

Zanda Daneberga

**THE FRAGILE X SYNDROME IN  
MENTALLY RETARDED  
PATIENTS FROM LATVIA**

Doctoral Thesis

Summary

Speciality: Medical Genetics

Rīga, 2011

Elaboration of the Study: Medical Genetics Clinic, University Children`s Hospital, Riga, Latvia

Scientific supervisor: *dr. med.* assoc. prof. **Rita Lugovska**, Medical Genetics Clinic, University Children`s Hospital, Riga, Latvia; Riga Stradins University,

Scientific advisor: *dr. med.* **Baiba Lāce**, Medical Genetics Clinic, University Children`s Hospital, Riga, Latvia; Lavian BioMedical Research and Study Center, Riga, Latvia

Official reviewers:

*dr. biol.*, assoc. prof. **Edvīns Miklaševičs**, Riga Stradins University

*dr. habil. biol.*, prof. **Nikolajs Sjakste**, University of Latvia, Faculty of Medicine

*dr. med.*, assoc. prof. **Jurate Kasnauskiene**, Vilnius University, Faculty of Medicine

Defence of the doctoral thesis will be held on 21<sup>st</sup> of December, 2011 at 14.00 o`clock in the Hippocratic lecture-hall of Riga Stradins University (RSU) Dzirciema Str.16.

The Doctoral Thesis is available in the library of RSU and at [www.rsu.lv](http://www.rsu.lv)

The work was supported by ESF Project “Support to doctor`s studies and acquiring an academic degree in Riga Stradins University” No. 2009/ 0147/ 1DP/ 1.1.2.1.2/ 09/ IPIA/ VIAA/ 009.



Secretary of Promotion Council:

*Dr. habil.med.*, prof. **Līga Aberberga - Augškalne**

## TABLE OF CONTENTS

1.	INTRODUCTION.....	6
1.1.	Aim of the Study.....	8
1.2.	Tasks of the Study.....	8
1.3.	Scientific Novelty of the Study.....	9
1.4.	Practical Novelty of the Study.....	9
1.5.	Author`s Contribution to the Work.....	10
1.6.	Elaboration of the Study.....	10
1.7.	Outline of the Thesis.....	11
2.	SUBJECTS AND METHODS.....	11
2.1.	Subjects.....	11
2.1.1.	Prevalence of the Fragile X Syndrome.....	11
2.1.2.	Variation of CGG Trinucleotide Repeats.....	12
2.1.3.	The Case-Control Study.....	12
2.1.4.	Genotype-Phenotype Correlation.....	13
2.2.	Molecular Studies.....	13
3.2.1.	DNA Extraction.....	13
3.2.2.	Routine Screening PCR Amplification.....	14
3.2.3.	Fluorescent PCR.....	14
3.2.4.	Southern Blotting.....	15
3.2.5.	AGG Interspersion Pattern Analysis.....	15
3.2.6.	Single Nucleotide Polymorphism Analysis.....	16
3.2.7.	Haplotype Analysis.....	17
3.2.8.	Statistical Data Analysis.....	19
3.	RESULTS.....	20
3.1.	Prevalence of the Fragile X Syndrome.....	20
3.2.	Variation of CGG Trinucleotide Repeats.....	21

3.3.	ATL1 SNP.....	22
3.4.	Repeat Structures of Grey-Zone Alleles.....	24
3.5.	DXS548-FRAXAC1-ATL1-FRAXAC2 Haplotypes.....	25
3.6.	Analysis of Molecular Variance.....	28
3.7.	Genotype-Phenotype Correlation.....	29
4.	DISCUSSION.....	31
4.1.	Prevalence of the Fragile X Syndrome.....	31
4.2.	Variation of CGG Trinucleotide Repeats.....	32
4.3.	ATL1 SNP.....	33
4.4.	Repeat Structures of Grey-Zone Alleles.....	33
4.5.	DXS548-FRAXAC1-ATL1-FRAXAC2 Haplotypes.....	34
4.6.	Genotype-Phenotype Correlation.....	37
5.	CONCLUSIONS.....	39
6.	PUBLICATIONS.....	40
7.	ACKNOWLEDGMENTS.....	42
8.	REFERENCES.....	44

## **ABBREVIATION**

AMOVA - Analysis of molecular variance

ARI - Acute respiratory infection

bp - Base pairs

CGG - Cytosine – guanine – guanine

CNS - Central nervous system

DNA - Deoxyribonucleic acid

EDTA - Ethylenediaminetetraacetic acid

FXS - Fragile X Syndrome

MR – Mental retardation

IQ - Intelligence Quotient

PCR - Polymerase Chain Reaction

PWS - Prader-Willy syndrome

RF – Relative frequency

SNP - Single-Nucleotide Polymorphism

STR – Short tandem repeat

# 1. INTRADUCTION

Mental retardation (MR) is a complex phenotype, characterized by suboptimal functioning of the central nervous system (CNS) resulting in significant limitations both in intellectual functioning and in adaptive behaviour. Mental retardation affects about 2–3% of people and about a quarter of cases are caused by genetic disorders. Mental retardation is the most frequent cause of severe handicap in children. Therefore ascertainment of mental retardation aetiology is an important task in paediatrics.

Fragile X syndrome (FXS; MIM #300624; FRAXA, Xq27.3) is well known and a common cause of X-linked mental retardation. The fragile X syndrome is caused by an expanded CGG repeat (> 200 units, full mutation) at the 5' end of the *FMRI* gene, which is associated with methylation of a CpG island upstream of the *FMRI* gene and down regulation of the transcription (Oberle et al., 1991; Poustka et al., 1991; Rousseau et al., 1991).

Amongst 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 (Eichler et al., 1996; Fu et al., 1991). 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 *FMRI* expression. CGG repeat expansion over 200 is the basis for CpG island methylation, leading to silencing of the *FMRI* gene (de Vries et al., 1998). 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 (Moutou et al., 1997; Rife et al., 2004; Sherman et al., 2005). These alleles usually have stable transmission, 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 three main factors related to this instability (Dombrowski et al., 2002; Eichler et al., 1996; Rife et al., 2004). The microsatellite markers DXS548-FRAXAC1-FRAXAC2 and the ATL1 SNP have previously been reported as markers associated with *FMRI* CGG repeat instability (Eichler et al., 1996; Gunter et al., 1998; Kunst et al., 1996; Macpherson et al., 1994; Murray et al., 1997; Oudet et al., 1993; Richards et al., 1991).

Haplotypes linked to FXS are widely described across Western European and Scandinavian populations; however, less is known regarding populations from North-Eastern Europe, including the Baltic States. This is the first study in the Baltic States region regarding *FMRI* linked haplotypes.

The first clinical indication of FXS is usually delay in child's developmental milestones and mental retardation. In addition to mental retardation, speech and language skills are severely affected. Most speech is poorly articulated and expressive language is often limited to three- or four-word sentences. FXS patients often repeat words or phrases, an attribute typically associated with autism. Indeed many FXS males present autistic type behaviour – gaze aversion, shyness, hand biting, hand flapping and rocking (Bardoni et al., 2000; Garber et al., 2008; Hernandez et al., 2009).

The phenotype is subtle in young children and evolves with age. Hyperextensibility of finger joints, *pectus excavatum*, mitral valve prolapse and strabismus are other possible prevalent features (de Vries et al., 1996; Larbrisseau et al., 1982; Phadke 2005). The clinical manifestations of this syndrome in adult males include an elongated and narrow face with a large forehead and prominent chin, large and anteverted ears, joints with increased mobility, and uni- or bilaterally large testes. Macroorchidism is an important feature in the post-pubertal age. However, it is not presented in all FXS males,

but it is specific for FXS. Between 25-30% of all patients with FXS do not have the typical faces of the syndrome. The secondary characteristics of FXS in turn include tallness, a soft and silky skin, widened fingertips and flat feet (Ridaura-Ruiz et al., 2009).

Early diagnosis of fragile X syndrome is crucial in order to inform other members of the family of their risk of having affected offspring. Therefore it is recommended that most fragile X diagnostic tests will be carried out on a very broad range of patients regardless of a consequently low detection rate.

Ten years of experience with molecular diagnostic of the fragile X syndrome in Latvia and number of diagnosed patients in this time period, revealed a low pickup rate of patients and insufficient clinical recognition of symptoms.

### **1.1.Aim of the Study**

Ascertain the prevalence of fragile X syndrome in Latvia, characterise genetic and clinical variability of the FRAXA locus *FMRI* gene in patients with unclear aetiology of mental retardation.

### **1.2.Tasks of the Study**

1. Estimate the prevalence of the fragile X syndrome in the entire Latvian male population.
2. Perform a distribution and structure study of CGG repeats among X chromosomes with normal CGG repeat alleles.
3. Characterise the ATL1 SNP/CGG repeat number correlation within chromosomes with a normal CGG repeat number and chromosomes with a full mutation.
4. Perform a case-control study of *FMRI* gene linked haplotypes based on STR and SNP markers, to identify specific haplotypes among Latvian



FXS patients and control group mentally retarded patients with a normal number of CGG repeats with respect to allelic stability.

