Universidade de Aveiro Departamento de Biologia
Ano 2017

Bárbara Luísa Cerqueira Rodrigues

The influence of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function

A influência da região repetitiva CGG do gene Fragile Mental Retardation-1 (*FMR1*) na função reprodutiva feminina

DECLARAÇÃO

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

_____

Bárbara Luísa Cerqueira Rodrigues

Universidade de Aveiro Departamento de Biologia
Ano 2017

Bárbara Luísa Cerqueira Rodrigues

The influence of *the* Fragile X Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function

A influência da região repetitiva CGG do gene Fragile X Mental Retardation-1 (*FMR1*) na função reprodutiva feminina

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Molecular e Celular, realizada sob a orientação científica do Prof. Doutor António José Arsénia Nogueira, professor associado com agregação do Departamento de Biologia da Universidade de Aveiro, e da Doutora Paula Jorge, Assistente Principal da Carreira Técnica Superior de Saúde – Ramo de Genética do Centro de Genética Médica Doutor Jacinto Magalhães – Centro Hospitalar do Porto (UMIB/ICBAS/UP) e da Profª. Doutora Maria de Lourdes Pereira, professora associada com agregação do Departamento de Biologia da Universidade de Aveiro.

Dedico este trabalho aos meus pais, Carlos e Luísa, e ainda, ao meu namorado, Luís Filipe, pelo incansável apoio e motivação, sem os quais era impossível alcançar os meus sonhos.

o júri

presidente

Profª. Doutora Maria Helena Abreu Silva
Professora auxiliar do Departamento de Biologia da Universidade Aveiro

Prof. Doutor João Paulo Soeiro Terra Teodoro
Professor auxiliar convidado do Departamento de Ciências da Vida da Universidade de Coimbra (arguente)

Doutora Paula Maria Vieira Jorge
Assistente Principal da Carreira Técnica Superior de Saúde – Ramo de Genética do Centro de Genética Médica Doutor Jacinto Magalhães, Centro Hospitalar do Porto (UMIB/ICBAS/UP) (orientadora)

agradecimentos

A terminar uma das etapas mais importantes da minha vida, chegou o momento de agradecer a todas as pessoas que me acompanharam, ajudaram, apoiaram e contribuíram para a realização deste trabalho. Assim gostaria de agradecer:

À minha orientadora, Doutora Paula Jorge, pela orientação prestada, espírito científico, ensinamentos, disponibilidade, apoio e amizade, incentivando-me diariamente a continuar e a concretizar esta etapa da minha vida. Fez-me sempre sentir parte da equipa e agradeço toda a confiança que depositou em mim.

Ao Dr. Nuno Maia, pelos ensinamentos, disponibilidade, apoio, motivação, incentivo e amizade. Obrigada pelos momentos de descontração que me proporcionou, todo o auxílio laboratorial, por ter deixado muitas vezes o seu trabalho para segundo plano para me poder ajudar. Obrigada por ter sempre acreditado em mim.

Ao Prof. Doutor António Nogueira por toda a contribuição, imprescindível, neste trabalho. Obrigada pelo incentivo, motivação, disponibilidade, ajuda e amizade.

À professora Dr.ª Isabel Sousa Pereira coordenadora do Banco Público de Gâmetas e à enfermeira Alexandra Marques, Centro Materno Infantil do Norte (CMIN), Centro Procriação Medicamente Assistida (CPMA), CHP E.P.E., por terem acedido a colaborar neste estudo de investigação e pela disponibilidade demonstrada

À Prof. Doutora Maria Lourdes Pereira, minha co-orientadora, queria agradecer o apoio, orientação durante o meu trabalho e formação académica.

A todos os elementos do Centro de Genética Médica Doutor Jacinto Magalhães que me acolheram com grande simpatia e amizade mostrando sempre dispostos a esclarecer qualquer dúvida.

Á Doutora Rosário Santos queria agradecer a oportunidade de pertencer ao seu grupo de trabalho, permitindo a elaboração desta dissertação.

agradecimentos

Queria agradecer à minha família e amigos pela paciência, carinho, atenção, força e coragem, foi essencial nos momentos mais difíceis.

Ao meu namorado, Luís Filipe, por me aturar nos piores momentos, me chamar à razão nas minhas incertezas, por estar sempre do meu lado incentivando-me a concretizar todos os meus sonhos.

Por último, mas não menos importante, um agradecimento do tamanho do mundo aos meus pais, Carlos e Luísa, por estarem sempre do meu lado, acreditarem em mim e que era possível.

palavras-chave

Gene *FMR1*, repetições CGG, padrão de interrupções AGG, função reprodutiva, marcadores de reserva ovárica, potenciais dadoras de oócitos.

resumo

A relação entre o número de repetições CGG do gene Fragile Mental Retardation-1 (*FMR1*) e a função reprodutiva em mulheres não é uma novidade. Está descrito que as portadoras de alelos com um número de repetições CGG entre 55 e 200, designados por pré-mutação, têm uma predisposição para desenvolver insuficiência ovárica primária ou menopausa precoce. Porém, a existência de risco de diminuição da função reprodutiva nas mulheres, com genótipos considerados "normais" ($CGG_{<54}$), e respetivos sub-genótipos ainda não é clara. Sabe-se que a presença de interrupções AGG confere a esses alelos uma maior estabilidade, impedindo a expansão do número de repetições CGG para um tamanho considerado patogénico. A forma como o número e o padrão de interrupções AGG poderá influenciar a função reprodutiva feminina, nunca foi estudada. No presente trabalho, os marcadores de reserva ovárica foram correlacionados com o número de repetições CGG e perfil das interrupções AGG. A população em estudo incluiu 50 mulheres jovens e saudáveis, candidatas à doação de oócitos. Dado que o número e o padrão das interrupções AGG não são determinados por rotina, foi então necessário implementar a sua análise, recorrendo a diferentes metodologias: 1) Triplet-Primed Polymerase Chain Reaction; 2) Sanger sequencing; and 3) Restriction Fragment-Length Analysis. Foi realizada uma projeção da associação entre o número de repetições CGG e os níveis hormonais, através de uma análise multivariável, considerando os novos sub-genótipos "normais" *FMR1* previamente definidos. Os níveis hormonais associados às diferentes amostras não foram suficientes para discriminar sub-genótipos, indicando que a individualização das amostras classificadas por sub-genótipos não era possível. Recorrendo a uma fórmula matemática que determina a pontuação alélica, tendo em consideração o tamanho total do alelo, e o número e o padrão de AGG. Após análise estatística foi possível dividir as amostras em dois grupos: um primeiro designado por grupo equivalente e um segundo designado por grupo oposto. O grupo equivalente, que é composto principalmente por amostras que possuem alelos do sub-genótipo *"normal" FMR1,* e o oposto, onde a maioria das amostras possui sub-genótipo *"normal/baixo" FMR1*. No grupo equivalente, observou-se correlação positiva e significativa entre número de folículos antrais e os níveis hormonais: prolactina e hormona luteinizante (LH). Assim, é possível prever o número de folículos antrais produzidos combinando os níveis de prolactina e LH. Estes resultados confirmam publicações anteriores, já que o sub-genótipo *"normal/baixo"* foi anteriormente associado a uma diminuição da reserva ovárica. No geral, este estudo confirma a associação da região repetitiva CGG do *FMR1* na função reprodutiva feminina e sugere que a estabilidade dos alelos é um fator determinante para o sucesso da resposta ovárica.

keywords

FMR1 gene, CGG repeats, AGG interruption pattern, reproductive function, ovarian reserve markers, potential oocyte donors.

abstract

The impact of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repeat number in the female reproductive function is well established. Carriers of a CGG repeat number between 55 and 200, designated a premutation, are prone to develop primary ovarian insufficiency or early menopause. Yet, an impact on the reproductive function in carriers of "normal" genotypes and sub-genotypes ($CGG_{<54}$) is controversial. The presence of AGG in normal-sized alleles confers stability, hampering the expansion of the repeat number in future generations. To the best of our knowledge testing the influence of the AGG number and pattern on the female reproductive function has never been endeavored. Herein, the ovarian reserve markers were correlated with CGG number as well as AGG number and pattern, in female carriers of *FMR1* normal-sized alleles. Our cohort comprised 50 healthy young females, candidates for oocyte donation. Considering AGG number and pattern are not routinely determined different methodologies were implemented: 1) Triplet-Primed Polymerase Chain Reaction; 2) Sanger sequencing; and 3) Restriction Fragment-Length Analysis. A projection of the association between the CGG repeat values and the hormonal levels, by multivariate analysis, was performed, considering the *FMR1* new "normal" sub-genotypes previously defined. The hormonal levels associated with the different samples were not sufficient to discriminate the sub-genotypes, indicating that the individualization of the samples classified by sub-genotype was not possible. Resorting to a mathematical formula that determines the allelic score, taking into account total allele size, and AGG number and pattern. After statistical analysis, it was possible to divide the samples into two groups: a first called an equivalent group and a second called an opposite group. The equivalent group is composed mainly of samples carrying alleles in the *normal FMR1* sub-genotype and the opposite, where most of the samples have an *FMR1 low/normal* sub-genotype. In the equivalent group, a positive and significant correlation was observed between the number of antral follicles and the hormonal levels: prolactin and luteinizing hormone (LH). Thus, it is possible to predict the largest number of antral follicles produced combining the levels of prolactin and LH. These results actually confirm prior publications as the *low/normal* sub-genotype has been previously associated with a diminished ovarian reserve. Overall, this study confirms the association of the *FMR1* CGG repetitive region in the female reproductive function and suggests that the stability of the alleles is a determining factor for the ovarian response success.

INDEX

List of Figures

List of Tables

List of Abbreviations and Sigla

AMH - Anti-Müllerian Hormone

CHP - Centro Hospitalar do Porto

CMIN – Centro Materno Infantil do Norte

DMSO - Dimethyl Sulfoxide

DNA - Deoxyribonucleic Acid

DOR - Diminished Ovarian Reserve

E2 - Estradiol

FM - *FMR1* Full Mutation

***FMR1* -** *Fragile Mental Retardation 1* gene

FMRP - Fragile Mental Retardation Protein

FOR - Function of Ovarian Reserve

FSH - Follicle-Stimulating Hormone

FXPOI - Fragile X - associated Primary Ovarian Insufficiency

FXS - Fragile X Syndrome

FXTAS - Fragile X - associated Tremor/Ataxia Syndrome

GnRH - Gonadotropin-Releasing Hormone

ID - Intellectual Disability

IVF - *In vitro* Fertilization

Kb - Kilobases

LH - Luteinizing Hormone

MAP - Medically Assisted Procreation

PCR - Polymerase Chain Reaction

PM - *FMR1* Premutation

POF - Premature Ovarian Failure

PRL - Prolactin

RNA - RiboNucleic Acid

tAFC - Total Antral Follicle Count in both ovaries.

TP-PCR - Triplet-Primed Polymerase Chain Reaction

UTR - Untranslated Region

CHAPTER 1.

Introduction

“The influence of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function”

Chapter 1. Introduction

1. Background

Infertility is clinically defined as a disease of the reproductive system characterized by the inability to achieve a pregnancy after a year or more of unprotected sexual activity (ESHRE, 2013). According to the most recent estimates about 9% of couples have infertility problems (Boivin *et al.*, 2007). Infertile female due to ovarian failure can pursue fertility treatments and conceive following *in vitro* fertilization (IVF) techniques resorting to oocyte donation (ESHRE, 2013). Oocyte donation is the process by which a fertile and healthy female allows several of her oocytes to be aspirated, usually following ovarian stimulation.

The Portuguese Public Bank of Gametes is a service provided by the Portuguese National Health Service responsible for the recruitment and selection of oocytes and sperm donors. Gametes resulting from volunteer donations are then used in Medically Assisted Procreation (MAP) techniques. The collection and preservation of the gametes is carried out in public specialized Centers, namely at the Banco de Gametas – Centro de Procriação Medicamente Assistida, Centro Materno Infantil do Norte, (CMIN), Centro Hospitalar do Porto (CHP, E.P.E.). In females, the process involves a set of three steps ranging from candidate selection, to hormonal stimulation and donation (oocytes harvest). In the selection phase, several studies are carried out including psychological, blood tests and reproductive function evaluation i.e. ovarian reserve tests. These are of importance in potential oocyte donors to evaluate the ovarian reserve and subsequently determine the gonadotrophin dosage prior to the stimulation. Blood tests include exclusion of sexually transmitted diseases and genetic analyses. The latter comprise karyotype and molecular screening for cystic fibrosis causing mutations and *FMR1* premutation or full mutation alleles. Testing for both genetic disorders is justifiable by the high incidence of carriers in the general population (Hantash *et al.*, 2011). In an oocyte donor, the presence of an *FMR1* premutation besides the risk of expanding to a full mutation causing fragile X syndrome in the fetus, predisposes to a premature ovarian failure (POF) causing a fertility issue in the potential donor (Schenkel *et al.*, 2016). The association between *FMR1* CGG number and infertility is not new. Recent studies however have shown some evidence that this gene can also be associated with other forms of ovarian dysfunction beyond FXPOI (Gleicher *et al.*, 2015).

The work described in this thesis aims to shed light on the influence of the *FMR1* gene CGG repeat length and stability in the female reproductive function. To the best of our knowledge this was never attempted before. The cohort of potential oocyte donors, which usually includes healthy females aging from 18 to 33 years, characterized on both hormonal and genetic profiles, was considered a good control population for these studies. Herein, data from fifty potential donors recruited at the Banco de Gametas, CMIN, CHP, E.P.E. - between 2014 and 2016, were used.

2. Fragile Mental Retardation-1 gene - *FMR1*

2.1.Fragile X Syndrome - Historical overview

Fragile X syndrome (FXS; OMIM #3000624) is the most common cause of inherited intellectual disability (ID) and the second most frequent cause of ID of genetic origin (Saldarriaga *et al.*, 2014). Anciently the diagnosis of FXS was made by cytogenetic techniques through the identification of an unusual constriction at the distal end of the long arm of the X chromosome (fragile site), which was called the "marker X chromosome" (Lubs, 1969). In 1991, the Fragile Mental Retardation-1 gene (*FMR1;* OMIM #309550*)* was identified and established as the underlying cause of FXS (Verkerk *et al.*, 1991).

2.2.*FMR1* gene and coded protein

FMR1 gene is located on the long arm of the X chromosome at Xq27.3 and is composed by 17 exons spanning approximately 40 Kilobases (Kb). Near the gene promoter at 5'UTR a CpG island and a triplet repetitive and polymorphic region composed by CGG is present (Jin and Warren, 2000; Wittenberger *et al*., 2007) (Figure I).

FMR1 gene codes for the fragile X mental retardation protein (FMRP), a selective RNA-binding protein that shuttles between the nucleus and the cytoplasm (Kline *et al.*, 2014).

FMRP is present in neurons, primordial germ cells and granulosa cells of developing ovarian follicles (Sullivan *et al.*, 2011).

