.016        | 0.012        | 0         | -            | -            |
| 6            | 4       | G    | 5       | 2             | 0.016        | 0.012        | 0         | -            | -            |
| 6            | 5       | A    | 5+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 3            | 4       | G    | 6+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 7            | 4       | A    | 5       | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 7            | 4       | G    | 7       | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 7            | 3       | A    | 5+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 6            | 4       | A    | 4+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 7            | 4       | A    | 3       | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 7            | 4       | A    | 3+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 9            | 4       | A    | 5+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 6            | 5       | G    | 6+      | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 6            | 5       | A    | 7       | 1             | 0.008        | 0.008        | 0         | -            | -            |
| 2            | 4       | G    | 5       | 0             | -            | -            | 1         | 0.091        | 0.091        |
| <b>Total</b> |         |      |         | <b>122</b>    | <b>1.000</b> | <b>0.374</b> | <b>11</b> | <b>1.000</b> | <b>0.304</b> |

*n* = number of chromosomes; RF = relative frequency; SD = standard deviation; \*\*\* *p* < 0.0001

mal CGG alleles. However, after Bonferroni correction, this finding was not considered to be statistically significant. Furthermore, haplotype 2-2-G-4 was found to be in positive association with full mutation CGG alleles and grey zone alleles featuring a long ( $\geq 18$  repeats) uninterrupted CGG tract at the 3' end. These findings imply that, in our population, haplotype 2-2-G-4 is a marker of CGG tract instability. Grey zone alleles with a long uninterrupted CGG tract at the 3' end associated with this haplotype have a higher likelihood of increasing the number of CGG repeats, leading to premutation or mutation over generations.

Several studies have identified specific haplotypes associated with FXS patients and normal CGG repeat alleles across European populations [22,25-29]. However, only a limited number of these studies focused on populations from Eastern and Northeastern Europe. Different loci were used for these haplotype analyses in European populations. Thus, comparison of our results with these analyses would prove difficult. There-

fore, we compared the results presented here with single locus data in the literature.

Allele 7 at the DXS548 locus was the most common allele (RF = 0.639) in our control group. Similar findings, albeit with slightly different frequencies, have been reported from Sweden [27], Czech Republic [28], Finland [30], France [16], Croatia [25], Russia [31], Norway [32] and Poland [29]. In contrast, our finding with respect to the most common allele in FXS patients (allele 2, RF = 0.909) was in agreement with only one of the analyses (Polish population, allele 2; Swedish population, alleles 7 and 6; Czech population, alleles 2, 6 and 7; French population, alleles 2, 6 and 7; Finnish population, almost exclusively allele 6; Croatian population, alleles 7, 6 and 3; Norwegian population, alleles 6 and 2).

For the FRAXAC1 locus, allele 4 was the most common allele (RF = 0.664) in our control group. Analysis of this locus in control chromosomes of a Czech population revealed allele 3 to be the most common allele [28]. This was also the case in populations from Russia [31],

Norway [32] and Croatia [25], while alleles 3 and 4 were prevalent in a Polish population [29]. Regarding Latvian FXS chromosomes, we found allele 2 to be the most common allele (RF = 0.818). However, this was not in line with the findings from other populations (Norwegian, allele 4; Polish, allele 4; Croatian, allele 3; Czech, alleles 3 and 4).

Compared to the microsatellite marker FRAXAC1, the FRAXAC2 locus has been more widely used in studies. Allele 5+ was the most common allele at this locus in our control group, detected at a RF of 0.467. This finding was not replicated in other populations: Finnish, allele 3 [30]; Swedish, Czech and French, allele 4+ [16,26,28]; Polish, alleles 7 and 7+ [29]. The prevalent allele in our FXS patients was allele 4. Allele 3 was identified in half of the investigated chromosomes in a Finnish FXS patient group, while alleles 4+ and 5 were prevalent in a Swedish population, alleles 4+ and 4 in a Czech population, alleles 3 and 4 in a French population, and allele 7 in a Polish population.

In summary, our finding regarding allele 7 being the most common allele at the DXS548 locus in Latvian control patients is in line with several other European population control groups. Furthermore, our FXS patient group finding (i.e. allele 2 being the prevalent allele at this locus) is consistent with data from a Polish FXS population. The FRAXAC1 and FRAXAC2 loci results for our control and FXS group differ to varying degrees from the data reported for other European populations.

As the analysed microsatellite loci and nomenclature assigned to alleles in the literature are different, confusion arises, which may lead to bias in

the interpretation of literature data comparing haplotypic results from different populations.