5. Identify the association of grey-zone allele structure and *FMRI* linked haplotypes.
6. Evaluate genotype-phenotype correlation in patients with full mutation and/or repeat size/methylation mosaic.

### **1.3.Scientific Novelty of the Study**

This study is the first study in the Baltic States region regarding *FMRI* linked haplotypes. Described haplotypes of Latvian fragile X syndrome patients differ from published studies in populations of Western European descent. Therefore this data provide evidence of different mutational pathways of CGG repeat expansion in the North-Eastern European region.

The estimated prevalence of fragile X syndrome in the Latvian male population contributes to the ascertainment of this disease distribution in our geographical region.

### **1.4.Practical Novelty of the Study**

The estimated prevalence of fragile X syndrome in the Latvian male population is in line with the prevalence of this syndrome in several other European populations. The low number of confirmed patients with fragile X syndrome in ten years, point to a low detection rate of patients in paediatrician, child psychiatric and child neurology practices.

Haplotypes linked to unstable CGG repeat alleles in the Latvian FXS male population are very useful in practical family cascade testing and for consultation of families at risk.

The newly adapted clinical questionnaire form shall contribute to an increase in the detection rate of patients with suspected fragile X syndrome by paediatricians, child psychiatric and child neurology practices.

### **1.5. Author's Contribution to the Work**

This PhD project was initiated in 2005, based on scientific elaboration forerun of project "Genomic studies of the Latvian population, their application for diagnosis and prevention of human pathology".

The author of this thesis performed the following laboratory investigations: DNA extraction (partly); routine screening PCR amplification; fluorescent PCR; Southern blotting (partly); AT1 SNP analysis; AGG interspersion pattern analysis; fluorescent PCR of microsatellite markers and haplotype analysis. Author performed retrospective data collection and data analysis for prevalence study. All statistical data analysis and AMOVA were done by the author of this thesis.

Clinical evaluation of patients was done by clinical geneticists and child psychiatrists.

### **1.6. Elaboration of the Study**

The current study was carried out in the Medical Genetics Clinic, University Children's Hospital, Riga, Latvia in collaboration with Children's Psychiatric Department, University Children's Hospital, Riga, Latvia.

Confirmation of FXS diagnosis by Southern blot analysis was done in the DNA Laboratory, Department of Medical Genetics, Ullevål University Hospital, Oslo, Norway and in the DNA Diagnostic Laboratory, University Medical Center Nijmegen, The Netherlands.

The Latvian Central Committee of Medical Ethics and the Riga Stradins University Committee of Medical Ethics approved the study.

## **1.7. Outline of the Thesis**

The thesis is composed on 124 pages in English, following classical scheme. The work is structured in ten chapters: Introduction; Literature review; Subjects and Methods; Results; Discussion; Conclusions; Publications; Acknowledgements; References and Appendixes. Text of thesis is supplemented by 19 Tables; 19 Figures and 14 Appendixes. Reference list consist of 131 cited references.

## **2. SUBJECTS AND METHODS**

### **2.1. Subjects**

#### **2.1.1. Prevalence of the Fragile X Syndrome**

The retrospective data of patients genotypes, analyzed in the Medical Genetic Clinic, University Children`s Hospital between 1998 and 2007, were summarized to assess the prevalence of FXS.

All patients were referred for exclusion/confirmation of fragile X syndrome by clinical geneticist at the Medical Genetic Clinic, University Children`s Hospital, by child psychiatrist for the hospitalized persons at the Children`s Psychiatric Department, University Children`s Hospital and by clinical geneticist at the children`s attending Social Care Centre Riga, Latvia.

Inclusion criteria for selecting patients' data were as follows:

- patients with mental retardation in various degrees with or without association with dysmorphic features
- MR patients with autism, autistic spectrum disorders and any type of behavioural disturbances
- genotype data with exact number of CGG repeats

Exclusion criteria for selecting patients' data were as follows:

- patient gender (female)

- consanguinity
- monogenic, chromosomal and metabolic diseases

The clinical features of the patients were assessed and family history obtained by clinical geneticist. The ethnical background of patients was not considered.

Based on inclusion/exclusion criteria 374 anonymous, unrelated male patients data were selected for the prevalence study. The age of patients at the time of the DNA diagnostic test varied between two and seventeen years.

### **2.1.2. Variation of CGG Trinucleotide Repeats**

To assess distribution of normal CGG repeat alleles retrospective data of patients genotypes, analyzed in the Medical Genetic Clinic, University Children's Hospital between 1998 and 2007, were used. Based on inclusion/exclusion criteria, 374 anonymous, unrelated male patient data were used. We considered selected data comparable, because for all 374 samples, both routine screening with PCR and fluorescent PCR following Applied Biosystems protocol for exact CGG repeat number detection, were performed.

### **2.1.3. The Case-Control Study**

For case-control study of *FMRI* linked haplotypes the control group of 122 unrelated male patients with normal number of CGG repeats were selected based on inclusion/exclusion criteria.

Inclusion criteria for selecting control group were as follows:

- parents or legal representatives of minors signed informed consent according regulations issued by Ethics Committee for participation in this study
- genotype data within a normal range of CGG repeats
- patients with mental retardation in various degrees with or without association with dysmorphic features

- MR patients with autism, autistic spectrum disorders and any type of behavioural disturbances

Exclusion criteria for selecting control group were as follows:

- patient gender (female)
- consanguinity
- monogenic, chromosomal and metabolic diseases

The case group consisted of 11 unrelated male patients with confirmed diagnosis (full mutation). Parents or legal representatives of minors signed informed consent according to regulations issued by Ethics Committee for participation in this study.

#### **2.1.4. Genotype-Phenotype Correlation**

Genotype-phenotype correlation was assessed for 12 male patients with confirmed diagnosis of FXS in time period from 1998 to 2010. In this group of study siblings were included. The age of patients at the moment of diagnosis varied between two and sixteen years (average =  $7.33 \pm 4.46$ ). Clinical information was obtained from case-records of patients by clinical geneticist or child psychiatrist. Anthropometric data were measured according to the “Smith recognizable patterns” and a methodology described by Krūmiņa, Kokare and Biķis (2007). IQ tests were performed based on the Woodcock – Johnson test and Wechsler Intelligence Scale for Children. Autistic spectrum disorders were evaluated according to the Autism Diagnostic Observation Schedule (ADOS).

## **2.2.Molecular Studies**

### **2.2.1. DNA Extraction**

Five millilitres of peripheral blood were collected in EDTA-coated tubes. Before DNA extraction, blood samples were kept frozen at -20°C. DNA

was extracted using “Genomic DNA Purification Kit” (Fermentas, Lithuania) according to the manufacturer protocol.

### **2.2.2. Routine screening PCR amplification**

For the amplification of normal CGG repeat allele within *FMRI* gene, primers sequence corresponding to position 212-241 and 599-571 at the 1kb *PstI* fragment plasmid Ep5.1, containing CpG island and CGG repeat region, were used. The reaction was performed according to the protocol described by Chong et al. (1994).

### **2.2.3. Fluorescent PCR**

For a precise determination of CGG repeat number, fluorescent PCR was carried out according to the Applied Biosystems (USA) protocol. The reagent solutions were in-house prepared. Approximately 100 diploid copies (0.67 ng) per  $\mu\text{l}$  of genomic DNA were used to carry out the reaction. This reaction was performed using the “Hot start” technique with Ampi Wax (Applied Biosystems, USA).

PCR products were separated on an ABI Prism<sup>®</sup> 310 genetic analyzer (Applied Biosystems, USA) under two different electrophoresis conditions.

Genotyping results were analysed by GeneScan<sup>™</sup> software (Applied Biosystems, USA). The corresponding peaks length was calculated according to the calibration curve of the fragile X size standard (50 bp-2500 bp). Apolipoprotein E (*ApoE*) gene CG rich fragment was used as an internal control for amplification. Presence of *Apo E* fragment indicates successful amplification and template quality.

#### **2.2.4. Southern blotting**

The FXS diagnosis was confirmed by sizing of the repeat array using methylation specific restriction enzyme digestion and genomic Southern blot hybridization according to the described protocol (Dracopoli and Haines, 1994).

For Southern blot analysis, 4 - 6 µg of genomic DNA were used.

To perform sizing of the repeat array, two different digestion reactions were used – EagI/EcoRI (methylation specific) and PstI. Products were separated on 0.8% agarose gel (at ~ 0.35 V/cm, overnight) and transferred to positive charged Nylon membrane by capillary transfer. To detect DNA fragment, labelled [<sup>32</sup>P] StB12.3 hybridisation probe was used. Unlabelled StB12.3 probe obtained from Prof. J. L. Mandel, Strasbourg. Labelling of the probe was done according to the described protocol (Sambrook and Russell, 2001). Analysis performed in the DNA Laboratory, Department of Medical Genetics, Ullevål University Hospital, Oslo, Norway.