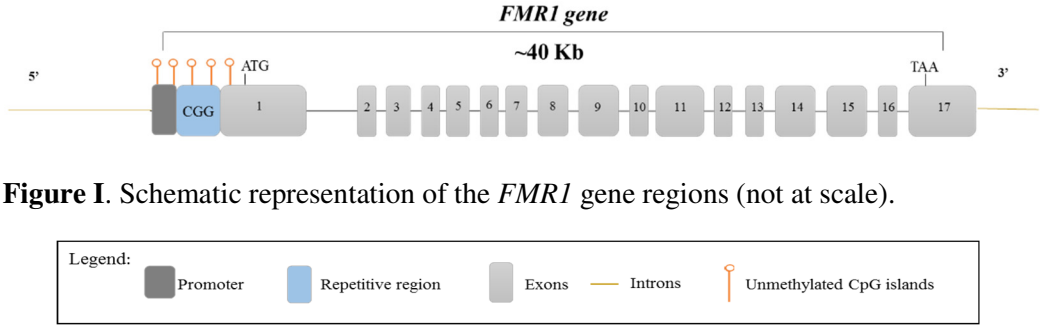


Figure I. Schematic representation of the *FMR1* gene regions (not at scale).

In neurons, FMRP can also be found at the base of dendritic spines and is suggested to participate in the process following synaptic activity. These findings indicate that the absence of FMRP may directly influence neuronal plasticity and thus cause ID in FXS patients (Wittenberger *et al.*, 2007; Hoyos and Thakur, 2017).

2.3. *FMR1* allele categories

Depending on the number of the CGG repeats, four types of alleles can be defined (Biancalana *et al.*, 2015). In the normal population both normal alleles, that have up to 44 CGG repeats, and intermediate or gray zone alleles, with 45 to 54 CGGs, can be observed. The latter can show some degree of instability particularly in meiosis. Alleles with 55 to 200 CGGs are defined as premutations (PM), while alleles containing more than 200 CGG repeats as full mutations (FM), these are usually completely methylated (Schenkel *et al.*, 2016). Both of these pathogenic mutations can be unstable upon transmission causing different phenotypes described in detail below (Biancalana *et al.*, 2015) (Figure II).

2.4. Mechanisms of CGG expansion

Normal-sized alleles are considered stable, e. g., there is usually no change in CGG number when transmitted to the next generation, unlike premutation or full mutation alleles that are considered unstable. The instability in the transmission of intermediate alleles is still debatable. Although the molecular causes of the CGG expansion are not fully understood, there are some risk factors (Maia *et al.*, 2016).

Commonly, the polymorphic region the *FMR1* gene 5'UTR is not composed by pure CGG triplets. The number of CGG repeats in the general population is highly variable (the most frequent allele has 29 or 30 repeats and carrying an AGG interruptions after the ninth or tenth CGG (Oostra and Willemsen, 2003; Maia *et al*., 2016). These AGG interspersions are proposed to be the anchor that avoid DNA slippage during DNA replication. Thus, it is believed that the repeats are more stable when interrupted with AGGs. On the contrary, the loss of an interruption resulting in a long pure CGG sequence at the 3′ end contributes to instability and replication errors. The longer the 3′ pure CGG, the more susceptible is to further expansion. The instability of this region seems to be associated not only with the number, but also with the pattern of the interruptions (Eichler *et al*., 1994; Poon *et al*., 2006). In the majority of premutation and full mutation alleles, no AGG interruption is observed (Penagarikano *et al.*, 2007).

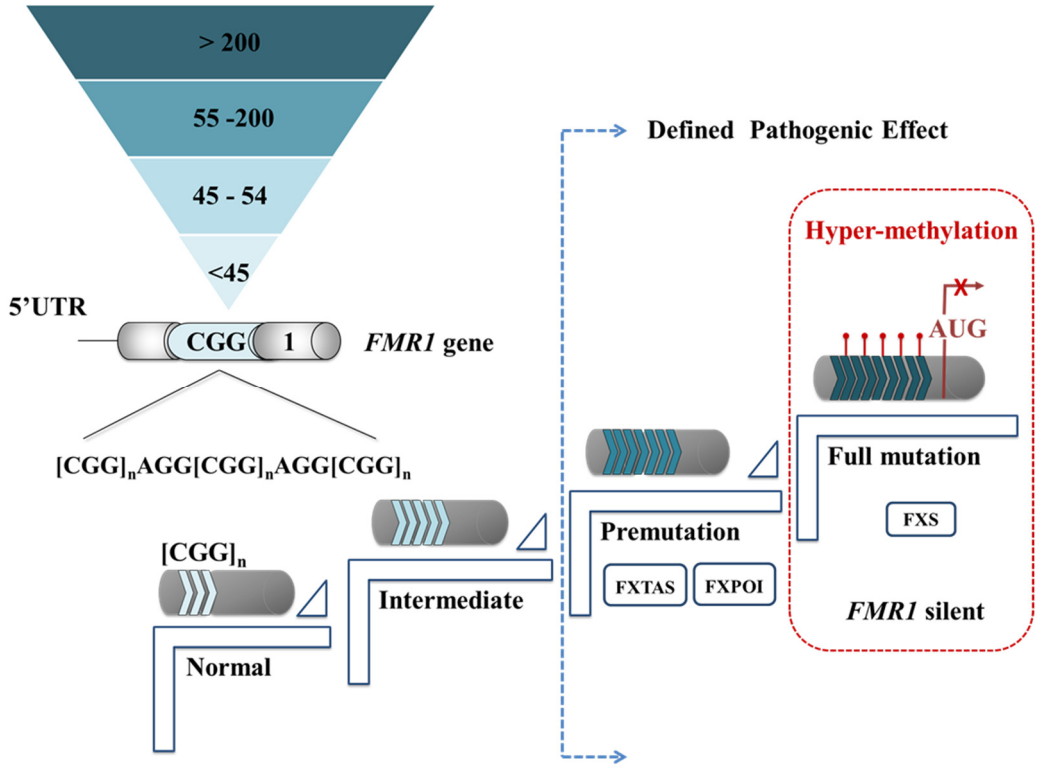


Figure II. Categories of *FMR1* alleles.

Allele classes divided according to CGG repeat number and respective clinical effect: FXTAS - Fragile X Tremor/Ataxia Syndrome; FXPOI - Fragile X Primary Ovarian Insufficiency; FXS - Fragile X Syndrome. Image provided by Laboratório de Genética Molecular, CGMJM, CHP, E.P.E.

2.5. Diseases caused by mutations in *FMR1* gene

2.5.1. Fragile X Syndrome (FXS)

Fragile X syndrome is a neurodevelopmental condition with an estimated incidence of 1 in 4000 males and 1 in 4000-8000 females (Jin and Warren, 2000). The typical characteristics of this syndrome include moderate to severe ID, macroorchidism and distinct facial features: long face and protruding ears (Saldarriaga *et al.*, 2014; Schenkel *et al*., 2016). The majority of FXS patients has an expansion of more than 200 CGGs. These CGG expansions result in the hyper-methylation of *FMR1* promotor which hampers transcription and consequently leads to gene silencing. The clinical features associated to the FXS are caused by absence of the protein coded by this gene, FMRP (Schenkel *et al*., 2016; Hoyos and Thakur, 2017). Other rare mutational mechanism, such as *FMR1* deletions or even rarely missense mutations can cause FXS (Schenkel *et al.*, 2016).

2.5.2. Fragile X - associated Tremor Ataxia Syndrome (FXTAS)

FXTAS (FXTAS; OMIM #300623) is a late-onset (> 50 years) progressive neurodegenerative disorder that affects premutation carriers of both sexes (Wittenberger *et al.*, 2007; Wang *et al*., 2017). This syndrome has a penetrance of around 40% in males and between 8% and 16% in females (Hoyos and Thakur, 2017). In males, the risk of developing FXTAS increases with age (Wittenberger *et al.*, 2007). The lower penetrance of FXTAS among females can be explained by the "protective" effect of the normal and active X chromosome due to the random X- chromosome inactivation (Wittenberger *et al.*, 2007). The premutation alleles are not methylated, therefore the synthesis of FMRP occurs at a level below the normal (Schenkel *et al.*, 2016).

Clinical features of FXTAS include progressive intention tremor, gait ataxia (resulting on frequent falls), cognitive deficits (executive function and memory problems), psychological features (anxiety, depression, mood lability, outburst or reclusive behavior), autonomic dysfunction, parkinsonian features and peripheral neuropathy with diminished sensation in the lower extremities. FXTAS brain magnetic resonance image shows progressive global brain atrophy and white matter disease of both cerebellum and cerebrum.

Before the onset of tremor or ataxia it is possible to detect cerebellar and brainstem atrophy and ventricular enlargement (Wittenberger *et al.*, 2007; Wang *et al*., 2017) .

2.5.3. Fragile X - associated Primary Ovarian Insufficiency (FXPOI)

Fragile X primary ovarian insufficiency (FXPOI; OMIM #311360) is a disorder defined as cessation of menses (menopause) before age 40 years, also known as premature ovarian failure (POF) (Voorhuis *et al.*, 2013).

The diagnosis is based on amenorrhea for at least 4 months and elevated serum follicle-stimulating hormone (FSH) levels in the menopausal range (usually above 40IU/mL). This syndrome presents other characteristics, such as shorted follicular phases and low anti-Müllerian (AMH) levels (Streuli *et al*., 2009; Tural *et al*., 2014; Eslami *et al*., 2016). Premature ovarian failure is clearly an heterogeneous disorder with at least 13 distinct loci identified so far (/www.omim.org/entry/311360; assessed on 2017/08/5). POF1 locus located within Xq26-q28 region is associated with the premutation in the *FMR1* gene (Eslami *et al.*, 2016).

Female full mutation carriers do not develop ovarian dysfunction, which suggests that the mechanism underlying this clinical entity is not caused by the absence of the FMRP (Streuli *et al.*, 2009). In the premutation range, FMRP is expressed but the *FMR1* mRNA levels are increased causing *FMR1* mRNA protein complexes, which accumulate and interfere in the normal cell function, and eventually cell death (Ruth *et al.*, 2016). The increase of *FMR1* mRNA transcription level is proportionally related to the number of CGG in the premutation range (Streuli *et al.*, 2009). However, the risk of FXPOI is greater in carriers of premutated alleles from 80 to 100 CGGs than in those with higher repeat number (Voorhuis *et al.*, 2013) (Figure III).

Ovarian dysfunction is believed to be caused by a toxic effect due to the accumulation of *FMR1* mRNA in ovarian cells (granulose and oocyte cells), leading to follicular atresia and impairing follicle development, by a mechanism that remains to be elucidated (Sullivan *et al*., 2005; Willemsen *et al.*, 2011).

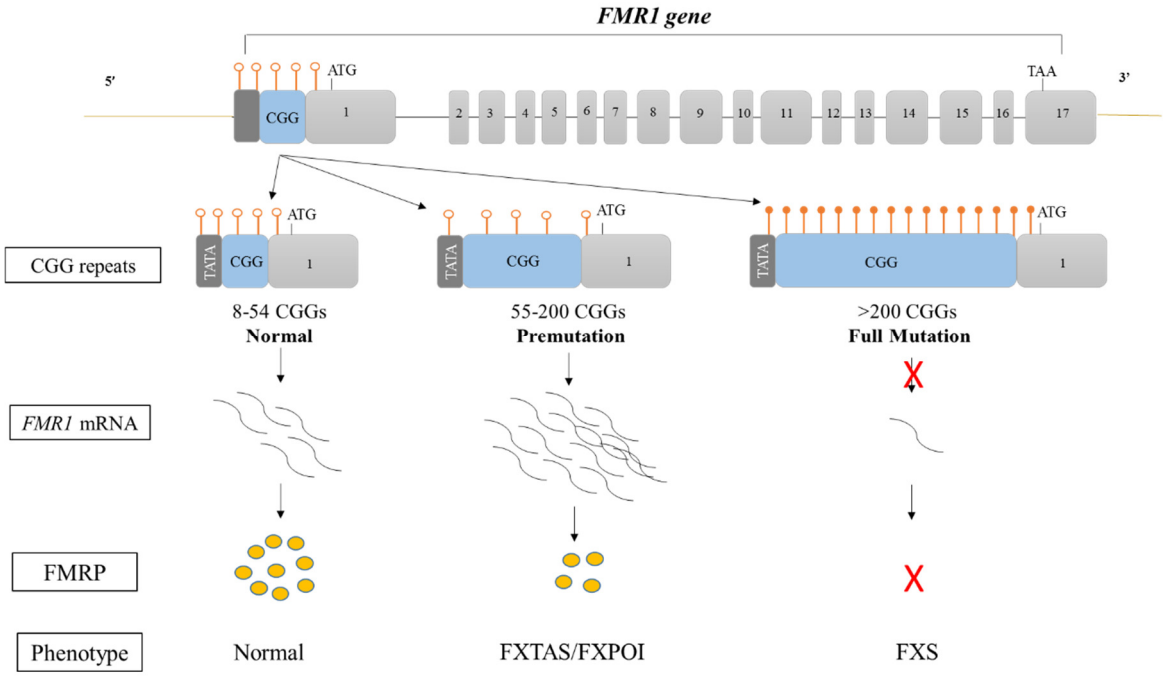


Figure III. Schematic representation of the effect caused by the different categories of *FMR1* alleles (adapted from Wittenberger *et al*., 2007).

2.6. *FMR1* in female fertility

The association of *FMR1* gene with female infertility is not surprising as it is long known that carriers of a premutation ($_{55\leq}[CGG]_{\leq 200}$) have increase risk for POF (Gleicher *et al.*, 2015). The association of the *FMR1* gene with other forms of ovarian dysfunction such as pathologic diminished ovarian reserve (DOR), have also been described but are less clear (Pfeifer *et al.*, 2015).

2.6.1. Female *FMR1* new "normal" genotypes

Based on their work on *FMR1* and ovarian function, recently Gleicher *et al*., (2014) suggested a new normal range sub-genotype: $_{26\leq}[CGG]_{\leq 34}$. These authors have shown that carriers of both or one allele with a repeat number above or below this "new" normal range are in risk of DOR (Gleicher *et al.*, 2014). Several new normal *FMR1* sub-genotypes are being considered (Table 1).

This subject is controversial and is still under discussion, since several studies seem to have contradictory conclusions. Some authors found no relation between *normal/high* or intermediate-sized alleles and premature ovarian failure or DOR (Bennett *et al.*, 2010; Murray *et al*., 2014 ; Kline *et al*., 2014 ; Voorhuis *et al*., 2014). A different study with 1,287 infertile woman suggests no correlation between alleles with a CGG number between 35 and 54 and the ovarian reserve markers FSH, AMH or antral follicles count (Schufreider *et al.*, 2015).

Thus, it is still unclear whether *FMR1* alleles in the normal range are associated with diminished ovarian reserve and if the repeat length confers risk for diminished fertility.