To the best of our knowledge, specific *FMR1*-linked haplotypes in the Baltic State region and Northeastern Europe have not been previously described. The present study is the first to report Latvian population *FMR1* haplotype data. Comparison of the data with those obtained from geographically close European populations highlights differences, particularly with the FXS patient group. Indeed, haplotype 2-2-G-4 appears to be exclusively found in Latvian FXS chromosomes. A small sample size making up the case group may have a negative effect on this result.

We conclude that a founder effect could not be an explanation of our findings on the basis of heterogeneity exhibited by the Latvian population and lack of studies across this geographical region. A larger study of *FMR1*-linked haplotypes in Eastern and Northeastern European regions may provide evidence of a specific mutational pathway for unstable CGG alleles in this geographical region.

### Acknowledgements

The study was supported by Riga Stradins University ESF projects No. 2004/0005/VPD1/ESF/PIAA/04/NP/3.2.3.1./0001/0004/0066 and No. 2009/0147/1DP/1.1.2.1.2/09/IPIA/VIAA/009. We are grateful to Dr. K. Eiklid, Ullevål University Hospital, Oslo, Norway, and Prof. R. A. Wevers and Dr. H. Yntema, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands for technical help and inspiration for this project.

### References

- [1] Crawford, D. C., Acuna, J. M. and Sherman, S. L. *FMR1* and the fragile X syndrome: Human genome epidemiology review. *Genetics in Medicine*, 2001, 3(5), 359-371
- [2] Oberle I., Rousseau F., Heitz D., Devys D., Hanauer A., Boue J., *et al.*, Instability of a 550-Base Pair DNA Segment and Abnormal Methylation in Fragile X Syndrome. *Science*, 1991, 252, 1097-1102
- [3] Poustka A., Dietrich A., Langenstein G., Toniolo D., Warren S. T., and Lehrach H., Physical map of human Xq27-qter: Localizing the region of the fragile X mutation. *Proc. Natl. Acad. Sci. USA*, 1991, 88, 8302-8306
- [4] Rousseau F., Heitz D., Oberle I., and Mandel J-L., Selection in blood cells from female carriers of the fragile X syndrome: inverse correlation between age and proportion of active X chromosomes carrying the full mutation. *J. Med. Genet.*, 1991, 28, 830-836
- [5] Eichler E. E., Macpherson J. N., Murray A., Jacobs P. A., Chakravarti A., and Nelson D. L., Haplotype and interspersed analysis of the *FMR1* CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. *Hum. Mol. Genet.*, 1996, 5, 319-330
- [6] Fu Y. H., Kuhl D. P. A., Pizzuti A., Pieretti M., Sutcliffe J. S., Richards S., *et al.*, Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. *Cell*, 1991, 67, 1047-1058
- [7] de Vries B. B. A., Halley D. J. J., Oostra B. A., and Niermeijer M. F., The fragile X syndrome. *J. Med. Genet.*, 1998, 35, 579-589
- [8] Moutou C., Vincent M. C., Biancalana V., and Mandel J-L., Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic. 1997, *Hum. Mol. Genet.*, 3, 971-979