Southern blot analysis for patients and their family members using probe pAO365 was performed in the DNA Diagnostic Laboratory, University Medical Center Nijmegen, The Netherlands.

#### **2.2.5. AGG interspersed pattern analysis**

Twenty-six alleles of grey zone (35-50 CGG repeat) were analysed for CGG repeat patterns. The AGG interspersed pattern was determined by sequencing of the CGG-repeat array. In brief, the CGG repeat and surrounding DNA sequences were amplified from genomic DNA by *Pfu* polymerase (Fermentas, Lithuania) with the PCR protocol described previously (Chong et al. 1994). The PCR products were run on 2.5% agarose gel at 5.5 V/cm for 60 min to check for amplification of a single allele. The PCR products were

concentrated and purified for sequencing by the Montage PCR centrifugal filter device (Millipore, USA).

The sequencing reaction was performed using concentrated and purified PCR products and the BigDye<sup>®</sup> Terminator v3.1 kit (Applied Biosystems, USA) according to the manufacturer's protocol.

All sequencing reactions were run on an ABI Prism<sup>®</sup> 310 genetic analyzer using 61cm x 50 $\mu$ m (50 cm well-to-read) capillary with POP-6<sup>™</sup> polymer and analyzed by ABI DNA<sup>™</sup> sequencing software.

The sequence pattern of the CGG repeat array was read from the first exon of *FMRI* gene. DNA sequence was established by visual interpretation of the electropherograms. Nucleotides were assigned in the sequence based on the highest fluorescence signal at each position, provided that the nucleotide peak exceeded the background level. To describe the sequence of the *FMRI*, the number of CGG repeats is denoted as a number, while AGG interruption is denoted as “+”.

#### **2.2.6. Single nucleotide polymorphism analysis**

The ATL1 polymorphism (alleles A/G located 5613bp upstream CGG repeat) was analysed by the allele-specific oligonucleotide PCR protocol, described by Dombrowski et al. (2002). PCR products were visualised on 1% agarose gel using ethidium bromide staining. Presence of PCR product in length of 385 bp was interpreted as a positive result for the specific allele. A lack of PCR product was interpreted as a negative result for a specific allele. The ATL1 SNP was identified by performing two PCR reactions for each chromosome.



### 2.2.7. Haplotype Analysis

For haplotype analysis among normal and mutant chromosomes, the microsatellite markers DXS548, FRAXAC1 and FRAXAC2 were used. The DXS548 microsatellite is located 189895bp downstream CGG repeat. The FRAXAC1 microsatellite is located 7221bp downstream CGG repeat and the FRAXC2 microsatellite is located 12418bp upstream of the CGG repeat.

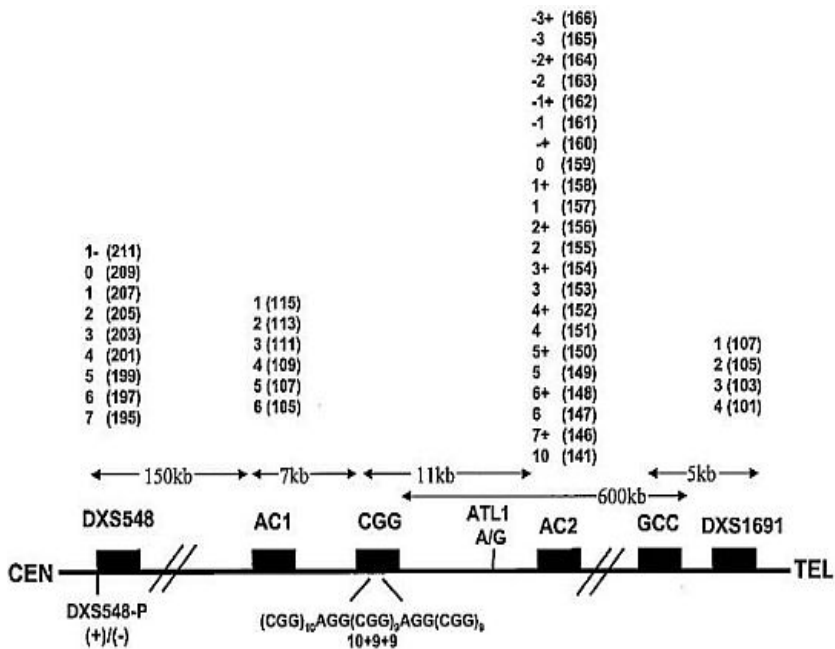
Multiplex fluorescent PCR for DXS548 and FRAXAC2 was performed using primers sequences described by Chiurazzi et al. (1999). The FRAXAC1 microsatellite marker was amplified separately under same reaction conditions. Primers sequences described by Chiurazzi et al. (1999) were used.

Genotyping results were analysed by GeneScan<sup>®</sup> Analysis software. The corresponding fragment length was calculated according to the calibration curve of the GeneScan<sup>™</sup> ROX 500<sup>™</sup> size standard. Nomenclature for alleles was adjusted to the nomenclature recommended by Macpherson et al. (1994) and Eichler et al. (1996) (Fig. 2.1.).

Validation of genotyping results was made by direct sequencing of random alleles for each microsatellite marker. For alleles of each marker the same PCR conditions, as described above for genotyping, was used. The exception was reverse primers A, which was not labelled with fluorescent dye. The PCR products obtained in reaction were concentrated and purified for sequencing reaction by the Millipore Montage PCR filter device.

The sequencing reaction was performed using concentrated and purified PCR products and the BigDye<sup>®</sup> Terminator v3.1 kit (Applied Biosystems, USA) according to the manufacturer's protocol.

All sequencing reactions were run on an ABI Prism<sup>®</sup> 310 genetic analyzer using 61cm x 50 $\mu$ m (50 cm well-to-read) capillary with POP-6<sup>™</sup> polymer and analyzed by ABI DNA<sup>™</sup> sequencing software.



**Fig. 2.1 .The location and nomenclature of CGG repeat pattern, STR- and SNP-based haplotypes according to Eichler et al. (1996), adapted from Crawford et al. (2000).**

Deviation of fragment length described by Macpherson et al. (1994) was detected by sequencing results. Nomenclature was set based to the repeat number in microsatellite markers sequence. Deviation in fragment length was found for FRAXAC1 and DXS548 loci. FRAXAC2 locus alleles corresponded to the previously described fragment length in bp. For the DXS548 locus fragments in length of 193 bp and 192 bp were denoted as allele 6, because both fragments contain 20 CA repeats and the 1bp difference rose from one extra G in forward amplified sequence.

The haplotype of each chromosome was formed by combining microsatellite markers and single nucleotide polymorphism according to their

position at the FRAXA locus. STR and SNP marker haplotypes were combined as follows: DXS548-FRAXAC1-ATL1-FRAXAC2.

### 2.3. Statistical Data Analysis

For AMOVA analysis, level of heterozygosity for all polymorphisms and calculation of pairwise genetic distances (*Fst*) the Arlequin 3.5 package (Excoffier and Lischer, 2010) was used. The significance of the haplotype results was tested by 10000 permutations. To detect critical value of *Fst*, the on-line statistical calculator for critical values of F-statistics by BioKin, Ltd. was used. The “degrees of freedom numerator” was set to 1. The “denominator” was set to 132.

Analysis of 27 haplotypes, derives from a total 133 chromosomes, was carried out by population splitting in two subgroups based on normal/mutated FRAXA alleles.

The case-control study data was analysed by Fisher’s exact test of 2×2 contingency tables and chi-square using GraphPad QuickCalcs on-line calculator. For statistical significance of results, the *p*-value was set less than 0.05. The Bonferroni correction was applied for multiple testing (Bland and Altman, 1995). According to the Bonferroni correction in four markers, the haplotype analysis *p*-value was set less than 0.0125.

For calculation of the 95% confidence interval the QuickCalcs on-line calculator was used. For the proportion calculation CI 95% was applied according to the modified Wald method by Agresti and Coull (1998).

In the retrospect study, estimation of FSX prevalence was based on data from the Central Statistical Bureau of Latvia and data obtained from this study. Prevalence per 100 000 males was calculated:

$$\text{Prevalence} = \left[ \frac{\text{number of estimated male patients during the period}}{\text{average number of male population during the period}} \right] \times 100\,000$$

Prevalence expressed as one affected male to the number of male persons in population was calculated:

$$\text{Prevalence} = \frac{\text{average number of male population during the period}}{\text{number of estimated male patients during the period}}$$

### 3. RESULTS

#### 3.1. Prevalence of the Fragile X Syndrome

The prevalence of the fragile X syndrome in male individuals with mental retardation and developmental disability was estimated retrospectively. In this study the estimation of the population prevalence was restricted to the data from male population, because females with a full mutation in the *FMR1* gene show an intellectual development from severely retarded to normal and cannot be identified solely by clinical data.