Table I. Normal *FMR1* sub-genotypes.

Sub-genotype	$_{8<}$**[CGG]**$_{<26}$	$_{26<}$**[CGG]**$_{<34}$	$_{34<}$**[CGG]**$_{<55}$
normal		A1 and A2	
normal/high		A1	A2
low/normal	A1	A2	
low/high	A1		A2
high/high			A1 and A2
low/low	A1 and A2		

A1- Allele with smallest CGG repeat number; A2- Allele with largest CGG repeat number.

2.6.2. Ovarian reserve decline

The ovary contains a finite number of oocytes, called the ovarian reserve, that are subject to maturation to reach the ovulatory stage. Ovarian reserve is characterized by the reproductive potential as a function of the number and quality of remaining oocytes (Pfeifer *et al.*, 2015 ; Eslami *et al*., 2016). A fecundity decay is associated with ovarian follicular depletion and decreased oocyte quality (Verissimo and Silvestre, 2016). A reduction in the quantity and quality of aged oocytes is a normal physiological occurrence, however some females show a premature decline of the ovarian reserve (Pastore *et al.*, 2014) (Figure IV). A diagnosis of DOR is attributed to females at reproductive age, having regular menses,

with reduced ovarian stimulation or fecundity, when compared with normal females of similar age (American Society for Reprodutive Medicine, 2015). Females with DOR may have reduced fecundity, or a limited response to ovarian stimulation with fertility medications. However, DOR does not directly imply an inability to conceive (Man *et al.*, 2017).

Various circumstances can have an adverse effect on the ovarian function, but in most cases the exact cause of DOR is unknown (Pfeifer *et al.*, 2015). Some examples of those conditions are: aging, environmental factors (e.g. smoking), autoimmune diseases, iatrogenic agents (e.g. chemotherapy), endometriosis with ovarian involvement and genetic abnormalities (Pfeifer *et al.*, 2015 ; Eslami *et al.*, 2016).

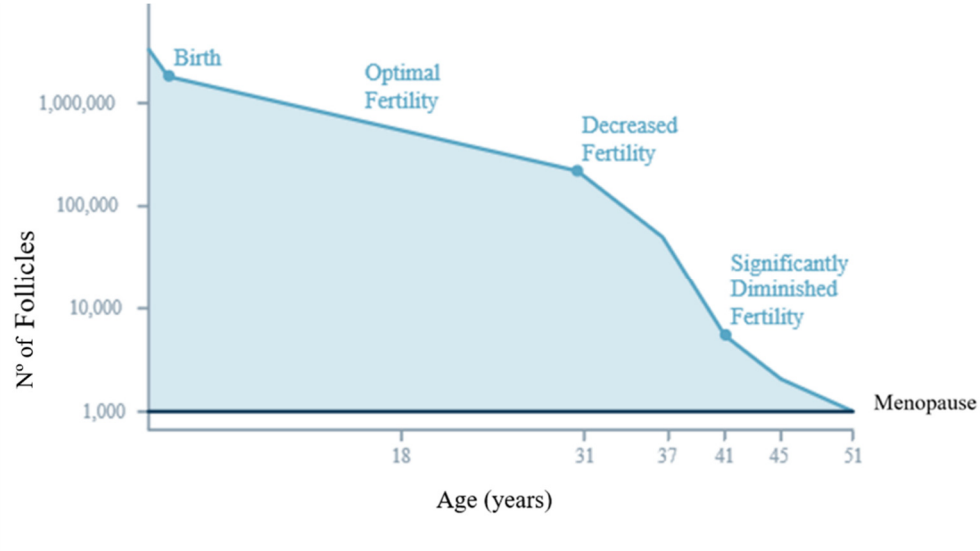


Figure IV. Graphical representation of the follicle number according to age (http://www.eggfreezing.org.uk/assessed on 2017/09/21).

2.6.3. Ovarian reserve testing and markers

The possibility of knowing the ovarian reserve at a given point assumes a clinical importance, not only to evaluate the follicular pool but also to predict the physiological ovarian failure and terminus of the conception capacity (Verissimo and Silvestre, 2016).

Ovarian reserve markers can be divided into: 1) biochemical markers and 2) ultrasound markers. Biochemical markers usually include measuring the serum levels of AMH, estradiol, FSH, inhibin B, LH and Prolactin. The determination of total antral follicle count in both ovaries corresponds to the ultrasonographic marker (Pfeifer *et al.*, 2015). Figure V shows the process of follicle maturation from the primordial follicle to the ovulatory follicle (folliculogenesis) and the role of different hormones in this process.

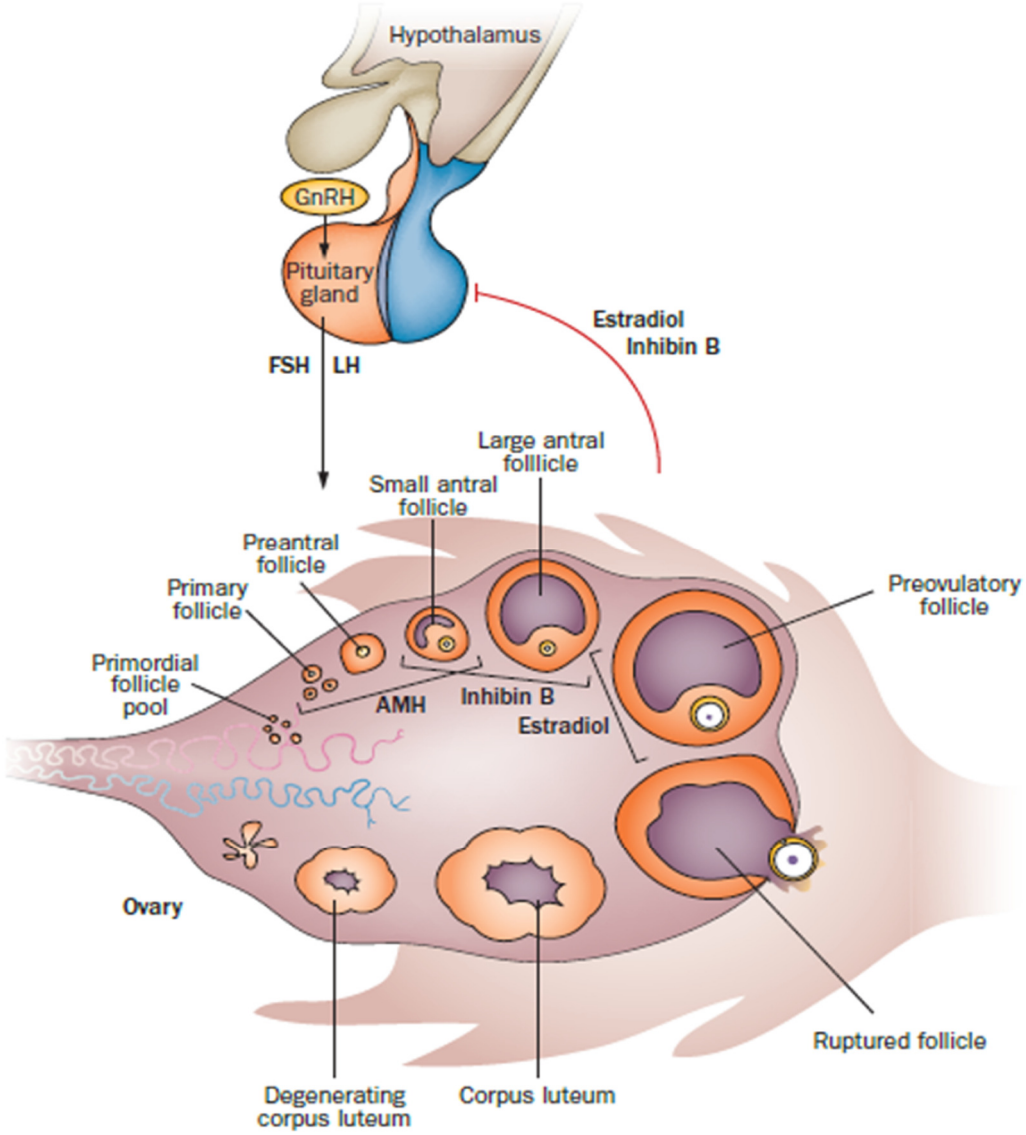


Figure V. Schematic representation of folliculogenesis (from Visser *et al.*, 2012)

2.6.3.1. Hormonal (biochemical) markers

Anti-Müllerian hormone (AMH) is a glycoprotein produced by granulosa cells of pre-antral follicles and small antral follicles as soon as the primordial follicles are recruited (Pfeifer *et al.*, 2015 ; Verissimo and Silvestre, 2016). AMH is a gonadotropin-independent hormone and therefore remains relatively consistent within and between menstrual cycles. Low AMH levels have been associated to a poor response to ovarian stimulation, embryo quality and pregnancy outcomes in IVF (Pfeifer *et al.*, 2015).

Estradiol (E2) is a steroid hormone produced by ovarian follicles of the granulosa cells (Verissimo and Silvestre, 2016). Basal estradiol has a poor inter- and intra-cycle reliability, thereby it should not be used alone as a measure for function ovarian reserve (Pfeifer *et al.*, 2015). In the early follicular phase, when the estradiol level is elevated and the basal FSH concentration is normal, there is no evidence of an association between poor ovarian response and low pregnancy rates in IVF treatments (Pfeifer *et al*., 2015 ; Verissimo and Silvestre, 2016).

Follicle-stimulating hormone (FSH) is secreted by the anterior pituitary gland and plays an important role in the recruitment and maturation of primordial follicles and the induction of ovarian estrogen productions via the aromatase enzyme. A diminished follicular pool leads to a decrease in the estradiol and inhibin B production and, consequently, an increase of FSH (Verissimo and Silvestre, 2016)**.** High FSH values have been associated to a diminished ovarian reserve. However, FSH test, shows significant inter-and-intra cycle variability, which is a limitation (Pfeifer *et al.*, 2015).

Inhibin B is secreted by the ovary granulosa and teak cells. It inhibits the FSH secretion and has a paracrine action in the follicles development. An inhibin B decrease on the third day of the menstrual cycle may predict a low ovarian reserve. However, inhibin B is not a reliable measure because its levels can increase after gonadotropin-releasing hormone (GnRH) or FSH stimulation, exhibiting high intra-cycle variability (Pfeifer *et al.*, 2015).

Luteinizing hormone (LH) is a hormone secreted by the adenohypophysis to stimulate the formation of the yellow body in the ovaries at the Graaff follicle rupture site, in order to produce progesterone. LH plays an essential role in oocyte maturing and development, in

IVF cycles (Canêdo, 2015). An elevated (>3,6 mUI/mL) FSH:LH ratio, determined on day 3 of the cycle, can be a signal of DOR (Mukherjee *et al.*, 1996 ; Roudebush *et al.*, 2008).

Prolactin (PRL) is the secreted by the adeno-hypophysis and its main function is to stimulate the production of milk by the mammary glands Hyperprolactinemia, can cause menses alteration and infertility (Melmed *et al.*, 2011).

2.6.3.2. Ultrasonographic marker

Antral Follicle Count (AFC) – The total antral follicle count corresponds to the total number of antral follicles, measuring between 2-10 mm in diameter, in both ovaries, seen with transvaginal ultrasound during the early follicle phase. A low AFC number (3-6 total antral follicles) is considered a predictor of poor ovarian response and pregnancy outcomes after IVF. A significant correlation was observed between the antral follicles and the number of oocytes obtained after FSH stimulation. The clinical utility of AFC is limited due to its intra- and inter- cycle variability (Pfeifer *et al.*, 2015 ; Verissimo and Silvestre, 2016).

Table II shows serum hormonal reference values and antral follicles number. Potential oocyte donors are exclued if one of these values falls outside reference value establishsed by Banco de Gametas, CMIN, CHP, E.P.E.

Table II. Reference values ovarian reserve markers.

Markers	Refence values*
FSH	<10 mIU/mL
Inhibin B	<40 pg/mL
Estradiol	<60 pg/mL
FSH:LH	<3,6 mUI/mL
Prolactin	<25 ng/mL
AFC	>6 tAFC

*Established by Banco de Gametas, CMIN, CHP, E.P.E.

Overall, there is no consensus on the test, or combination of measures to best evaluate the ovarian reserve. Several authors suggest that AFC combined with serum AMH show a good descriminatory potential as ovarian reserve predictors. The AMH has no inter cyclic variability and AFC has the advantage of being noninvasive (compared with hormonal tests) and its variation between cycles do not affect the ovarian response (Verissimo and Silvestre, 2016).

3. Objectives of the thesis

The present thesis endeavors to evaluate the influence of the *FMR1* gene on the ovarian reserve. The association between *FMR1* CGG number and the reproductive function in healthy females will be investigated. Considering the CGG repeat is interrupted by AGG, in the majority of normal-sized/control alleles, not only the size of the repeat but also the pattern of the AGG interruptions will be considered. No previous literature reports were found describing a similar approach in normal sized-alleles.

Anonymized data from 50 potential oocyte donors were obtained from Banco de Gametas of the Centro de Procriação Medicamente Assistida – CMIN, CHP, E.P.E. The data includes: age at first consultation, total antral follicle count and serum levels of Estradiol, FSH, LH, and Prolactin. *FMR1* CGG repeat number (100 alleles) was determined at Centro de Genética Médica Doutor Jacinto Magalhães – CHP, E.P.E.

Specific aims of the present work include:

a) Statistical analysis aiming a correlation between hormonal levels and *FMR1* CGG repeat number.

b) Determination of the AGG pattern of the 100 alleles using different methodologies: Triplet-primed PCR, Sanger sequencing and restriction fragment analysis.

c) Establish the influence of *FMR1* CGG number and AGG pattern on the ovarian reserve and ultimately in the female fertility.

4. Thesis organization

The present thesis is organized in three chapters. The first chapter provides a broad introduction addressing literature review on *FMR1* gene, associated phenotypes and its role in female fertility. Chapter 2 is structured as a scientific paper submitted to Fertility and Sterility Journal. A broader discussion of results provided in the paper as well as future perspectives of this thesis work are detailed in Chapter 3. Annexes 1-3 contain supplementary data of chapter 2 results. Annex 4 include other scientific work focusing on subjects not directly linked to the main goal of this thesis.

References

American Society for Reprodutive Medicine. Diagnostic evaluation of the infertile female: A committee opinion. *Fertil Steril* 2015;**103**:44–55.

Bennett CE, Conway GS, MacPherson JN, Jacobs PA, Murray A. Intermediate sized CGG repeats are not a common cause of idiopathic premature ovarian failure. *Hum Reprod* 2010;**25**:1335–1338.

Biancalana V, Glaeser D, McQuaid S, Steinbach P. EMQN best practice guidelines for the molecular genetic testing and reporting of fragile X syndrome and other fragile X-associated disorders. *Eur J Hum Genet* 2015;**23**:417–425.

Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: Potential need and demand for infertility medical care. *Hum Reprod* 2007;**22**:1506–1512.

Canêdo M. Uso do hormônio luteinizante recombinante em ciclos de fertilização assistida. 2015, 16-52.

Eichler EE, Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeu SN, Richards CS, Ward PA, Nelson DL. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. *Nat Genet* 1994;**8**:88–94.

ESHRE. Failures (with some successes) of assisted reproduction and gamete donation programs. *Hum Reprod Update* 2013;**19**:354–365.

Eslami A, Farahmand K, Totonchi M, Madani T, Asadpour U, Zari Moradi S, Gourabi H, Mohseni-Meybodi A. FMR1 premutation: not only important in premature ovarian failure but also in diminished ovarian reserve. *Hum Fertil* 2016;**10**:1–6.

Gleicher N, Kushnir VA, Weghofer A, Barad DH. How the FMR1 gene became relevant to female fertility and reproductive medicine. *Front Genet* 2014;**5**:1–5.

Gleicher N, Yu Y, Himaya E, Barad DH, Weghofer A, Wu Y, Albertini DF, Wang VQ, Kushnir VA. Early decline in functional ovarian reserve in young women with low (CGGn < 26) FMR1 gene alleles. *Transl Res* 2015;**166**:502–507.

Hantash FM, Goos DM, Crossley B, Anderson B, Zhang K, Sun W, Strom CM. FMR1 premutation carrier frequency in patients undergoing routine population-based carrier screening: Insights into the prevalence of fragile X syndrome, fragile X-associated tremor/ataxia syndrome, and fragile X-associated primary ovarian insufficiency i. *Genet Med* 2011;**13**:39–45.

Hoyos LR, Thakur M. Fragile X premutation in women: recognizing the health challenges beyond primary ovarian insufficiency. *J Assist Reprod Genet* 2017;**34**:315–323.

Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet* 2000;**9**:901–908.

Kline JK, Kinney AM, Levin B, Brown SA, Hadd AG, Warburton D. Intermediate CGG repeat length at the FMR1 locus is not associated with hormonal indicators of ovarian age. *J North Am Menopause Soc* 2014;**21**:740–748.

Lubs H. A marker X chromosome. *Am J Hum Genet* 1969;**21**:231–244.

Maia N, Loureiro JR, Oliveira B, Marques I, Santos R, Jorge P, Martins S. Contraction of fully expanded FMR1 alleles to the normal range: predisposing haplotype or rare events? *J Hum Genet* 2016;**62**:1–7.

Man L, Lekovich J, Rosenwaks Z, Gerhardt J. Fragile X-Associated Diminished Ovarian Reserve and Primary Ovarian Insufficiency from Molecular Mechanisms to Clinical Manifestations. *Front Mol Neurosci* 2017;**10**:1–17.

Melmed S, Casanueva FF, Hoffman AR, Kleinberg DL, Montori VM, Schlechte JA, Wass JAH. Diagnosis and treatment of hyperprolactinemia: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011;**96**:273–288.

Mukherjee T, Copperman AB, Lapinski R, Sandler B, Bustillo M, Grunfeld L. An elevated day three follicle-stimulating hormone:luteinizing hormone ratio (FSH:LH) in the presence of a normal day 3 FSH predicts a poor response to controlled ovarian hyperstimulation. *Fertil Steril* 1996;**65**:588–593.

Murray A, Schoemaker MJ, Bennett CE, Ennis S, Macpherson JN, Jones M, Morris DH, Orr N, Ashworth A, Jacobs PA, Swerdlow AJ. Population-based estimates of the prevalence of FMR1 expansion mutations in women with early menopause and

primary ovarian insufficiency. *Genet Med* 2014;**16**:19–24.

Oostra BA, Willemsen R. A fragile balance: FMR1 expression levels. *Hum Mol Genet* 2003;**12**:R249-57.

Pastore LM, McMurry TL, Williams CD, Baker VL, Young SL. AMH in women with diminished ovarian reserve: potential differences by FMR1 CGG repeat level. *J Assist Reprod Genet* 2014;**31**:1295–1301.

Penagarikano O, Mulle JG, Warren ST. The Pathophysiology of Fragile X Syndrome. *Annu Rev Genomics Hum Genet* 2007;**8**:109–129.

Pfeifer S, Butts S, Dumesic D, Fossum G, Giudice L, Gracia C, Barbera A La, Odem R, Pisarska M, Rebar R, *et al.* Testing and interpreting measures of ovarian reserve: a committee opinion. *Fertil Steril* 2015;**103**:e9–e17.

Poon PMK, Chen QL, Zhong N, Lam S. S, Lai KYC, Wong CK, Pang CP. AGG interspersion analysis of the FMR1 CGG repeats in mental retardation of unspecific cause. *Clin Biochem* 2006;**39**:244–248.

Roudebush WE, Kivens WJ, Mattke JM. Biomarkers of ovarian reserve. *Biomark Insights* 2008; 1–23.

Ruth KS, Bennett CE, Schoemaker MJ, Weedon MN, Swerdlow AJ, Murray A. Length of FMR1 repeat alleles within the normal range does not substantially affect the risk of early menopause. *Hum Reprod* 2016;**31**:2396–2403.

Saldarriaga W, Tassone F, González-Teshima LY, Forero-Forero J V., Ayala-Zapata S, Hagerman R. Fragile X syndrome. *Colomb médica* 2014;**45**:190–198.

Schenkel LC, Schwartz C, Skinner C, Rodenhiser D, Ainsworth P, Pare G, Sadikovic B. Clinical Validation of Fragile X Syndrome Screening by DNA Methylation Array. *J Mol Diagnostics* 2016;**18**:834–841.

Schufreider A, McQueen DB, Lee SM, Allon R, Uhler ML, Davie J, Feinberg EC. Diminished ovarian reserve is not observed in infertility patients with high normal CGG repeats on the fragile X mental retardation 1 (FMR1) gene. *Hum Reprod* 2015;**30**:2686–2692.

Streuli I, Fraisse T, Ibecheole V, Moix I, Morris MA, Ziegler D de. Intermediate and premutation FMR1 alleles in women with occult primary ovarian insufficiency. *Fertil Steril* 2009;**92**:464–470.

Sullivan AK, Marcus M, Epstein MP, Allen EG, Anido AE, Paquin JJ, Yadav-Shah M, Sherman SL. Association of FMR1 repeat size with ovarian dysfunction. *Hum Reprod* 2005;**20**:402–412.

Sullivan S, Welt C, Sherman S. FMR1 and the Continuum of Primary Ovarian Insufficiency. *Semin Reprod Med* 2011;**29**:299–307.

Tural S, Tekcan A, Kara N, Elbistan M, Guven D, Tasdemir HA. FMR1 gene mutation screening by TP-PCR in patients with premature ovarian failure and fragile X. *Gynecol Endocrinol* 2014;**3590**:1–5.

Verissimo R, Silvestre M. Ovarian reserve markers and antral follicle count. *Acta Obstet e Ginecol Port* 2016;**10**:308–316.

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 1991;**65**:905–914.

Voorhuis M, Onland-Moret NC, Fauser BCJM, Ploos Van Amstel HK, Schouw YT Van Der, Broekmans FJ. The association of CGG repeats in the FMR1 gene and timing of natural menopause. *Hum Reprod* 2013;**28**:496–501.

Voorhuis M, Onland-Moret NC, Janse F, Ploos Van Amstel HK, Goverde AJ, Lambalk CB, Laven JSE, Schouw YT Van Der, Broekmans FJM, Fauser BCJM. The significance of fragile X mental retardation gene 1 CGG repeat sizes in the normal and intermediate range in women with primary ovarian insufficiency. *Hum Reprod* 2014;**29**:1585–1593.

Wang JY, Trivedi AM, Carrillo NR, Yang J, Schneider A, Giulivi C, Adams P, Tassone F, Kim K, Rivera SM, Lubarr N, Wu CY, Irwin RW, Brintn RD, Olicheney JM, Rogawski MA, Hagerman RJ. Open-Label Allopregnanolone Treatment of Men with Fragile X-Associated Tremor/Ataxia Syndrome. *Neurotherapeutics* 2017;1–11.

Willemsen R, Levenga J, Oostra B. CGG repeat in the FMR1 gene: Size matters. *Clin Genet* 2011;**80**:214–225.

Wittenberger MD, Hagerman RJ, Sherman SL, McConkie-Rosell A, Welt CK, Rebar RW, Corrigan EC, Simpson JL, Nelson LM. The FMR1 premutation and reproduction. *Fertil Steril* 2007;**87**:456–465.

WEB: (https://www.sns.gov.pt/cidadao/banco-publico-de-gametas-2/ assessed on 2017/10/19)

“The influence of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function”

CHAPTER 2.

The impact of the *FMR1* gene AGG interspersions in the female reproductive function

“The influence of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function”

Chapter 2. The impact of the *FMR1* gene AGG interspersions in the female reproductive function

Abstract

Objectives – Albeit still controversial an association between *FMR1* gene CGG allele size below 26 and a low ovarian reserve has been reported. Furthermore, to the best of our knowledge, the effect of the AGG interspersions of those normal-sized alleles and their impact on the ovarian function was never tested. The aim of this study was to assess the influence of the *FMR1* gene in the female reproductive function and ovarian reserve by establishing a measure of the allele complexity, combining GGC number and AGG pattern.

Design - Young and healthy females.

Setting – Centro de Genética Médica Doutor Jacinto Magalhães – Centro Hospitalar do Porto.

Patients – Fifty potential oocyte donors aged between 18 and 33 years and corresponding results of hormonal, genetic, and reproductive studies previously obtained as part of routine donor selection protocol. *FMR1* allele sizes were organized by hitherto defined classes of sub-genotypes.

Results - The multivariate analysis showed that the hormonal levels are not sufficient to discriminate samples according to the *FMR1* sub-genotypes. We hypothesized that besides the length of the CGG tract also the AGG interspersion pattern could be implicated in the female reproductive function. A mathematical formula was further established to score the allelic complexity. At an allelic score value of 122, two distinct groups emerged: the equivalent, composed mainly by samples carrying alleles in the *normal* sub-genotype and the opposite, where most of the samples have a low/*normal* allele.

Conclusion – Overall, this study confirms the importance of *FMR1* gene CGG tract in the female reproductive function and highlights that the stability of those alleles, determined by the AGG number and pattern, is also an important factor. To assess *FMR1* influence in the female ovarian function we suggest the additional determination of the AGG number and

pattern as a measure of allelic complexity. To establish whether the AGG pattern can determine the ovarian response in IVF treatments or oocyte donor selection and be meaningful to changes in the clinical practice, studies in larger and at-risk cohorts (e.g. infertile couples) are needed.

Keywords – CGG repeat number; reproductive function; AGG interspersions; ovarian reserve markers; oocyte donors

Funding

Unit for Multidisciplinary Research in Biomedicine (UMIB), ICBAS/UP, is supported by national funds: FCT Strategic Project UID/Multi/00215/2013.

Compliance with ethical standards

Conflicts of Interest: All authors have read the journal's policy on disclosure of potential conflicts of interest and have none to declare.

This study has been approved by the medical ethical committee of the Centro Hospitalar do Porto (CHP, E.P.E.) - REF 2017-060 (054-DEFI/052-CES) and is part of Rodrigues B, Master's Degree research project.

1. INTRODUCTION

Fragile X-associated primary ovarian insufficiency (FXPOI; OMIM#311360) is characterized by an ovarian reduced function and accounts for about 5% of all cases of primary ovarian insufficiency (1). FXPOI can cause early menopause, irregular cycles, elevated follicle stimulating hormone (FSH) levels and ultimately lead to infertility (2–4). Fragile mental retardation-1 gene (*FMR1*) mutations in the premutation range (CGG number between 55 and 200) augment the risk of FXPOI, in about 15 to 20%, when compared to an healthy population (3,5). Shortened follicular phases and changes in several ovarian function biochemical markers, such as raised FSH and low anti-Mullerian hormone (AMH) levels, have also been described in *FMR1* premutation carriers (2). Bellow 54 CGG repeats, alleles are classified as normal and usually carry two or more AGG interruptions which is assumed to give stability, hampering the expansion to pathogenic ranges (6,7). Among the normal range, alleles with 45 to 54 CGGs are known as intermediate or gray zone, due to the likelihood of expansion (in two or three generations) (8,9). When this expansion is in full mutation range the simultaneous hypermethylation of the repetitive region and the *FMR1* promoter, as well as gene silencing, occur. *FMR1* transcriptional inactivation and the consequent absence of the coded protein, FMRP, is the cause of the intellectual disability in patients with fragile X syndrome [FXS; OMIM#300624] (9,10). FXS includes also learning problems, autistic behaviour and typical physical features, such as long and narrow face and protruding ears (9,11). FMRP plays a role in the development of connections at synapses (10,12). Although arising from mutations in the same gene, different mechanisms lead to FXS and FXPOI (2). The FMRP implication in the ovarian function, remains to be unravelled, although it is established that the premutation triggers the overproduction of *FMR1* mRNA that leads to a process of RNA toxicity (5,13,14).

Recent reports of phenotypes that overlap to those seen in females with a premutation or are exclusive to normal/intermediate size carriers has grown interest in this latter range of alleles. Conclusions however, are controversial, not only regarding influence of *FMR1* in ovarian reserve but also on definition of a "new" normal repeat range applicable exclusively to the female reproductive function. Gleicher, and co-workers (2015), published several studies showing the influence of the CGG repeat number in the ovarian reserve. In another study, an AMH decline, suggestive of diminished ovarian reserve, was observed to occur more rapidly in oocyte donor candidates carrying one allele with a CGG number below 26

(15). In a cohort of infertile women, lower AMH levels were associated with presence of one allele with less than 28 and the other with more than 33 repeats (16). Spitzer *et al.*, on the contrary, has found no such associations when studied a similar but larger cohort (17) .

To better understand the impact of the *FMR1* normal range genotypes and sub-genotypes we investigated if the AGG interspersions also play a role in the female ovarian function. Since hormonal levels and antral follicle counts seem to be associated with those "new-normal" *FMR1* genotypes we studied the effect of AGG number and pattern in a pilot study of 50 potential oocyte donors. Overall, this study confirms the association of the *FMR1* CGG repetitive region in the female ovarian function and suggests that the stability of the alleles – determined by AGG number and pattern – is also a determining factor for the ovarian response success.

2. MATERIAL AND METHODS

2.1. Pilot study

The Public Bank of Gametes is a service provided by the National Health Service responsible for the recruitment and selection of oocytes and sperm donors. Gametes resulting from volunteer donations are then used in Medically Assisted Procreation techniques. The collection and preservation of the gametes is carried out in public specialized Centers, namely Centro Materno Infantil do Norte (CMIN) of Centro Hospitalar do Porto (CHP, E.P.E.). Herein, results from 50 potential oocyte donors, aged between 18 and 33 years (mean 25,4 ± 3,9), obtained as routine protocol for donor selection were used, namely hormonal (Estradiol, FSH, LH and Prolactin levels), reproductive (number of antral follicles, tAFC) and genetic (*FMR1* gene allele sizing) (annex-1).