- [9] Rife M., Badenas C., Quinto L.I., Puigoriol E., Tazon B., Rodriguez-Revenge L., *et al.*, Analysis of CGG variation through 642 meioses in Fragile X families. *Mol. Hum. Reprod.*, 2004, 10, 773-779
- [10] Sherman S., Plecher B. A., and Driscoll D. A., Fragile X syndrome: Diagnostic and carrier testing. *Genetics in Medicine*, 2005, 7, 584-587
- [11] Dombrowski C., Levesque S., Morel M. L., Rouillard P., Morgan K., and Rousseau F., Premutation and intermediate-size FMR1 alleles in 10 572 males from the general population: loss of an AGG interruption is a late event in the generation of fragile X syndrome alleles. *Hum. Mol. Genet.*, 2002, 11, 371-378
- [12] Gunter C., Paradee W., Crawford D. C., Meadows K. A., Newman J., Kunst C. B., *et al.*, Re-examination of factors associated with expansion of CGG repeats using a single nucleotide polymorphism in FMR1. *Hum. Mol. Genet.*, 1998, 7, 1935-1946
- [13] Kunst C. B., Zerynick C., Karickhoff L., Eichler E., Bullard J., Chalifoux M., *et al.*, FMR1 in Global Populations. *Am. J. Hum. Genet.*, 1996, 58, 513-522
- [14] Macpherson J. N., Bullman H., Youngs S. A., and Jacobs P. A., Insert size and flanking haplotype in fragile X and normal population: possible multiple origins for the fragile X mutations. *Hum. Mol. Genet.*, 1994, 3, 399-405
- [15] Murray A., Macpherson J. N., Pound M. C., Sharrock A., Youngs S. A., Dennis N. R., *et al.*, The role of size, sequence and haplotype in the stability of FRAXA and FRAXE alleles during transmission. *Hum. Mol. Genet.*, 1997, 6, 173-184
- [16] Oudet C., Mornet E., Serre J. L., Thomas F., Lentes-Zengerling S., Kretz C., *et al.*, Linkage disequilibrium between the fragile X mutation two closely linked CA repeats suggests that fragile X chromosomes are derived from a small number of founder chromosomes. *Am. J. Hum. Genet.*, 1993, 52, 297-304
- [17] Richards R. I., Holman K., Kozman H., Kremer E., Pritchard M., Yu S., *et al.*, Fragile X syndrome: genetic localisation by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. *J. Hum. Genet.*, 1991, 28, 818-823
- [18] Chong S. S., Eichler E. E., Nelson D. L., and Hughes M. R., Robust amplification and ethidium – visible detection of the fragile X syndrome CGG repeat using Pfu polymerase. *Am. J. Med. Genet.*, 1994, 51, 522-526
- [19] Crawford D. C., Schwartz C. E., Meadows K. L., Newman J. L., Taft L. F., Gunter C., *et al.*, 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. Med. Genet.*, 2000, 66, 480-493
- [20] Chiurazzi P., DNA panel to interlaboratory standardization of haplotype studies on the fragile X syndrome and proposal for a new allele nomenclature. *Am. J. Med. Genet.*, 1999, 83, 347-349
- [21] Excoffier L. and Lischer H.E. L., Arlequin suite ver. 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources.*, 2010, 10, 564-567
- [22] Arrieta I., Penagarikano O., Telez M., Ortega B., Flores P., Criado B., *et al.*, The FMR1 CGG repeat and linked microsatellite markers in two Basque valleys, *Heredity*, 2003, 90, 206-211
- [23] Curlis Y., Zhang C., Holden J. J. A., Kirkby K., Loesch D., and Mitchell R. J., Haplotype study of intermediate-length alleles at the fragile X (FMR1) gene: ATL1, FMRb, and microsatellite haplotypes differ from those found in common-size FMR1 alleles. *Hum. Biol.*, 2005, 77, 137-151
- [24] Zhou Y., Tang K., Law H. Y., Ng I. S. L., Lee C. G. L., and Chong S. S., FMR1 CGG repeat patterns and flanking haplotypes in three Asian populations and their relationship with repeat instability. *Ann. Hum. Genet.*, 2006, 70, 784-796
- [25] Dokic H., Barišić I., Čulić V., Lozić B., and Hecimovic S., Haplotype and AGG interspersed analysis of FMR1 alleles in a Croatian population: no founder effect detected in patients with fragile X syndrome. *Hum. Biol.*, 2008, 80, 581-587
- [26] Malmgren H., Gustavson K. H., Oudet C., *et al.*, Strong founder effect for the fragile X syndrome in Sweden. *Eur. J. Hum. Genet.*, 1994, 2, 103-109
- [27] Peixoto A., dos Santos M. R., Seruca R., Amorim A., and Castedo S., Analysis of FMR1 and flanking microsatellite markers in normal and fragile X chromosomes in Portugal: evidence for a “protector” haplotype. *Eur. J. Hum. Genet.*, 1998, 6, 518-522.
- [28] Pekarik V., Blazkova M., and Kozak L., Haplotype analysis of the fragile X syndrome gene FMR1 in the Czech Republic. *Am. J. Med. Genet.*, 1999, 84, 214-21

- 
- [29] Rajkiewicz M., Molecular analysis of the *FMR1* gene and searching for the premutations in the large group of Polish ataxia patients and group of women with premature ovarian failure. PhD thesis, Department of Genetics, Institute of Psychiatry and Neurology, Warsaw, Poland, 2008
- [30] Haataja R., Vaisanen M. L., Li M., Muiyao L., Ryyanen M., and Leisti J., The fragile X syndrome in Finland: demonstration of a founder effect by analysis of microsatellite haplotypes. *Hum. Genet.*, 1994, 94, 479-483
- [31] Drozd O. V., Strelnikov V. V., Babenko O. V., Zemlyakova V.V., Nemtsova M. V., and Zaletaev D. V., Allelic polymorphism of the five X-linked (CA)<sub>n</sub> dinucleotide repeats in Russia. *Genetica.*, 2003, 39, 433-438
- [32] Larsen L. A., Vuust J., Nystad M., Evseeva I., Van Ghelue M., and Tranebjaerg L., Analysis of *FMR1* (CGG)<sub>n</sub> alleles and DXS548-FRAXAC1 haplotypes in three European circumpolar populations: traces of genetic relationship with Asia. *Eur. J. Hum. Genet.*, 2001, 9, 724-727