In the group of unrelated, mentally retarded males (n = 374), 10 (95% CI 4.80 – 18.39) in ten years time period newly diagnosed fragile X syndrome patients had a relative prevalence of 0.0267 (2.67%). According to the data from the Central Statistical Bureau of Latvia, during the ten years time period 1998 to 2007, 10503 patients with psychological development disorders or behavioural and emotional disorders were diagnosed (with onset usually occurring in childhood and adolescence, new cases excluded those caused by alcoholism and dependency upon narcotic and psychoactive substances). Gender structure for the diagnosed cases was not available. Based on a theoretical gender structure of the population (1:1) and a developmental disability diagnosis of 1.25 male to 1 female (Raymond, 2006), we estimated that there were 6295 male patients (95% CI 5690-7430) with diagnosis of developmental disability in Latvia. According to the calculated relative

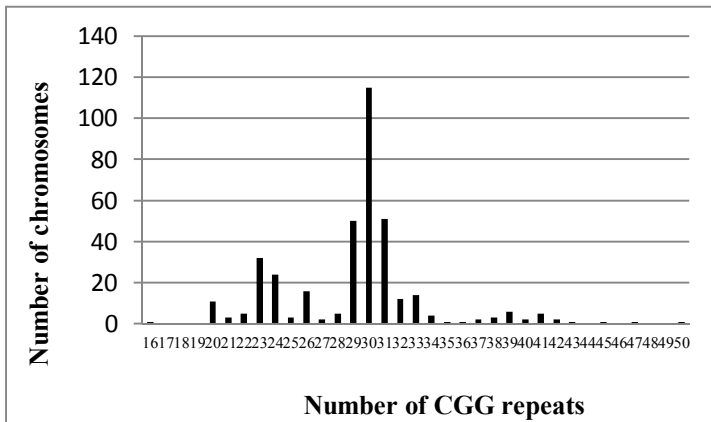
prevalence of disease from our laboratory data, we further estimated that this population included 168 (95% CI 143 – 195) FXS male patients.

For Latvia, with an average 1 079 941 male residents (based on data of the Central Statistical Bureau of Latvia for 1998 - 2007), and assumed 168 male patients with the fragile X syndrome, gives a result in prevalence of 1/6428 males (95% CI 5538-7552) or 15.55/100 000 males (95% CI 13.24 – 18.05).

### **3.2. Variation of CGG Trinucleotide Repeats**

In total, 374 patients were analysed with PCR screening and for all those patients an exact CGG repeat number within *FMRI* gene were detected. Distribution of alleles was as follow: 90.37% of alleles fell in group of normal CGG repeat number, 6.95% were grey-zone alleles and 2.67% of alleles revealed full CGG repeat expansion. The highest incidence among all analysed chromosomes were observed for allele 30 (29.95%), allele 31 (13.10%) and allele 29 (12.83%).

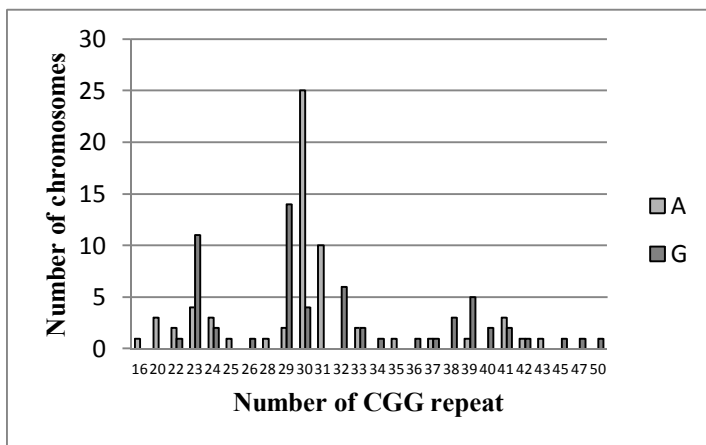
From 374 patients, analysed with PCR screening, 364 were detected having CGG repeat alleles within a non-pathogenic range (5 – 50 repeats). Twenty-six different alleles were observed. The smallest repeat size identified within the normal range was 16 CGG repeats. There were absences of alleles with 17; 18; 19; 44; 46; 48 and 49 CGG repeats. The most common allele in the normal range was allele 30 (30.77%). Comparably prevalent were alleles 29 (13.19%), 31 (13.46%), 23 (8.52%) and 24 (6.32%). Distribution of non-pathogenic range CGG repeat alleles is shown in Fig. 3.1.



**Fig. 3.1. Distribution of CGG alleles within non-pathogenic range.**

### 3.3. ATL1 SNP

The control group CGG repeat alleles (n = 122) were analysed with respect to ATL1 SNP alleles. Sixty-two chromosomes in total had ATL1 polymorphism A which results in a frequency of 56%. Sixty chromosomes were detected with ATL1 polymorphism G (44%). Distribution of ATL1 alleles A and G among control group CGG repeat alleles is shown in Figure 3.2. Polymorphism A was observed in 17 different individual CGG repeat alleles and the G polymorphism in 19 alleles respectively. Statistically significant association for individual CGG alleles and ATL1 SNP was found for alleles 29 and G ( $p = 0.001$ ); 30 CGG repeats and A ( $p < 0.0001$ ) and allele 31 with A ( $p = 0.0013$ ). For allele 23 significant associations with the ATL1 allele G was not confirmed.



**Fig. 3.2. Distribution of ATL1 alleles A and G among non-pathogenic CGG repeat alleles.**

To discover ATL1 SNP distribution among stable CGG repeats and presumably unstable CGG repeats, the control group was divided into two subgroups according to CGG allele repeat number – normal size alleles and grey-zone alleles (Table 3.1). In grey-zone subgroup polymorphism A and polymorphism G were found with a relative frequency of 0.31 (30.76%) and 0.69 (69.24%) respectively. Distribution of polymorphism G significantly differ between normal size CGG alleles and grey-zone CGG alleles ( $p = 0.0271$ ).

Table 3.1.

**Frequencies of ATL1 SNP's among CGG Alleles**

ATL1	Normal CGG alleles (16-34 repeats)			Grey-zone CGG alleles (35-50 repeats)		
	n	RF	%	n	RF	%
<b>A</b>	54	0.5625	56.25	8	0.3076	30.76
<b>G</b>	42	0.4375	43.75	18*	0.6924	69.24
<b>Total</b>	96	1.000	100	26	1.000	100

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

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

### 3.4. Repeat Structure of Grey-Zone Alleles

26 grey-zone alleles were analysed using direct sequencing, to characterise CGG repeat interruption by AGG trinucleotides. In twelve chromosomes, a CGG interspersed pattern with three AGG's was detected. Twelve chromosomes with two AGG, one chromosome with one AGG and one pure CGG tract were also detected. Ten different AGG interruption patterns were detected (Table 3.2.). For all chromosomes, loss of AGG was detected on 3' end of the sequence.

Table 3.2.

**The Structure of Grey-Zone Allele CGG Array's**

Pattern of AGG interruption	Number of chromosomes	Rel. frequency
Pure	1	0.038
9+n	1	0.038
9+9+n	10	0.384
9+9+9+n	4	0.154
9+10+6+n	2	0.077
10+9+n	2	0.077
10+9+9+n	2	0.077
10+9+10+n	2	0.077
10+9+n+n	1	0.038
10+n+n+n	1	0.038

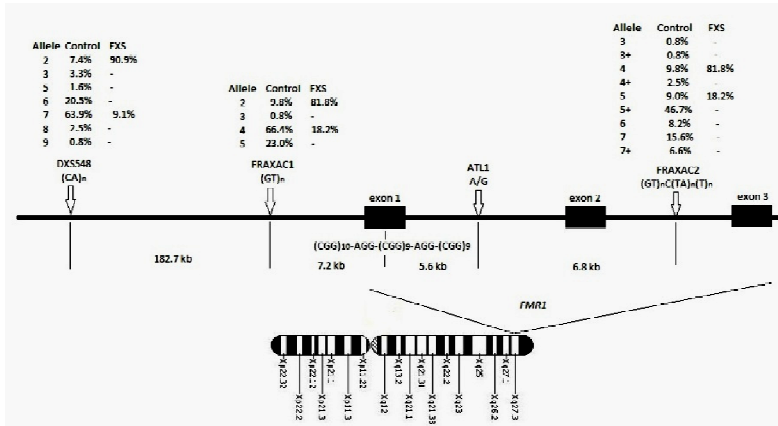
AGG pattern of CGG tract, the digit to the CGG repeat number, "n" corresponds to an uninterrupted CGG repeat number and "+" denote the AGG interspersed position.