Further molecular studies were performed on stored genomic DNA previously extracted from peripheral blood using ArchivePure DNA Cell/Blood Kit (5 Prime® GmbH, Hamburg, Germany).

2.2. Triplet-primed polymerase chain reaction (TP-PCR)

AGG interspersion pattern was determined by TP-PCR, using primers and components described on Table 1 and 2, respectively. The amplification program included 10 min incubation at 98°C, and 48 cycles of denaturation at 98°C for 1 min, annealing at 60°C for 1 min and extension at 68°C for 6 min, the final extension of 10 min was performed at 68°C. Products were resolved on ABI PRISM® 3130*xl* Genetic Analyzer (Applied Biosystems™, Foster City, CA, USA), AGG pattern was analyzed using GeneMapper® software version 4.0 (Applied Biosystems™) and allele sizing were obtained by comparison with the 500 ROX™ size standard (Gene Scan™, Warrington, UK) (annex-2).

Table 1. Sequence of primers used in TP-PCR.

Primer	**Sequence**
g.FMR1_CGG_F	5'-GACGGAGGCGCCGCTGCCAGG-3'
g.FMR1_CGG_R* (6-FAM labeled)	5'-TACGCATCCCAGTTTGAGACG-3'
g.FMR1_P4_CGG	5'-TACGCATCCCAGTTTGAGACGCGGCGGCGGCGGCG-3'

Table 2. Triplet-primed PCR components.

Reagents	
Multiplex PCR Master Mix (Promega®, Madison, WI, USA)	1X
Betaine (Sigma-Aldrich®, St. Louis, Missouri, EUA)	1,1 M
DMSO (Sigma-Aldrich®)	11%
g.FMR1_TP-PCR_F	5 pmol
g.FMR1_ TP-PCR _R [6-FAM labeled]	5 pmol
g.FMR1_P4_CGG	2,5 pmol
DNA	150 ng

2.3. Sanger Sequencing

To validate TP-PCR results, samples were sequenced using standard conditions (Tables 3 and 4). Symmetric PCR products were purified using the enzymatic PCR clean up technology -Illustra™ ExoProStar™ 1-step (GE Healthcare Life Sciences, Little Chalfont, Buckinghamshire, UK) according to the manufacturer's instructions. Asymmetric amplification (Table 5) consisted of 6 min incubation at 95°C and 35 cycles of denaturation at 96°C for 10 sec, annealing at 50°C for 10 sec and extension at 60°C for 4 min, the final extension of 10 min was performed at 60°C. Following gel filtration spin columns cleanup (DyeEx®96 kit, QIAGEN GmbH, Hilden, Germany), products were electrophoresed on an ABI PRISM® 3130*xl* Genetic Analyzer (Applied Biosystems™) and analyzed using SeqScape® software version 2.5 (Applied Biosystems™) (*FMR1* cDNA reference sequence NM_002024).

Table 3. Sequence of primers used in sequencing PCR reactions.

Primer	**Sequence**
g.FMR1_CGG_F	5'-TGACGTGGTTTCAGTGTTTAC-3'
g.FMR1_CGG_R	5'-AGCCAAGTACCTTGTAGAAAGCG-3'

Table 4. Symmetric PCR components.

Reagents	
Distilled water	5,8 µL
Accu Taq LA Buffer (Sigma-Aldrich®)	1X
Betaine (Sigma-Aldrich®)	1,5 M
dATP/dCTG/dTTP/dGTP (Bioline, London, UK)	0,16 mM
DMSO (Sigma-Aldrich®)	6,3%
g.FMR1_CGG_F	10 pmol
g.FMR1_CGG_R	10 pmol
DNA	150 ng
Accu Taq LA DNA polymerase (Sigma-Aldrich®)	2 U

Table 5. Asymmetric PCR components.

Reagents	
Purified symmetric PCR product	7 µL
BigDye® Terminador version 3.1 Cycle Sequencing Mix (Applied Biosystems™)	3 µL
DMSO (Sigma-Aldrich®)	3,4%
Betaine (Sigma-Aldrich®)	1 M
g.*FMR1*_CGG_F/R	5 pmol

2.4. Restriction-fragment length analysis (RFLA)

Samples showing a complex AGG pattern were further analyzed by RFLA after *Mnl*I digestion of amplification products using the primers described on table 3. The incubation was performed at 37°C overnight using the total PCR volume (20µl) in 1X NEBuffer 4 (New England BioLabs®, Ipswich, Massachusetts, EUA) and 0,015 U of *Mnl*I endonuclease (New England BioLabs®). Products were analyzed by agarose gel electrophoresis for sizing, by comparison with GeneRuler™ 100 bp Plus (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Fragment images were obtained by the Fujifilm Luminescent Image Analyzer LAS-3000 v2.2 (Fujifilm, Tokyo, Japan). *Mnl*I recognizes the sequence 3'…GGAG$(N)_6$…5', cutting the amplified sequence in a fragment of 141bp (in case of alleles with 29 pure CGGs) and fragments of 37, 33, 25, 20 and 13 bp, irrespectively of the CGG repeat number. In the presence of 2 AGG interruptions a 29 CGG allele will yield fragments with 58, 53 and 30bp. The expected size of each allele was calculated using the NEBcutter® version 2.0 simulator (New England BioLabs®).

2.5. Statistical analyses

Analyses were carried out using *FMR1* genotypes and hormonal levels data of all 50 potential oocyte donors (100 alleles) irrespective of the possible inclusion as donors.

Several sets of analyses were carried out:

a) Association between CGG repeat number and hormonal levels.

Principal component analysis was used to arrange the donors in a multi-dimensional space defined by the donor's hormone levels and classified based on the combination of repeat number in both alleles, using the Canoco for Windows, version 4.5.

b) Allelic complexity score values were calculated for both alleles using AGG number and pattern breakpoints as:

$$Allelic\ Score = \left(\sum_{i=1}^{n} R_i \times c\ 4^{i-1}\right) + (R_{n+1} \times\ 4^n)$$

R_i: number of CGG repeats before the first AGG interruption of order i (counting from 5' to 3')

n: total number of AGG interspersions.

R_{n+1}: number of CGG repeats after the last AGG interruption.

This mathematical formula simultaneously combines the allele size and the AGG interspersion number and pattern. Thus, the allelic score reflects the structure and complexity of the AGG interspersion pattern.

c) Correlation between AGG interspersion number and pattern, the antral follicle count and hormonal levels.

Parametric statistics and multiple linear regression were calculated using Minitab® statistical software, version 16 (Minitab® Inc., State College, USA). A significance level of 0.05 was considered for all the analyses.

3. RESULTS & DISCUSSION

Summary of *FMR1* genotyping results in shown in Table 6. Data are divided according to the *FMR1* sub-genotypes previously defined by others (18). At least one sample was identified in each of the six allelic categories. The smallest allele has 15 CGGs and the largest 48 repeats. As expected, the number of CGG repeats is highly variable with the most frequent allele being that with 29 repeats (n=37), followed by alleles with 19 (n=15) and 28 (n=13) CGGs (19). In our cohort, eleven homoallelic samples (22%) were identified, of which ten belong to the normal group and one to the low/low group (Table 6). One sample belonging to the high/high group has one allele in the intermediate range (D50 - annex 1). Based only on allele sizing and on the *FMR1* coding DNA Reference Sequence, GenBank NM_002024.5, the recommended nomenclature to describe the alleles present in that sample is c.-128_-126[40]; [48]. Carriers of intermediate size alleles are excluded from donation, as alleles of this size can expand to full mutation in two generations.

Table 6. Summary of *FMR1* genotyping data in the cohort of 50 potential oocyte donor samples.

Alleles	**CGG repeat number**	***FMR1* sub-genotypes classification**	N	Homoallelic samples (%)
A1	**26<CGG<34**	***normal***	23	44
A2				
A1	**8<CGG<25**	***low/normal***	17	0
A2	**26<CGG<34**			
A1	**8<CGG<26**	***low/high***	4	0
A2	**35<CGG<55**			
A1	**26<CGG<34**	***normal/high***	3	0
A2	**35<CGG<55**			
A1	**8<CGG<26**	***low/low***	2	50
A2				
A1	**34<CGG <55**	***high/high***	1	0
A2				

A1- Allele 1 (smallest in size); A2- Allele 2 (largest in size); N – number of samples.

Table 7. Summary of donor characteristics and biochemical data.

N=50		**Reference values***
Age	25,4 ± 3,9	18-33 years
Number of antral follicles	8 ± 4,4	> 6
FSH	5,8 ± 1,7	<10 mIU/mL
LH	5,9 ± 5,1	<10 mUI/mL
Estradiol	40,6 ± 24,8	< 60 pg/mL
Prolactin	14,1 ± 6,4	< 25 ng/mL

*Adapted from (20,21). Values are expressed in mean ± SD.

All samples presented FSH levels within reference values except for D49. D9 revealed elevated prolactin plasma level and D25, D26, D30 and D33, showed estradiol levels above normal. Above normal were also the LH levels of D42, D30 and D23. D2, D32 and D25 disclosed antral follicles below 6, being the last two in the *low/normal FMR1* sub-genotype and D2 in the *normal FMR1* sub-genotype (annex 3).

An exploratory approach to identify an association between the CGG number and hormonal levels was performed. The six categories of *FMR1* sub-genotypes were defined as species and each biochemical parameter (FSH, LH, estradiol and prolactin levels) as supplementary explanatory variants. The values were centered and standardized within Canoco but were not transformed, yielding a biplot correlation. In standardization, all variables were considered equally important regardless of their variability.

The biplot in Figure 1 depicts the association between samples (grouped according with CGG repeat number and corresponding *FMR1* sub-genotype) and hormonal levels. The distribution of the samples is determined mainly by estradiol (first axis) and prolactin (second axis). However, the hormonal profile is not able to separate the groups defined by the CGG number. Among the hormones selected for the current study the estradiol alone explained 93,2% of the total variance.

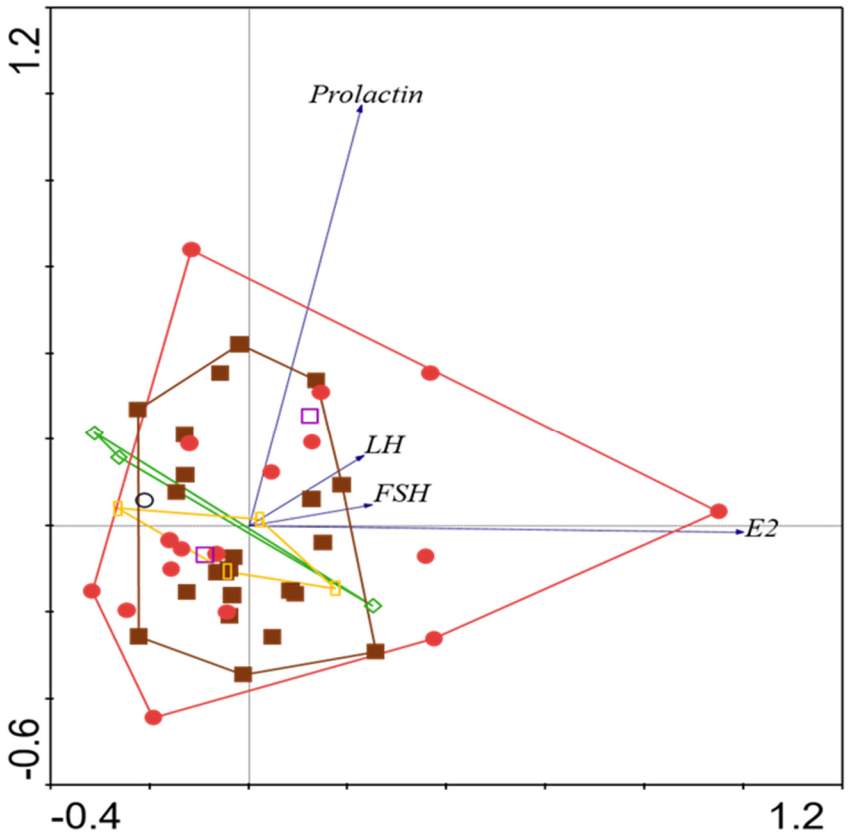


Figure 1. Biplot of *FMR1* sub-genotypes biochemical results (for the 50 potential oocyte donors).

Multivariate analysis to project the association between the CGG repeat number and the different species, hindered the individualization of the samples classified by *FMR1* sub-genotype. The biplot shows that the hormonal levels are not sufficient to discriminate samples according to the *FMR1* sub-genotypes which may be due to the large variability of the hormonal levels observed among the different samples or to the fact that the CGG repeat number and the hormonal levels are independent variables.

The absence of an association prompted further investigation of other variables besides the length of the CGG tract. Napierala, and co-workers (2005), have demonstrated that the presence of AGG interruptions in the *FMR1* repetitive region can influence FXTAS clinical outcome in male premutation carriers, by weakening the *FMR1* mRNA structure (22). We hypothesized that by a similar mechanism, both the length of the CGG tract and the pattern

of the AGG interspersions, could play a role in the female reproductive function. Thus, a mathematical formula was stablished to score *FMR1* alleles according to the CGG number and AGG number and pattern. The score was named allelic complexity value. Using this approach not only the size but also the stability - as determined by the AGG number and pattern - were considered. A clear correlation could not be found between the allelic scores when they are plotted one against the other (Figure 2).

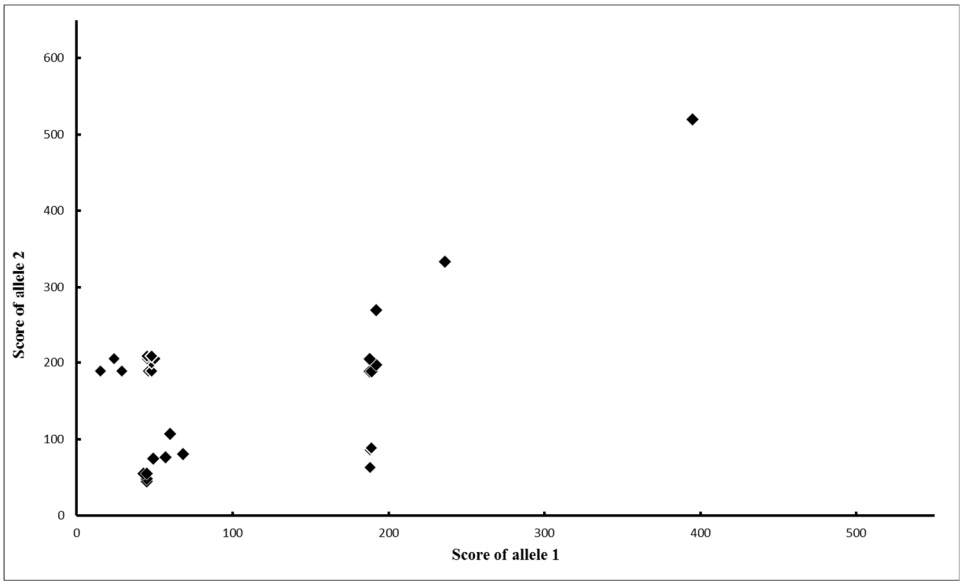


Figure 2. Allelic complexity score value of each sample.