The CGG repeat structure was analysed with respect to ATL1 SNP alleles. Significant associations were found, firstly for the allele A and a repeat array with a 10+n structure ( $p = 0.001$ ) and secondly for the allele G and a repeat array with 9+n structure ( $p = 0.004$ ).



### 3.5. DXS548-FRAXAC1-ATL1-FRAXAC2 Haplotypes

Microsatellite markers and ATL1 SNP were analysed in both a control group and FXS patient group (Fig. 3.3.)



**Fig. 3.3. The location, nomenclature and frequencies of *FMR1* linked STR- and SNP-polymorphisms, tested in this study.**

The haplotype of each chromosome was formed by combining microsatellite markers and single nucleotide polymorphism according to their position at the FRAXA locus. STR and SNP marker haplotypes were combined as follows: DXS548-FRAXAC1-ATL1-FRAXAC2. In total 27 different haplotypes were detected – 26 in the control group and three in the FXS group. Only one haplotype from the FXS group was unique (Table 3.3.).

Among the FXS patients, haplotype 2-2-G-4 was found at a relative frequency of 0.818 ( $p < 0.0001$ ). The most common haplotype among control group chromosomes was 7-4-A-5+ (RF = 0.327;  $p = 0.0336$ ). Corrected by Bonferroni this haplotype association is not significant for stable CGG repeat alleles in our population.

Table 3.3.

**Detected DXS548-FRAXAC1-ATL1-FRAXAC2 Haplotypes in Control group and FXS Group**

Haplotype	Control group			FXS group		
	n	RF	S.D.	n	RF	S.D.
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; \* -  $p < 0.0001$ .

To discover haplotype distribution among stable CGG repeats and presumably unstable CGG repeats, the control group was divided into two subgroups according to CGG allele repeat number – normal size alleles and grey-zone alleles.

Haplotype 7-4-A-5+ with a relative frequency of 0.354 was found to be prevalent among normal CGG repeats (16 - 34) (Figure 3.4.). Compared to distribution of haplotypes in grey-zone alleles, this finding was not significant.

Haplotype analysis in grey zone alleles showed the following results (Figure 3.5.). The most common haplotypes in this subgroup were 2-2-G-4 with a relative frequency of 0.308 and 7-4-A-5+ with a relative frequency of 0.231. All alleles with the 2-2-G-4 haplotype had long (> 18 CGG repeat) uninterrupted sequence on 3' end ( $p = 0.0022$ ). Six alleles out of 18, associated with other haplotypes, had the same, uninterrupted CGG repeat pattern.

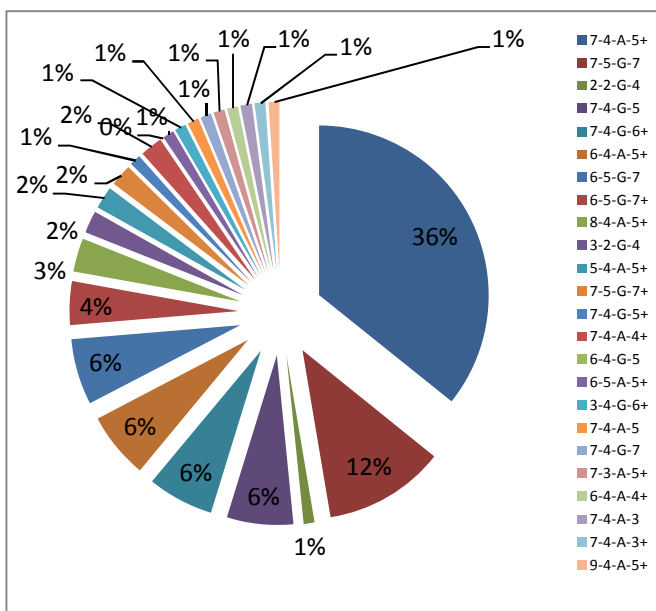


Fig. 3.4. Distribution of *FMRI* Linked Haplotypes among Stable CGG Alleles

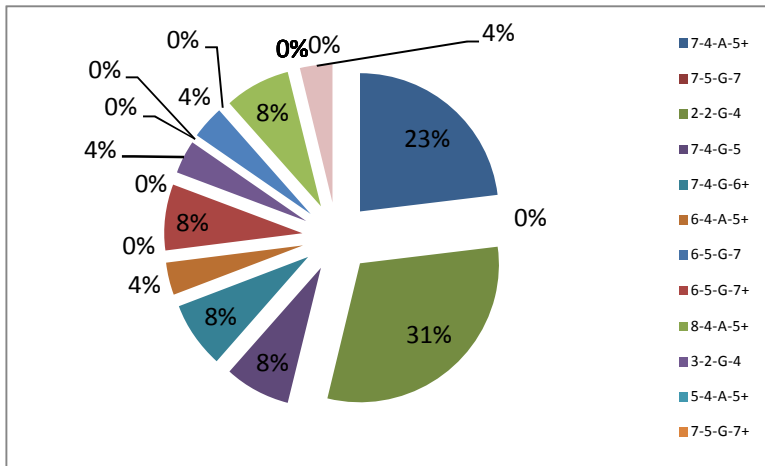


Fig. 3.5. Distribution of *FMRI* Linked Haplotypes in Grey-Zone Alleles

### 3.6. Analysis of Molecular Variance

The level of heterozygosity for all polymorphisms was examined under finite island model and compared. Expected heterozygosity was calculated within control and FXS groups based on allele frequency (Table 3.4.).

Table 3.4.  
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
DXSS48	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.

AMOVA analysis revealed that molecular variation among groups was 27.04%. Molecular variation within groups was 72.96%. The fixation index  $F_{st}$  was calculated based on haplotype frequencies between control and FXS groups and found to be 0.27042 ( $p < 0.001$ ). The critical value of  $F_{st}$  to confirm the null hypothesis, was calculated to be 0.0683 ( $\alpha = 0.01$ ).

### 3.7. Genotype-Phenotype Correlation

Clinical data based on case-records of twelve confirmed FXS male patients were analysed. The age of patients at the moment of diagnosis varied between two and sixteen years (average =  $7.33 \pm 4.46$ ). Molecular diagnostic results for these patients revealed different patterns of CGG repeat expansion. A full repeat size mutation ( $> 200$  CGG repeats) with fully methylated gene promoter region was found in nine patients. Two patients showed premutation/full repeat size mutation mosaic with methylation mosaicism. One patient had a full repeat size mutation with methylation mosaicism (up to 80% unmethylated).

Major clinical symptoms of FXS were analysed for twelve patients. Eight patients out of twelve were tested for IQ. Results revealed, IQ level of patients ranged from 34 to 74 with an average IQ level of  $52.75 (\pm 12.75)$ . In the group of psychomotor symptoms mental retardation, learning difficulties, speech delay and attention-deficit/hyperactivity were observed all patients. From other clinical symptom groups only hypotonia was found in all examined patients.

In order to assess the genotype – phenotype correlation among full mutation alleles and CGG repeat size and/or methylation mosaicism alleles, clinical symptoms were compared between patients with full mutation in lymphocytes and patients with repeat size and/or methylation mosaic. Genotype-phenotype comparison did not reveal significant differences among

patients with full mutation of *FMRI* CGG repeats and patients with CGG repeats and/or methylation status mosaic (Table 3.5.).

Table 3.5.

**Genotype – phenotype comparison among patients with full mutation and patients with mosaic**

<b>Clinical symptoms</b>	<b>Full mutation n = 8</b>	<b>Rel. Frequency</b>	<b>Mosaic n = 4</b>	<b>Rel. Frequency</b>
Psychomotor and neurological symptoms				
Mental retardation	8	1.00	4	1.00
Learning difficulties	8	1.00	4	1.00
Motor development delay	7	0.88	4	1.00
Speech delay/difficulties	8	1.00	4	1.00
Autistic features	6	0.75	1	0.25
Attention-deficit/hyperactivity	8	1.00	4	1.00
Seizure	2	0.25	2	0.50
Balance disturbance	1	0.13	0	0.00
PWS-like phenotype	3	0.38	0	0.00
Dysmorphic features and connective tissues				
Long face	7	0.88	4	1.00
Large ears	4	0.50	3	0.75
High, wide forehead	6	0.75	4	1.00
Prognathia	0	0.00	0	0.00
Hyper elasticity of joints	6	0.75	4	1.00
Flatfoot	4*	0.66	3	0.75
Hypotonia	8	1.00	4	1.00
Recurrent ARI/otitis	5	0.63	2	0.50

\* = 6 patient examined

For improvement of recognition and screening of fragile X syndrome among mentally retarded patients, a clinical questionnaire check-list was adapted from literature and translated into Latvian, for use by family doctors, paediatricians, child neurologists and child psychiatrists.