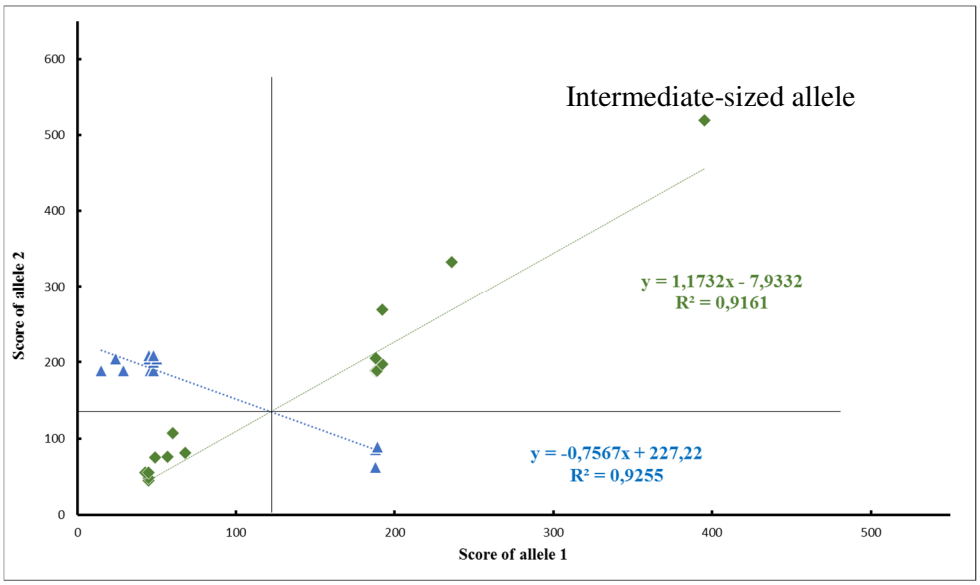


Figure 3. Allelic complexity score value based on the allele size and AGG interruption number and pattern. Samples carrying alleles of equivalent AGG pattern are represented with green lozenges and those with an opposite pattern with blue triangles.

Nevertheless, when the graph is divided in quadrants cantered at an allelic score of 122 (Figure 3), two distinct patterns emerge: one where a similar AGG interspersion pattern can be observed for both alleles (**equivalent group**) and a second in which both alleles present a different AGG interspersion pattern (**opposite group**). The equivalent group includes samples where both alleles share similar number of AGG interruptions (e.g. one or two). Sample D45 containing three AGGs was removed from this analysis (see annex-3 for supplementary data).

Table 8. Samples distributed according to allelic complexity score groups and *FMR1* CGG sub-genotypes.

***FMR1* sub-genotype**	**N Equivalent group**	**%**	**N Opposite group**	**%**	***p* value**
normal	**15**	**31**	8	16	0,0615
low/normal	5	10	**12**	**25**	**0,0274**
low/hight	1	2	3	6	---
normal/high	1	2	1	2	---
low/low	2	4	0	0	---
high/high	1	2	0	0	---
TOTAL (N=49)	25	51	24	49	

N- number of samples.

As shown in Table 8, equivalent group is mainly composed by samples carrying alleles in the normal *FMR1* sub-genotype. In the opposite group, the most frequent samples show an *FMR1 low/normal* sub-genotype (p=0.027).

To strengthen our hypothesis that AGG can influence the ovarian function a correlation between the number of antral follicles and the hormonal levels was attained in the equivalent group, using Minitab® 16 statistical software.

This process was initiated by a stepwise regression performed with all quantified hormones. A positive correlation between number of antral follicles, and prolactin and LH levels was obtained. A multiple regression for the number of antral folicules using prolactin

and LH as descriptors was performed to establish a statistically significant model (p=0.030) that predicted the number of antral follicles based on the LH and prolactin levels (Figure 4):

$$\mathbf{tAFC = 3.62+0.523\times LH+0.210\times PRL}$$

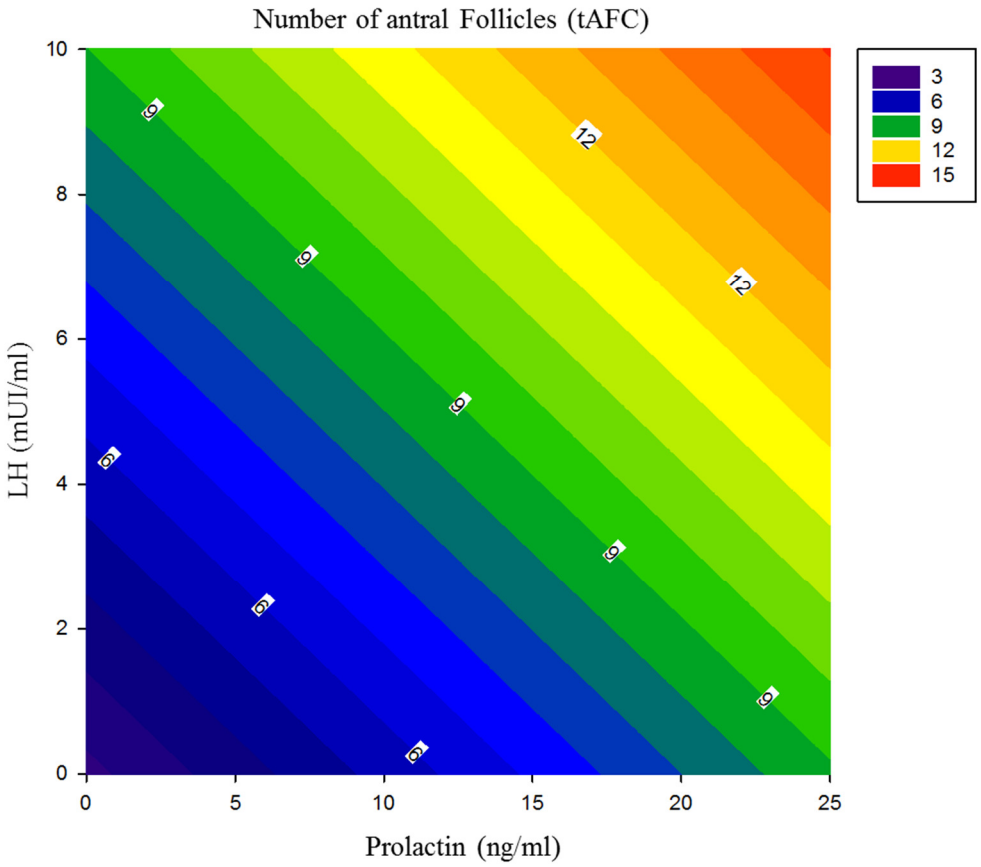


Figure 4. Isobologram showing the visual representation of the mathematical formula. Axes show the LH and Prolactin levels. Each color is associated with a specific number of antral follicles. A low number of follicles is represented by the blue color and the maximum in red. tAFC - Total Antral Follicle

According to our model, it is possible to determine the largest number of antral follicles produced combining the levels of prolactin and LH in the group of females that show an equivalent AGG pattern. Interestingly, most of samples of the opposite group carry alleles in the *low/normal FMR1* sub-genotype (one allele with less than 26 CGGs) (p=0.027). This observation is in line with previous publications that suggest a negative influence of a

low *FMR1* CGG number on the ovarian reserve, notwithstanding other elements seem to be contributing to this correlation (23).

In conclusion, the present study shows that the *FMR1* CGG repetitive region has an impact on the female reproductive function and that AGG interspersions can be used to assess the ovarian response success. Our results indicate that in the presence of an equivalent AGG pattern, combining the levels of prolactin and LH validly measures the antral follicles number. However, this mathematical model, as well as the hypothesis of underlying association between *FMR1* alleles with CGG<26 with low ovarian reserve and poor ovarian response has to be confirmed in future studies. Researchers interested in assessing *FMR1* influence in female ovarian function should consider interpreting the AGG number and pattern as one subscale for measuring allelic complexity, beyond allele size. Furthermore, a comprehensive evaluation of *FMR1* CGG tract impact should be accompanied by the assessment of clinical-relevant and female related factors such as the pattern of X-chromosome inactivation. To establish whether the AGG pattern can determine the ovarian response in IVF treatments or oocyte donor selection and be meaningful to changes in clinical practice, studies in larger and at-risk cohorts (e.g. infertile couples) are needed.

References

1. Murray A, Schoemaker MJ, Bennett CE, Ennis S, Macpherson JN, Jones M, et al. Population-based estimates of the prevalence of FMR1 expansion mutations in women with early menopause and primary ovarian insufficiency. Genet Med 2014;16(1):19–24.

2. Streuli I, Fraisse T, Ibecheole V, Moix I, Morris MA, de Ziegler D. Intermediate and premutation FMR1 alleles in women with occult primary ovarian insufficiency. Fertil Steril 2009;92(2):464–70.

3. Tural S, Tekcan A, Kara N, Elbistan M, Guven D, Tasdemir HA. FMR1 gene mutation screening by TP-PCR in patients with premature ovarian failure and fragile X. Gynecol Endocrinol 2014;3590(3):1–5.

4. Eslami A, Farahmand K, Totonchi M, Madani T, Asadpour U, Zari Moradi S, et al. FMR1 premutation: not only important in premature ovarian failure but also in diminished ovarian reserve. Hum Fertil 2016;10(2):1–6.

5. Sullivan AK, Marcus M, Epstein MP, Allen EG, Anido AE, Paquin JJ, et al. Association of FMR1 repeat size with ovarian dysfunction. Hum Reprod 2005;20(2):402–12.

6. Oostra BA, Willemsen R. A fragile balance: FMR1 expression levels. Hum Mol Genet 2003;12(2):R249-57.

7. Maia N, Loureiro JR, Oliveira B, Marques I, Santos R, Jorge P, et al. Contraction of fully expanded FMR1 alleles to the normal range: predisposing haplotype or rare events? J Hum Genet 2016;62(2):1–7.

8. Poon PMK, Chen QL, Zhong N, Lam S. S, Lai KYC, Wong CK, et al. AGG interspersion analysis of the FMR1 CGG repeats in mental retardation of unspecific cause. Clin Biochem 2006;39(3):244–8.

9. Schenkel LC, Schwartz C, Skinner C, Rodenhiser D, Ainsworth P, Pare G, et al. Clinical Validation of Fragile X Syndrome Screening by DNA Methylation Array. J Mol Diagnostics 2016;18(6):834–41.

10. Hoyos LR, Thakur M. Fragile X premutation in women: recognizing the health challenges beyond primary ovarian insufficiency. J Assist Reprod Genet 2017;34(3):315–23.

11. Saldarriaga W, Tassone F, González-Teshima LY, Forero-Forero J V., Ayala-Zapata S, Hagerman R. Fragile X syndrome. Colomb médica 2014;45(4):190–8.

12. Wittenberger MD, Hagerman RJ, Sherman SL, McConkie-Rosell A, Welt CK, Rebar RW, et al. The FMR1 premutation and reproduction. Fertil Steril 2007;87(3):456–65.

13. Willemsen R, Levenga J, Oostra B. CGG repeat in the FMR1 gene: Size matters. Clin Genet 2011;80(3):214–25.

14. Ruth KS, Bennett CE, Schoemaker MJ, Weedon MN, Swerdlow AJ, Murray A. Length of FMR1 repeat alleles within the normal range does not substantially affect the risk of early menopause. Hum Reprod 2016;31(10):2396–403.

15. Gleicher N, Yu Y, Himaya E, Barad DH, Weghofer A, Wu Y, et al. Early decline in functional ovarian reserve in young women with low (CGGn < 26) FMR1 gene alleles. Transl Res 2015;166(5):502–7.

16. Gleicher N, Weghofer A, Oktay K, Barad DH. Revelance of triple CGG repeats in the FMR1 gene to ovarian reserve. Acta Obstet Gynecol Scand 2009;88(9):1024–30.

17. Spitzer LT, Johnstone EB, Huddleston HG, Cedars LM, Davis G, Fujimoto V. FMR1 Repeats and Ovarian Reserve: CGG Repeat Number does not Influence Antral Follicle Count. J Fertil Vitr 2012;2(3):10–3.

18. Gleicher N, Kushnir VA, Weghofer A, Barad DH. How the FMR1 gene became relevant to female fertility and reproductive medicine. Front Genet 2014;5(284):1–5.

19. Jorge P, Oliveira B, Marques I, Santos R. Development and validation of a multiplex-PCR assay for X-linked intellectual disability. BMC Med Genet 2013;14(80):11.

20. Melmed S, Casanueva FF, Hoffman AR, Kleinberg DL, Montori VM, Schlechte JA,

et al. Diagnosis and treatment of hyperprolactinemia: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2011;96(2):273–88.

21. Pfeifer S, Butts S, Dumesic D, Fossum G, Giudice L, Gracia C, et al. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril 2015;103(3):e9–17.

22. Napierala M, Michalowski D, de Mezer M, Krzyzosiak WJ. Facile FMR1 mRNA structure regulation by interruptions in CGG repeats. Nucleic Acids Res 2005;33(2):451–63.

23. Kline JK, Kinney AM, Levin B, Brown SA, Hadd AG, Warburton D. Intermediate CGG repeat length at the FMR1 locus is not associated with hormonal indicators of ovarian age. J North Am Menopause Soc 2014;21(7):740–8.

CHAPTER 3.

Discussion and future perspectives

“The influence of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function”

Chapter 3. Discussion and future perspectives

The implication of the *FMR1* gene premutation, 55<CGG<200, in the female reproductive function has been extensively studied and reported (Man *et al.*, 2017). Still under debate is the correlation of the "normal" CGG repeat range with the diminished ovarian reserve and other ovarian dysfunction (Bretherick *et al.*, 2005; Bodega *et al.*, 2006; Gleicher *et al.*, 2009 and Schufreider *et al.*, 2015; Pastore *et al.*, 2017).

This controversy led us to test whether an association between a normal number of CGG repeats and hormonal levels could be found, in normal and healthy females. A projection of the association between the CGG repeat values and the hormonal levels, by multivariate analysis, was performed, considering the *FMR1* sub-genotypes previously defined. The hormonal levels associated with the different samples were not sufficient to discriminate the sub-genotypes, indicating that the individualization of the samples classified by sub-genotype was not possible. An explanation is the large variability of the hormonal levels or otherwise, CGG repeat number and the hormone levels are independent. AMH seems to have the best discriminatory potential as the basal value remains constant inter- and intra-cycle, allowing a reliable evaluation of the ovarian reserve (Gleicher *et al.*, 2015; Pfeifer *et al.*, 2015; Verissimo and Silvestre, 2016). As AMH has been recently reported as the *gold standard* to predict the ovarian reserve it should be incorporated in future studies.