## **4. DISCUSSION**

### **4.1. Prevalence of the Fragile X Syndrome**

Ten years of experience with molecular diagnostics of the fragile X syndrome in Latvia and a comparable low number of diagnosed patients with this disease led to the question, how prevalent is fragile X syndrome in the Latvian population? Lack of studies in our geographical region was one more factor that inspired us for this study.

As with other published studies, the target population of our study were patients with mental retardation and/or developmental disabilities, as these characteristics are the main symptoms of FXS. It is of prime importance to screen patients demonstrating symptoms of fragile X syndrome and whilst at the same time increase the detection rate for this disease.

This study of mentally retarded males results in 2.67% of prevalence in the target population. These results are in line with findings from other research group studies of populations with a similar clinical symptom range. Our result proves the importance of clinical symptoms recognition related to FXS syndrome in clinical practices and necessity to suggest a check-list of symptoms for clinical specialists to allow easier detection of patients with suspected FXS.

In this study, to assess prevalence of FXS in entire male population, we attributed our detected prevalence of FXS in target population to the total number of patients with psychological development disorders or behavioural and emotional disorders diagnosed in Latvia in same time period. These data were obtained from published data source of the Central Statistical Bureau of Latvia. The gender structure and a detailed overview of included diagnosis were not available, but the overall patient description was the most appropriate for the comparison with a target population of our study. This can be a source of inaccuracy of calculated prevalence in entire male population. As it was

mentioned before, the calculated prevalence in different studies correlate with the spectrum of clinical symptoms chosen for target population.

Our results are consistent with FXS prevalence in Europe. Crawford and colleagues in their fragile X syndrome epidemiology review indicated the necessity of a large population screening for the complete ascertainment of disease prevalence. We completely agree with such a necessity and, even more important, to find out the prevalence of the premutation carrier women in a population of different ethnical backgrounds.

#### **4.2. Variation of CGG Trinucleotide Repeats**

In our study we analysed the distribution of normal CGG repeat alleles among unrelated mentally retarded male patients. The prevalent allele detected in this study agrees with reports from populations across Europe and Western European descents from America, and it is allele 30. If we compare results from our study with results from a study by Estonian colleagues, there are no significant differences. Allele 30 was found in 29.30% of all chromosomes in Estonian patients (Puusepp et al., 2008), and 29.95% of all chromosomes in Latvian patients.

The distribution of normal CGG repeat alleles are described in different populations. It is possible that discrepancy of data in one single repeat unit rose from different methods used for CGG number detection and genotyping errors.

Distribution of CGG repeats in Latvian X chromosomes did not reveal any significant differences among our data and data from European populations. The total heterogeneity of CGG allele distribution in our population was assumed to be in line with data from European populations.



### **4.3. ATL1 SNP**

Previous studies have suggested linkage of CGG tract instability with three factors: the G allele of ATL1 SNP; specific microsatellite marker haplotypes; and a CGG tract AGG interspersion pattern exhibiting a long uninterrupted CGG repeat at the 3' end (Arrieta et al., 2003; Crawford et al., 2000; Curlis et al., 2005; Dombrowski et al., 2002; Eichler et al., 1996; Gunter et al., 1998; Zhou et al., 2006).

Interesting study proposed the hypothesis that polymorphism A originated as a mutation in the 30 CGG repeat array linked with the haplotype 7-3-4+. Through either selection or genetic drift, the allele A become the prevalent allele associated with normal CGG repeats in Western European descent populations (Gunter et al., 1998). If we compare this finding with results from our study, the haplotype 7-4-5+ is the most prevalent among the 30 CGG alleles (51.72%) and all alleles with this haplotype were associated with ATL1 polymorphism A.

Our results revealed a statistically significant prevalence of the ATL1 polymorphism G among grey zone alleles and full mutation alleles, which is an indicator of instability.

### **4.4. Repeat Structure of Grey-Zone Alleles**

One of the tasks in this study was to detect the structure of CGG repeats among X chromosomes with normal CGG repeat alleles. Since normal range CGG repeats (5 – 34) is considered to be stable in transmission, we decided to detect the CGG array just in grey-zone alleles (35 – 50). Grey-zone alleles are normally expressed and do not leads to FXS phenotype but, these alleles may show instability in transmission. For this reason these alleles are good a target for the study of instability factors.

The absence of AGG interruptions in long tracts of CGG repeats have been described as the main factors related to this instability (Rife et al., 2004). It is hypothesised that CGG expansion occurs only at the 3' end of the triplet array. There are various patterns of AGG interruptions of the array that are believed to be responsible for “stabilizing” the alleles (Dombrowski et al., 2002). In our study ten different CGG arrays for alleles with 35 to 50 CGG repeats were found. One allele had a pure CGG tract. The most common pattern of grey-zone alleles was 9+9+n. The association of CGG tract pattern results and specific haplotypes is discussed in the next section.

#### **4.5. DXS548-FRAXAC1-ATL1-FRAXAC2 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 *FMRI* gene CGG repeats. To achieve this task a case-control study was made. The AMOVA data based on calculated  $F_{st}$  suggested that the differences between detected haplotypes within the control and FXS groups were significant and both our population subgroups show different genetical backgrounds.

Several studies have identified specific haplotypes associated with FXS patients chromosomes and normal CGG repeat alleles across European populations (Arrieta et al., 2003; Dokic et al., 2008; Malmgren et al., 1994; Peixoto et al., 1998; Pekarik et al., 1999; Rajkiewicz, 2008). In Caucasians, haplotypes 6-4-4; 6-4-5 and 2-1-3 were reported as haplotypes positively associated with full CGG repeat expansion (Eichler et al., 1996). However, only a limited number of these studies focused on populations from Eastern and North-eastern Europe.

Different microsatellite markers were used for these haplotype analyses in European Caucasian populations. Thus, comparison of our results with these analyses would prove difficult. 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.

In Latvian population, 7-4-A-5+ was determined as the prevalent haplotype for normal CGG alleles. However, after the Bonferroni correction, this finding was not considered to be statistically significant. From literature data the prevalent haplotypes in English population and Western European descended populations of the USA were 7-3-4 and 7-3-4+. These haplotypes were not detected in our control group. Though, taking in to account that most published data from European populations analysed two microsatellite markers, we can therefore compare our results with two marker haplotypes. The haplotype 7-4 constructed from two microsatellite markers DXS548 and FRAXAC1 respectively, was the prevalent one in studies from Basque valleys, Czech, Croatia, Poland and Portugal (Arrieta et al., 2003; Dokic et al., 2008; Pekarik et al., 1999; Peixoto et al. 1998; Rajkiewicz, 2008). In our study, DXS548-FRAXAC1 haplotype 7-4 was found in 72 out of 122 control group chromosomes. This finding shows prevalence of the haplotype 7-4 in our control group.

Furthermore, haplotype 2-2-G-4 was found to be in positive association with full mutation CGG alleles in Latvian FXS chromosomes. In contrast, haplotypes 6-4-4; 6-4-5 and 2-1-3 were reported as positively associated with FXS in Western European descents. These haplotypes were not detected in our FXS group but, nevertheless haplotype 6-4-5 was detected in two grey-zone alleles (41 and 42 CGG repeats respectively). Both alleles were associated with ATL1 polymorphism G, and both had CGG tract pattern

9+9+n. All these findings for 6-4-5 haplotype linked alleles may provide evidence of possible instability for these alleles in later generations.

Grey zone alleles featuring a long ( $\geq 18$  repeats) uninterrupted CGG tract at the 3' end were found to be in positive association with the haplotype 2-2-G-4. In grey-zone alleles haplotype 2-2-G-4 was detected in 8 alleles. All these alleles had a CGG tract pattern where the first AGG interspersion occurred after nine CGG repeats (9+ structure). This finding is in line with Gunter's and colleagues suggested "positively" associated haplotypes with the fragile X mutation. Haplotype 7-4-A-5+ was found in six grey-zone alleles and all alleles had CGG tract pattern where the first AGG interspersion occurred after ten CGG repeats (10+ structure). This haplotype might be a "protective" haplotype for CGG tract stability.

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 either premutation or mutation over generations.

To the best of our knowledge, specific *FMRI*-linked haplotypes in the Baltic State region and North-eastern Europe have not been previously described. The present study is the first to report Latvian population *FMRI* 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. We conclude that a founder effect could not be an explanation of our findings, on the basis of heterogeneity exhibited in the Latvian population and on the basis of a lack of studies across this geographical region. The small number of FXS chromosomes analysed in this study was restricted by a low pick up rate of the fragile X syndrome in our

population. It could however, be a source of incomplete data for the FXS chromosomes linked haplotypes. A larger study of *FMRI*-linked haplotypes in Eastern and North-eastern European regions may provide more accurate data. Nevertheless we consider that our data provide evidence of a specific mutational pathway for unstable CGG alleles in our geographical region.

#### **4.6. Genotype-Phenotype Correlation**

Clinical symptoms are crucial for patients with fragile X syndrome detection among the mentally retarded population. Recognition of these symptoms is a first step towards a successful diagnosis of the fragile X syndrome and the subsequent cascade testing among family members. Assessment of a genotype – phenotype correlation among diagnosed patients can help predict a prognosis for FXS patients and allow exploration for a diversity of symptoms.

In our study we detected three patients with repeat size/methylation mosaic. The observed unmethylated premutation repeat size varied from 78 to 150 CGG repeats. The phenotype of patients with repeat size/methylation mosaic did not revealed significant differences from full mutation FXS phenotype. The only remarkable observation was a lack of autistic features for these patients. At the same time we should admit that for one patient with full mutation autistic features also were also not observed. Neither of our patients diagnosed with repeat size and /or methylation mosaicism showed any signs of milder phenotype which was apparent for patients diagnosed with full mutation and methylation. This can probably be explained by mitotic instability of the expanded CGG tract and possible mosaic in different tissue. Based on these observations we do not suggest making predictions on patient phenotype based solely on genotype data obtained from leucocytes DNA study.

There is a report of eight FXS patients with “Prader-Willi like” phenotype (de Vries et al., 1993). The patients had features resembling the Prader-Willi syndrome (PWS), such as truncal obesity, hypogenitalism, and small hands and feet. In our study three patient with “PW- like” phenotype were diagnosed having fragile X syndrome. Three patients out of twelve diagnosed with “PW- like” phenotype is remarkable number. In our opinion all patients with PWS phenotype showing mental retardation and/or autistic features should be tested for fragile X syndrome.

The low detection rate for patients with fragile X syndrome demonstrated in our study led to the conclusion that fragile X syndrome is generally clinically unrecognised. To help recognise patients with fragile X syndrome among the mentally retarded male population we adapted clinical check list based on a review of relevant literature and translated it into Latvian.

In our opinion, the low number of diagnosed patients was not only due to the failure to clinically recognise fragile X syndrome, but also due to the attitude of society toward mentally handicapped people and their families. In our experience, families with diagnosed FXS patients refuse to inform relatives at risk thereby preventing family genetic consultation. For clinical specialists, a lack of specific treatment for fragile X syndrome place this diagnose in line with other psychiatric diagnosis with symptomatic treatment.

## 5. CONCLUSIONS

1. The prevalence of fragile X syndrome in the Latvian male population was estimated to be 1/6428 (95% CI 5538-7552) or 15.55/100 000 males (95% CI 13.24 – 18.05). The prevalence of the fragile X syndrome among mentally retarded male patients was estimated to be 2.67%.
2. The highest incidence among all analysed normal CGG repeat chromosomes was observed for allele 30 (29.95%), allele 31 (13.10%) and allele 29 (12.83%).
3. For individual CGG alleles with a normal CGG repeat number, a statistically significant association with ATL1 SNP was found for allele 29 and G ( $p = 0.001$ ); allele 30 and A ( $p < 0.0001$ ) and allele 31 with A ( $p = 0.0013$ ). Polymorphism G was found to be associated with grey-zone CGG alleles ( $p = 0.0271$ ) and exclusively associated with all FXS alleles.
4. In the case-control study, haplotype 7-4-A-5+ was determined as the prevalent haplotype for normal CGG alleles in the Latvian population. However, after Bonferroni correction, this finding was not considered to be statistically significant. Results of this study imply that in the Latvian population, haplotype 2-2-G-4 is a marker of CGG tract instability. AMOVA results revealed distinct genetic background for FXS chromosomes.
5. Analysing the structure of grey-zone alleles, revealed ten different CGG arrays. The most common pattern of grey-zone alleles was 9+9+n. The CGG array with a 9+n structure associated with the haplotype 2-2-G-4 was recognised as the pattern positively associated with CGG repeat instability. The CGG array with 10+n structure associated with the haplotype 7-4-A-5+ was recognised as a “protective” pattern.

6. The results of genotype-phenotype analysis did not revealed significant correlation among clinical symptoms, observed in FXS patients, and distinct patterns of CGG repeat expansion.

## 6. PUBLICATIONS

### Publications

- Zanda Daneberga, Zita Krūmiņa, Baiba Lāce, Daiga Bauze, Rita Lugovska. Fragilās X hromosomas sindroms pacientiem ar neskaidras etioloģijas garīgo atpalcību Latvijā. *RSU Zinātniskie raksti 2007*. RSU, 2008: 224-227
- Zanda Daneberga, Zita Krūmiņa, Baiba Lāce, Daiga Bauze, Natālija Proņina, Rita Lugovska. Fragile X syndrome in mentally retarded patients from Latvia *Proceedings of the Latvian Academy of Science, Section B*, 2009 (63):70-72
- Zanda Daneberga, Natalija Pronina, Baiba Lace, and Rita Lugovska. *FMRI* Linked Haplotype Analysis in a Mentally Retarded Male Population: *Central European Journal of Medicine*. 2011, 6(6):750-757
- Zanda Daneberga, Zita Krūmiņa, Baiba Lāce, Daiga Bauze, Natālija Proņina, Rita Lugovska. The Fragile X Syndrome: 13 Years of Experience. *Proceedings of the Latvian Academy of Science, Section B*, 2011. *In press*.

### Abstracts

- Daneberga Z., Lace B., Pronina N., Lugovska R. *FMRI* haplotype analysis in mentally retarded male population. *European Journal of Human Genetics* Vol 19 Supp. 2, 2011, pp. 334.
- Bauze D., Ķevere L., Kronberga Z., Riževs A., Jeļisejevs S., Daneberga Z., Dzalbs A., Ločmele Dz., Krūmiņa Z., Andrēziņa R., Lāce B. Autisma un autiskā spektra traucējumu fenotipisko pazīmju, antropometrisko mērījumu un bioķīmisko rādītāju salīdzinājums un analīze. RSU scientific conference, Abstracts, 2011, pp.239.



- Daneberga Z., Lāce B., Proņina N., Lugovska R. Ar *FMRI* gēnu saistīto haplotipu analīze. RSU scientific conference, Abstracts, 2011, pp.266.
- Daiga Bauze, Zanda Daneberga, Arnis Riževs, Raisa Andrēziņa, Rita Lugovska, Baiba Lāce. Fragilās X hromosomas sindroma (FXS) genotipa un fenotipa salīdzinājums bērniem ar psihiskās attīstības traucējumiem. RSU scientific conference, Abstracts, 2010, pp.113.
- Daneberga Z., Krumina Z., Lāce B., Bauze D., Pronina N., Lugovska R. Prevalence of Fragile X syndrome in mentally retarded patients from Latvia. *European Journal of Human Genetics* Vol 17 Supp. 2, 2009, pp. 134.
- Daneberga Z., Krumina Z., Lāce B., Bauze D., Pronina N., Lugovska R. Molecular diagnostic of XLMR in mentally retarded males from Latvia. *European Journal of Human Genetics* Vol 16 Supp. 2, 2008, pp. 62.
- Daneberga Z., Krumina Z., Lāce B., Bauze D., Pronina N., Lugovska R. FRAXA locus investigation of mentally retarded patients. *European Journal of Human Genetics* Vol 15 Supp. 1, 2007, pp.198
- Daneberga Z., Krumina Z., Lāce B., Pronina N., Lugovska R. Molecular diagnosis of fragile X syndrome in mentally retarded patients in Latvia. *Laboratorine Medicina* 1(29), 2006, pp.20
- Daneberga Z., Krumina Z., Lāce B., Pronina N., Lugovska R. Molecular diagnosis of fragile X syndrome in mentally retarded patients in Latvia. *European Journal of Human Genetics* Vol 14 Supp. 1, 2006, pp.251

### **Approbation**

- Pre-defence of this thesis was held in joint meeting of Department of Biology and Microbiology, Riga Stradins University, Latvian Society of Medical Genetics and Latvian Association of Human Genetics. July 1<sup>st</sup>, 2011., Department of Biology and Microbiology, Riga Stradins University, Riga, Latvia.

- Zanda Daneberga, Zita Krūmiņa, Baiba Lāce, Daiga Bauze, Natālija Proņina, Rita Lugovska. Fragile X syndrome in Latvia: clinical, molecular and population genetic aspects. 11<sup>th</sup> Conference of Baltic Child Neurology Association. June 16<sup>th</sup> -18<sup>th</sup>, 2011, Riga, Latvia.
- Zanda Daneberga, Baiba Lāce, Natālija Proņina, Rita Lugovska. Ar *FMRI* gēnu saistīto haplotipu analīze. RSU scientific conference, April 14<sup>th</sup> - 15<sup>th</sup>, 2011. Riga, Latvia.
- Zanda Daneberga, Zita Krūmiņa, Baiba Lāce, Daiga Bauze, Natālija Proņina, Rita Lugovska. „Molecular diagnostic of X-linked mental retardation” 9<sup>th</sup> Baltic Congress of Laboratory Medicine, September 18<sup>th</sup> - 20<sup>th</sup>, 2008. Jurmala, Latvia.
- Zanda Daneberga, Zita Krūmiņa, Baiba Lāce, Daiga Bauze, Natālija Proņina, Rita Lugovska. „Fragile X syndrome in mentally retarded patients from Latvia” IV Baltic Genetical Congress, October 9<sup>th</sup> - 12<sup>th</sup>, 2007. Daugavpils, Latvia

## **7. ACKNOWLEDGEMENTS**

Many people have contributed their time and knowledge into this thesis. It is a pleasure to thank those who made this thesis possible.

I would like to thank my supervisor, assoc. prof. Rita Lugovska, for making possible to develop my PhD thesis and grow up in both professional and scientific field at the Medical Genetics Clinic, University Children`s Hospital.

I owe my deepest gratitude to my scientific advisor, Dr. Baiba Lāce, for the time she spent guiding me in statistical methods, reading my thesis and for her selfless contribution to improve this work.

I would like to show my gratitude to Dr. Kristin Eiklid, Ulleval University Hospital, Oslo, Norway for her open attitude, kindly offered technical help with confirming diagnosis and hosting me in her laboratory.

It is an honour for me to express my thanks to Prof. R. A. Wevers and Dr. H. Yntema, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands for technical help with confirming diagnosis and inspiration for this work.

I am grateful to all clinicians at the Medical Genetic Clinic for contribution to collect samples, evaluation of patient's clinical symptoms and my guidance in clinical data analysis. Special thanks to Dr. Zita Krūmiņa, Dr. Daiga Bauze and Dr. Dzintra Ločmele.

I wish to thank Irēna Rogovska and Liāna Pliss for time and advices they gave me in data statistical processing.

I would like to thank laboratory assistant of DNA laboratory, Kristina Morozova, for the preliminary preparation of samples and technical assistance.

I am indebted to many of my colleagues in Laboratory of Genetical Biochemistry and DNA laboratory for support, understanding, advice and patience. Special thanks to Pārsla Vēvere and Natālija Proņina.

I wish to thank all patient families for participating in this research, their cooperation are highly appreciated.

Finally, very special gratitude goes to my family, my partner and my friends for faith in me, for understanding and patience during the many hours I spent in laboratory and at my computer.

## 8. REFERENCES

1. Agresti A. and Coull B. A., Approximate is better than "Exact" for interval estimation of binomial proportions. *The American Statistician*. 1998; 52:119-126.
2. Arrieta, I., O. Penagarikano, M. Telez et al. The FMR1 CGG repeat and linked microsatellite markers in two Basque valleys. *Heredity* 2003; 90:206-211.
3. Bardoni B., Mandel J-L. and Fish G. S. FMR1 gene and fragile X syndrome. *Am. J. Med. Genet.* 2000; 97:153-163.
4. Bland J. M., Altman D. G. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 1995; 346: 1085-1087.
5. 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.
6. Chong S. S., Eichler E. E., Nelson D. L. et al. Robust amplification and ethidium – visible detection of the fragile X syndrome CGG repeat using Pfu polymerase. *Am. J. Med. Genet.* 1994; 51:522-526.
7. 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.
8. Crawford, D. C., C. E. Schwartz, K. L. Meadows 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.
9. Curlis Y., Zhang C., Holden J. J. A., Kirkby K., Loeesch D., 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(1):137-151.
10. de Vries B. B. A., Fryns J-P., Butler M. G., Canziani F., Wesby-van Swaay E., van Hemel J. O., Oostra B. A., Halley D. D. J, Niermeijer M. F. Clinical and molecular studies in fragile X patients with a Prader-Willi-like phenotype. *J Med Genet* 1993; 30: 761-766.
11. de Vries B. B. A., Jansen C. C. A. M., Duits A. A., Verheij C., Willemsen R., van Hemel J. O., van den Ouweland A. M. W., Niermeijer M. F., Oostra B. A., Halley D. D. J. Variable FMR1 gene methylation of large expansions leads to variable phenotype in three males from one fragile X family. *J Med Genet* 1996; 33:1007-1010.
12. de Vries, B. B. A., Halley D. J. J., Oostra B. A. et al. The fragile X syndrome. *J. Med. Genet.* 1998; 35:579-589.

13. Dokic, H., Barišić I., Čulic V., Lozic B., Hecimovic S. Haplotype and AGG interspersions analysis of FMR1 alleles in a Croatian population: no founder effect detected in patients with fragile X syndrome. *Hum. Biol.* 2008; 80:581-587.
14. Dombrowski, C., Levesque S., Morel M. L., Rouillard P., Morgan K., 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. Dracopoli and Haines, 1994).
15. Eichler, E. E., Macpherson J. N., Murray A., Jacobs P. A., Chakravarti A., Nelson D. L. Haplotype and interspersions 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.
16. 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.
17. Fu, Y. H., Kuhl D. P. A., Pizzuti A., Pieretti N., Sutcliffe J. S., Verkerk A. J. M. H., 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.
18. Garber K. B., Visootsak J. and Warren S. T. Fragile X syndrome. *Eur J Hum Genet.* 2008; 16:666-672.
19. 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.
20. Hernandez R. N., Feinberg R. L., Vaurio R., Passanante N. M., Thompson R. E., Kaufmann W. E. Autism Spectrum Disorder in Fragile X Syndrome: A Longitudinal Evaluation. *Am J Med Genet A.* 2009; 149A(6):1125-1137.
21. Krūmiņa Dž., Kokare I. and Biķis E. Latvijas bērnu fiziskās attīstības novērtēšana. SIA Medicīnas apgāds, 2007, Rīga
22. Kunst, C. B., Zerylnick C., Karickhoff L. et al. FMR1 in Global Populations. *Am. J. Hum. Genet.* 1996; 58:513-522.
23. Larbrisseau A., Jean P., Messier B., Richer C-L. Fragile X chromosome and X-linked mental retardation. *CMA Journal* 1982; 127:123-126.
24. Macpherson, J. N., Bullman H., Youngs S. A. et al. 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.

25. Malmgren, H., K. H. Gustavson, C. Oudet et al. Strong founder effect for the fragile X syndrome in Sweden. *Eur. J. Hum. Genet.* 1994; 2:103-109.
26. Moutou C., Vincent M. C., Biancalana V., Mandel J-L.. Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic. *Hum. Mol. Genet.* 1997; 3:971-979.
27. 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.
28. Oberle I., Rousseau F., Heitz D., Kretz C., Devys D., Hanauer A., et al. Instability of a 550-Base Pair DNA Segment and Abnormal Methylation in Fragile X Syndrome. *Science.* 1991; 252:1097-1102.
29. Oudet C., Mornet E., Serre J. L., Lentes-Zengerling F. T. S., Kretz C., Deluchatt C., Tejada I., Bouet J., Bouet A., Mandel J. L. Linkage Disequilibrium between the Fragile X Mutation and 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.
30. Peixoto A., dos Santos M. R., Seruca R. Amorim A., 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.
31. 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.
32. Phadke S. R. Fragile X syndrome. Orphanet encyclopaedia. 2005; <http://www.orpha.net/data/patho/GB/uk-Fragile-X.pdf>
33. Poustka A., Dietrich A., Langenstein G., Toniolo D., Warren S. T., 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.
34. Puusepp H., Kahre T., Sibul H., Soo V., Lind I., Raukas R., Õunap K. Prevalence of the Fragile X Syndrome Among Estonian Mentally Retarded and the Entire Children's Population *J. Child. Neurol.* 2008; 23:1400-1405.
35. Rajkiewicz, M. 2008. 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. Ph.D. dissertation. Department of Genetics, Institute of Psychiatry and Neurology, Warsaw, Poland.

36. Richards R. I., Shen Y., Holman K., Kozman H., Hyland V. J., Mulley J. C., Sutherland G. R. 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.
37. Ridaura-Ruiz L., Quinteros-Borgarello M., Berini-Aytés L., Gay-Escoda C. Fragile X-syndrome: Literature review and report of two cases. *Med. Oral. Patol. Oral. Cir. Bucal.* 2009; 14(9):434-439.
38. Rife M., Badenas C., Quinto L. et al. Analysis of CGG variation through 642 meioses in Fragile X families. *Mol. Hum. Reprod.* 2004; 10:773-779.
39. Rousseau F., Heitz D., Oberle I., 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.
40. Sambrook J. and Russell D. W. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, 2001. Cold Spring Harbor, NY.
41. Sherman S., Plecher B. A. and Driscoll D. A. Fragile X syndrome: Diagnostic and carrier testing. *Genetics in Medicine* 2005; 7:584-587.
42. Zhou Y., Tang K., Law H. Y., Ng I. S. L., Lee C. G. L., 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.