AGG interspersions present in the CGG tract have long been considered as stabilizers. Their presence in premutated alleles seems to reduce the risk of expansion upon maternal transmission (Maia *et al.*, 2016). Thus, we decided to analyze the impact of the AGG number and pattern in the ovarian function resorting to a mathematical formula that determines the allelic score, taking into account allele size and AGG number and pattern. This allelic score reflects the structure and complexity of both alleles of each sample. At an allelic score of 122, two distinct groups emerged: the equivalent composed mainly of samples carrying alleles in the normal *FMR1* sub-genotype and the opposite, where most of the samples have an *FMR1 low/normal* sub-genotype (p=0.027). The two groups were individually studied to correlate the number of antral follicles and the hormonal levels. In the equivalent group the number of antral follicles can be predicted using LH and prolactin levels. The opposite group, showing no such correlation, consists mostly of females with one allele above the

normal range and another allele below the normal range (*low/normal*). Interestingly, this particular sub-genotype has been associated with a diminished ovarian reserve pattern (Gleicher *et al.*, 2015).

How the stability of the *FMR1* alleles can be influencing the ovarian function in females is difficult to determine. Another important aspect that needs to be assessed particularly in females of the opposite group is the pattern of the X-chromosome inactivation. In the presence of a skewed inactivation pattern, the active allele can be contributing differently depending on the AGG number and pattern. One can speculate that a less stable allele could lead to mRNA accumulation. Then, the resulting toxic effect (either by an RNA gain-of-function or diminished FMRP synthesis) could cause early follicle atresia and ultimately a diminished ovarian reserve/function. Nevertheless, our results highlight the importance of the AGG triplets in the ovarian function and suggest that the determination of allele size alone, can provide partial results and mislead conclusions.

Data presented here confirm adverse *FMR1* effect in females carrying low CGG repeat number and suggest that AGG pattern can be influencing this negative outcome. If further studies confirm data reported herein, we suggest that *FMR1* testing in oocyte donor candidates should include AGG pattern analysis as an additional tool of selection.

Several weaknesses of this pilot study were drawn in this section. Possibly the most important being the small size of our cohort. On the contrary, the statistical power is noteworthy. These results are in line with prior publications and confirm the increasingly significant influence of the *FMR1* gene on female reproductive function, that can ultimately lead to infertility and add the AGG interspersions to the equation of fertility treatments success (Kushnir *et al.*, 2014).

References

Bodega B, Bione S, Dalprà L, Toniolo D, Ornaghi F, Vegetti W, Ginelli E, Marozzi A. Influence of intermediate and uninterrupted FMR1 CGG expansions in premature ovarian failure manifestation. *Hum Reprod* 2006;**21**:952–957.

Bretherick KL, Fluker MR, Robinson WP. FMR1 repeat sizes in the gray zone and high end of the normal range are associated with premature ovarian failure. *Hum Genet* 2005;**117**:376–382.

Gleicher N, Weghofer A, Oktay K, Barad DH. Revelance of triple CGG repeats in the FMR1 gene to ovarian reserve. *Acta Obstet Gynecol Scand* [Internet] 2009;**88**:1024–1030.

Gleicher N, Yu Y, Himaya E, Barad DH, Weghofer A, Wu Y, Albertini DF, Wang VQ, Kushnir VA. Early decline in functional ovarian reserve in young women with low (CGGn < 26) FMR1 gene alleles. *Transl Res* 2015;**166**:502–507. Elsevier Inc.

Kushnir VA, Yu Y, Barad DH, Weghofer A, Himaya E, Lee HJ, Wu YG, Shohat-Tal A, Lazzaroni-Tealdi E, Gleicher N. Utilizing FMR1 gene mutations as predictors of treatment success in human in vitro fertilization. *PLoS One* 2014;1-**9**.

Maia N, Loureiro JR, Oliveira B, Marques I, Santos R, Jorge P, Martins S. Contraction of fully expanded FMR1 alleles to the normal range: predisposing haplotype or rare events? *J Hum Genet* 2016;**62**:1–7.

Man L, Lekovich J, Rosenwaks Z, Gerhardt J. Fragile X-Associated Diminished Ovarian Reserve and Primary Ovarian Insufficiency from Molecular Mechanisms to Clinical Manifestations. *Front Mol Neurosci* 2017;**10**:1–17.

Pastore LM, Young SL, Manichaikul A, Baker VL, Wang XQ, Finkelstein JS. Distribution of the FMR1 gene in females by race/ethnicity: women with diminished ovarian reserve versus women with normal fertility (SWAN study). *Fertil Steril* 2017;**107**:205–211.

Pfeifer S, Butts S, Dumesic D, Fossum G, Giudice L, Gracia C, Barbera A La, Odem R, Pisarska M, Rebar R, *et al.* Testing and interpreting measures of ovarian reserve: a committee opinion. *Fertil Steril* 2015;**103**:e9–e17.

Schufreider A, McQueen DB, Lee SM, Allon R, Uhler ML, Davie J, Feinberg EC. Diminished ovarian reserve is not observed in infertility patients with high normal CGG repeats on the fragile X mental retardation 1 (FMR1) gene. *Hum Reprod* 2015;**30**:2686–2692.

Verissimo R, Silvestre M. Ovarian reserve markers and antral follicle count. *Acta Obstet e Ginecol Port* 2016;**10**:308–316.

ANNEXES

“The influence of the Fragile Mental Retardation-1 (*FMR1*) gene CGG repetitive region in the female reproductive function”

Annex 1. Summary of data regarding potential oocyte donors.

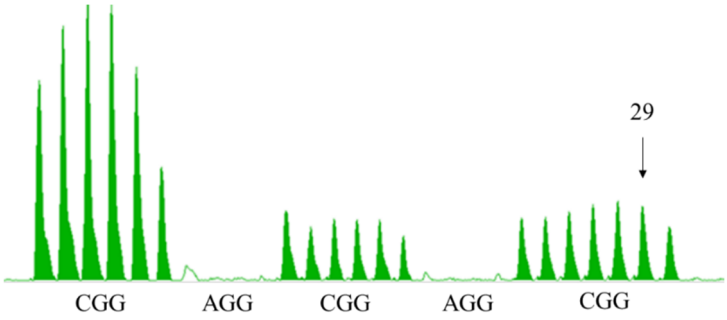
								Allele size	
Samples	**Age (y)**	**tCAF**	**FSH (mUI/mL)**	**LH (mUI/mL)**	**FSH/LH (mUI/mL)**	**E2 (pg/mL)**	**PRL (ng/mL)**	**A1**	**A2**
D01	20	10	5,4	4,8	1,1	39,9	4,0	**28**	**29**
D02	30	5	6,2	5,4	1,1	55,5	24,2	**29**	**29**
D03	30	10	5,0	6,7	0,7	57,7	13,6	**29**	**29**
D04	22	7	6,3	5,6	1,1	15,8	5,0	**28**	**29**
D05	28	12	8,9	6,5	1,4	51,5	9,8	**29**	**34**
D06	28	9	5,2	5,2	1,0	33,7	10,3	**29**	**31**
D07	19	8	4,7	3,4	1,4	37,1	9,1	**29**	**32**
D08	25	9	7,0	4,5	1,6	70,2	7,3	**29**	**29**
D09	28	PO	3,2	5,0	0,6	24,0	15,0	**29**	**29**
D10	29	9	5,8	3,2	1,8	37,4	11,6	**29**	**30**
D11	31	7	5,5	2,7	2,0	54,9	16,5	**29**	**29**
D12	18	10	5,3	3,0	1,8	14,8	20,0	**29**	**32**
D13	28	10	7,0	4,6	1,5	61,8	17,7	**28**	**29**
D14	33	6	6,5	3,4	1,9	33,5	23,5	**28**	**30**
D15	21	9	5,6	2,7	2,1	35,5	10,5	**29**	**29**
D16	21	11	5,7	2,7	2,1	37,9	25,7	**29**	**29**
D17	29	PO	4,9	3,9	1,3	46,5	6,9	**29**	**29**
D18	22	11	4,1	4,3	1,0	25,9	16,3	**29**	**29**
D19	23	PO	4,1	3,7	1,1	25,7	19,0	**28**	**29**
D20	30	6	7,1	4,7	1,5	36,5	7,6	**29**	**31**
D21	26	7	6,7	6,6	1,0	50,4	10,0	**28**	**28**
D22	21	PO	5,9	5,9	1,0	36,2	10,6	**29**	**30**
D23	25	10	3,7	38,0	0,1	26,8	8,7	**28**	**30**
D24	23	16	4,9	8,5	0,6	26,4	30,9	**24**	**30**
D25	23	4	3,1	2,4	1,3	83,5	9,1	**22**	**29**
D26	26	PO	3,7	6,9	0,5	81,6	26,2	**19**	**29**
D27	24	6	6,2	3,6	1,7	25,3	11,4	**19**	**30**
D28	25	NA	5,7	3,0	1,9	33,4	11,6	**19**	**29**
D29	28	12	5,7	2,9	2,0	45,6	17,7	**20**	**28**
D30	28	9	9,5	10,6	0,9	148,0	20,7	**19**	**31**
D31	30	11	6,3	7,4	0,9	56,6	23,4	**19**	**28**
D32	31	5	4,8	5,8	0,8	5,0	7,3	**19**	**28**
D33	28	11	5,2	6,7	0,8	81,2	14,1	**17**	**32**
D34	26	9	6,7	6,3	1,1	23,0	9,8	**23**	**28**
D35	22	7	6,0	8,8	0,7	35,9	7,6	**15**	**29**
D36	22	12	6,8	5,2	1,3	22,5	11,7	**19**	**29**
D37	23	12	9,5	9,3	1,0	54,7	19,8	**19**	**29**
D38	18	16	5,7	5,8	1,0	26,7	18,3	**24**	**30**
D39	23	PO	2,8	1,7	1,6	13,2	6,8	**19**	**31**
D40	31	10	4,5	3,9	1,2	19,5	NA	**17**	**29**
D41	26	10	4,6	3,3	1,4	36,0	10,6	**19**	**37**
D42	20	PO	7,6	11,7	0,6	42,9	13,9	**22**	**36**
D43	25	6	4,3	5,3	0,8	10,6	13,1	**19**	**37**
D44	27	15	7,3	5,4	1,4	60,7	10,8	**19**	**35**
D45	22	12	6,4	6,3	1,0	10,7	16,4	**29**	**38**
D46	23	10	3,3	2,6	1,3	69,4	10,4	**28**	**38**
D47	31	9	5,1	4,5	1,1	5,0	17,8	**31**	**38**
D48	21	13	6,7	3,4	2,0	54,2	21,9	**19**	**19**
D49	31	10	11,1	8,0	1,4	30,6	11,0	**19**	**22**
D50	26	PO	7,0	9,7	0,7	16,5	13,7	**40**	**48**

PO- Polycystic ovary; NA- Not available

Annex 2. Results of AGG interrution patterns.

a) A of cases studied by TP-PCR (n=45):

1. Sample D08.

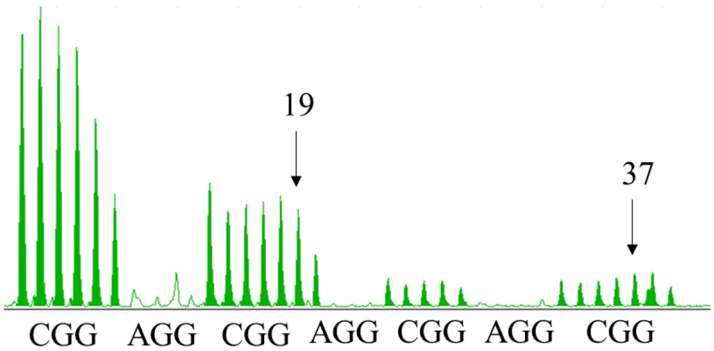


AGG interruption pattern:

Allele 1 - [CGG]9AGG[CGG]9AGG[CGG]9

Allele 2 - [CGG]9AGG[CGG]9AGG[CGG]9

2. Sample D43.



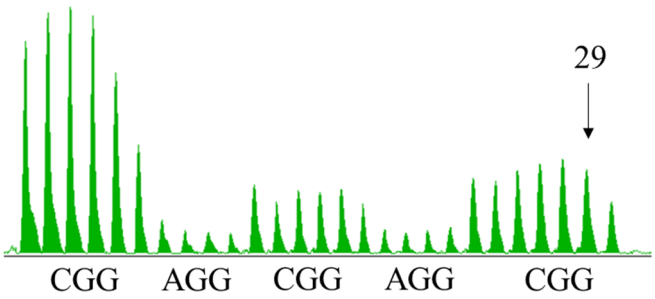
AGG interruption pattern:

Allele 1 - [CGG]10AGG[CGG]8

Allele 2 - [CGG]8AGG[CGG]6AGG[CGG]9AGG[CGG]9

b) Exemples of cases studied by TP-PCR and RFLA (n=4):

1. Sample D02

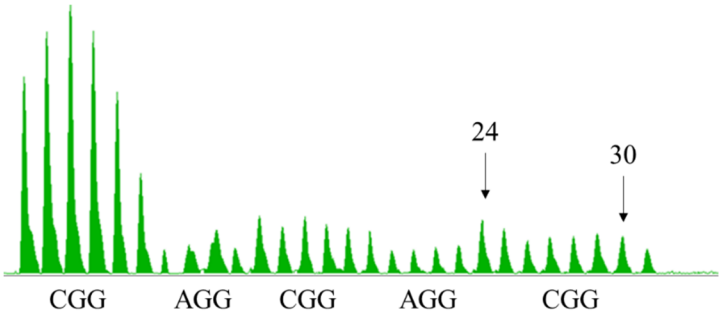


AGG interruption pattern:

Allele 1 - [CGG]29 or [CGG]19AGG[CGG]9

Allele 2 - [CGG]9AGG[CGG]9AGG[CGG]9 or [CGG]9AGG[CGG]19

2. Sample D24.

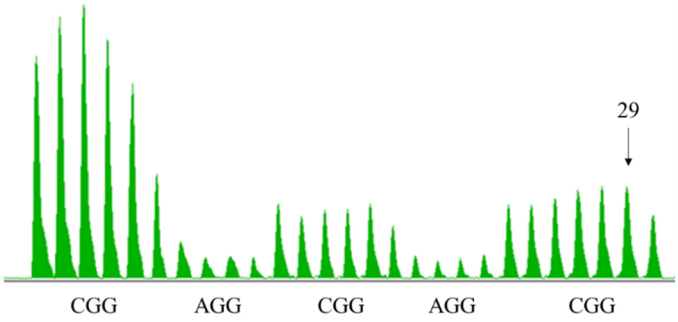


AGG interruption pattern:

Allele 1 - [CGG]24 or [CGG]14AGG[CGG]9

Allele 2 - [CGG]10AGG[CGG]9AGG[CGG]9 or [CGG]10AGG[CGG]19

3. Sample D18.

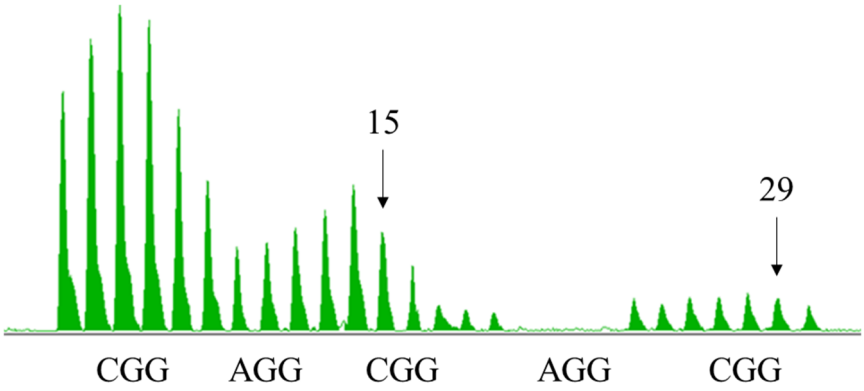


AGG interruption pattern:

Allele 1 - **[CGG]29** or [CGG]19AGG[CGG]9

Allele 2 - **[CGG]9AGG[CGG]9AGG[CGG]9** or [CGG]9AGG[CGG]19

4. Sample D35.

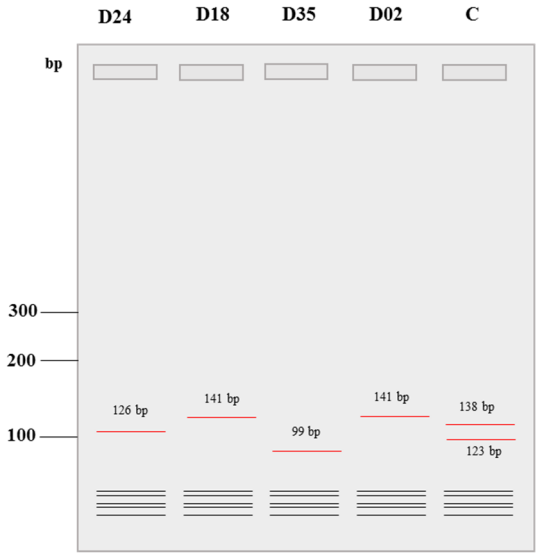


AGG interruption pattern:

Allele 1 - **[CGG]15** or [CGG]5AGG[CGG]9

Allele 2 - **[CGG]9AGG[CGG]9AGG[CGG]9** or [CGG]9AGG[CGG]19

5. Size of the expected fragment if the smallest allele is pure.



Samples	Allele size (CGGs)	Expected fragment size (bp)*	Fragments obtained independently of the CGG number or AGG pattern
D24	24/30	126	37, 33, 25, 20 and 13 bp
D18	29/29	141	
D35	15/29	99	
D02	29/29	141	
Control (C)	23/28 (both pure)	123 and 138 (respectively)	

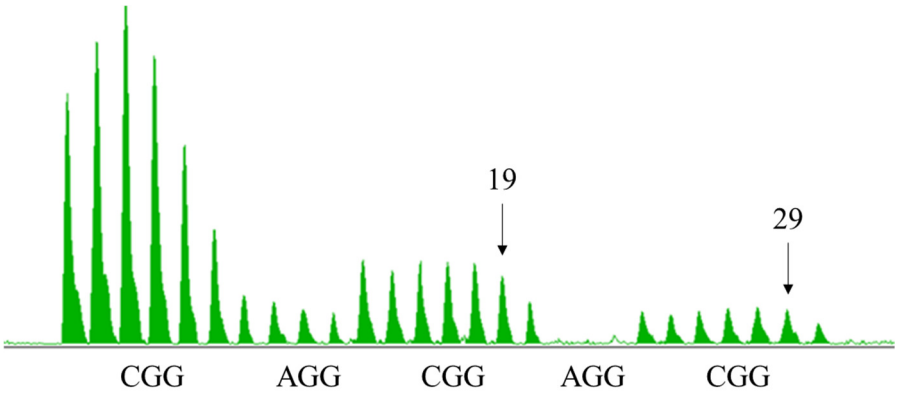
*If the smallest allele is pure

Image adapted from the simulator NEBcutter®.

Legend:
- Fragments obtained independently of the CGG number or AGG pattern
- Expected fragment if the smallest allele is pure

c) Example of cases studied by TP-PCR (in-house and commercial methods) (n=1):

1. Sample D37.

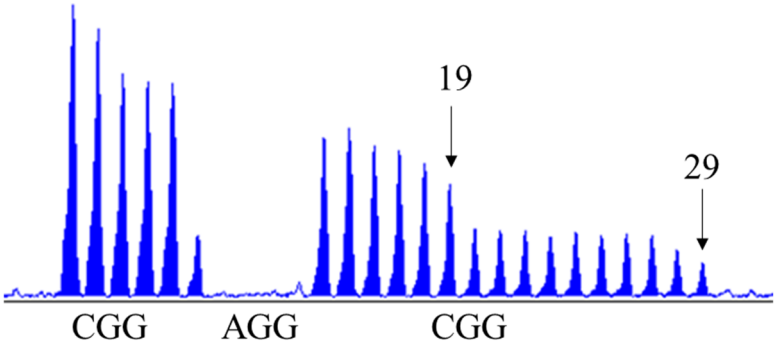


AGG interruption pattern:

Allele 1 – [CGG]9AGG[CGG]9 or [CGG]19

Allele 2 - [CGG]9AGG[CGG]19 or [CGG]9AGG[CGG]9AGG[CGG]9

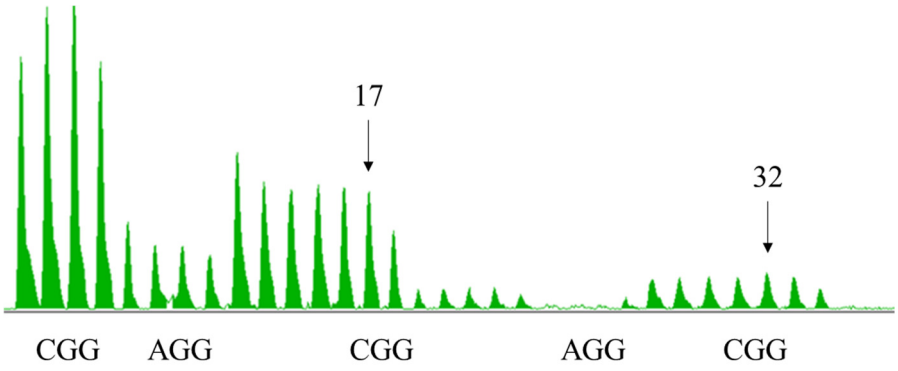
2. Sample D37.



FRAXA PCR kit (LabSscan™)

d) Case studied by TP-PCR and Sanger Sequencing (n=1):

1. Sample D33.

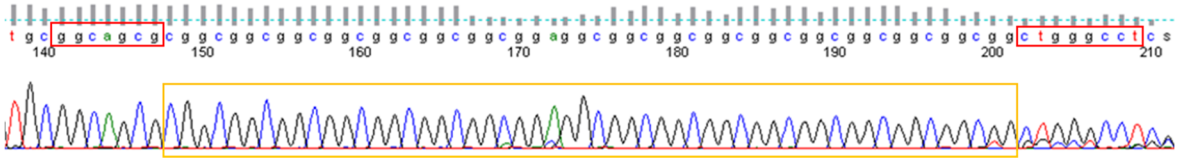


AGG interruption pattern:

Allele 1 - [CGG]9AGG[CGG]7 or [CGG]17

Allele 2 - [CGG]8AGG[CGG]23 or [CGG]8AGG[CGG]15AGG[CGG]7

2. Sample D33.



Partied electropherogram of Sample D33.

Annex 3. Summary of data organized according AGG interruption pattern and allelic score.

								Allele Size		AGG interspersion pattern TP-PCR; Sanger sequencing; RFLA; TP-PCR (commercial kit)	
	Samples	Age (y)	tAFC	FSH (mUl/ml)	LH (mUl/ml)	E2 (pg/ml)	PRL (ng/ml)	A1	A2	Allele 1	Allele 2
Equivalent pattern group	D48	21	13	6,7	3,4	54,2	21,9	19	19	[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9
	D49	31	10	11,1	8,0	30,6	11,0	19	22	[CGG9AGG[CGG]9	[CGG]9AGG[CGG]12
	D33	28	11	5,2	6,7	81,2	14,1	17	32	[CGG]9AGG[CGG]7	[CGG]8AGG[CGG]23
	D37	23	12	9,5	9,3	54,7	19,8	19	29	[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]19
	D29	28	12	5,7	2,9	45,6	17,7	20	28	[CGG]10AGG[CGG]9	[CGG]16AGG[CGG]11
	D25	23	4	3,1	2,4	83,5	9,1	22	29	[CGG]12AGG[CGG]9	[CGG]16AGG[CGG]12
	D34	26	9	6,7	6,3	23,0	9,8	23	28	[CGG]15AGG[CGG]8	[CGG]18AGG[CGG]9
	D42	20	PO	7,6	11,7	42,9	13,9	22	36	[CGG]13AGG[CGG]8	[CGG]24AGG[CGG]11
	D21	26	7	6,7	6,6	50,4	10,0	28	28	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]9AGG[CGG]9AGG[CGG]8
	D13	28	10	7,0	4,6	61,8	17,7	28	29	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]9AGG[CGG]9AGG[CGG]9
	D19	23	0P	4,1	3,7	25,7	19,0	28	29	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]9AGG[CGG]9AGG[CGG]9
	D03	30	10	5,0	6,7	57,7	13,6	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D08	25	9	7,0	4,5	70,2	7,3	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D09	28	PO	3,2	5,0	24,0	15,0	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D11	31	7	5,5	2,7	54,9	16,5	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D15	21	9	5,6	2,7	35,5	10,5	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D16	21	11	5,7	2,7	37,9	25,7	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D17	29	PO	4,9	3,9	46,5	6,9	29	29	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D06	28	9	5,2	5,2	33,7	10,3	29	31	[CGG]9AGG[CGG]10AGG[CGG]8	[CGG]9AGG[CGG]11AGG[CGG]9
	D20	30	6	7,1	4,7	36,5	7,6	29	31	[CGG]9AGG[CGG]10AGG[CGG]8	[CGG]9AGG[CGG]11AGG[CGG]9
	D04	22	7	6,3	5,6	15,8	5,0	28	29	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]10AGG[CGG]9AGG[CGG]8
	D14	33	6	6,5	3,4	33,5	23,5	28	30	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]10AGG[CGG]9AGG[CGG]9
	D05	28	12	8,9	6,5	51,5	9,8	29	34	[CGG]9AGG[CGG]10AGG[CGG]8	[CGG]14AGG[CGG]9AGG[CGG]9
	D47	31	9	5,1	4,5	5,0	17,8	31	38	[CGG]12AGG[CGG]9AGG[CGG]8	[CGG]18AGG[CGG]9AGG[CGG]9
	(D45)	22	12	6,4	6,3	10,7	16,4	29	38	[CGG]10AGG[CGG]9AGG[CGG]8	CGG]8AGG[CGG]9AGG[CGG]9AGG[CGG]9
	D50	26	PO	7,0	9,7	16,5	13,7	40	48	[CGG]22AGG[CGG]9AGG[CGG]7	[CGG]30AGG[CGG]8AGG[CGG8
Opposite pattern group	D35	22	7	6,0	8,8	35,9	7,6	15	29	[CGG]15	[CGG]9AGG[CGG]9AGG[CGG]9
	D02	30	5	6,2	5,4	55,5	24,2	29	29	[CGG]29	[CGG]9AGG[CGG]9AGG[CGG]9
	D18	22	11	4,1	4,3	25,9	16,3	29	29	[CGG]29	[CGG]9AGG[CGG]9AGG[CGG]9
	D24	23	16	4,9	8,5	26,4	30,9	24	30	[CGG]24	[CGG]10AGG[CGG]9AGG[CGG]9
	D26	26	PO	3,7	6,9	81,6	26,2	19	29	[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D28	25	NA	5,7	3,0	33,4	11,6	19	29	[CGG]9AGG[CGG]9	[CGG]9AGG[CGG]9AGG[CGG]9
	D40	31	10	4,5	3,9	19,5	NA	17	29	[CGG]10AGG[CGG]6	[CGG]9AGG[CGG]9AGG[CGG]9
	D36	22	12	6,8	5,2	22,5	11,7	19	29	[CGG]10AGG[CGG]8	[CGG]9AGG[CGG]9AGG[CGG]9
	D39	23	PO	2,8	1,7	13,2	6,8	19	31	[CGG]10AGG[CGG]8	[CGG]9AGG[CGG]12AGG[CGG]8
	D27	24	6	6,2	3,6	25,3	11,4	19	30	[CGG]9AGG[CGG]9	[CGG]10AGG[CGG]9AGG[CGG]9
	D01	20	10	5,4	4,8	39,9	4,0	28	29	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]9AGG[CGG]9AGG[CGG]9
	D30	28	9	9,5	10,6	148,0	20,7	19	31	[CGG]9AGG[CGG]9	[CGG]10AGG[CGG]10AGG[CGG]9
	D38	18	16	5,7	5,8	26,7	18,3	24	30	[CGG]9AGG[CGG]14	[CGG]10AGG[CGG]9AGG[CGG]9
	D31	30	11	6,3	7,4	56,6	23,4	19	28	[CGG]10AGG[CGG]8	[CGG]8AGG[CGG]9AGG[CGG]9
	D32	31	5	4,8	5,8	5,0	7,3	19	28	[CGG]10AGG[CGG]8	[CGG]8AGG[CGG]9AGG[CGG]9
	D23	25	10	3,7	38,0	26,8	8,7	28	30	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]19AGG[CGG]10
	D10	29	9	5,8	3,2	37,4	11,6	29	30	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]20AGG[CGG]9
	D22	21	PO	5,9	5,9	36,2	10,6	29	30	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]20AGG[CGG]9
	D07	19	8	4,7	3,4	37,1	9,1	29	32	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]22AGG[CGG]9
	D12	18	10	5,3	3,0	14,8	20,0	29	32	[CGG]9AGG[CGG]9AGG[CGG]9	[CGG]22AGG[CGG]9
	D46	23	10	3,3	2,6	69,4	10,4	28	38	[CGG]9AGG[CGG]9AGG[CGG]8	[CGG]29AGG[CGG]8
	D44	27	15	7,3	5,4	60,7	10,8	19	35	[CGG]10AGG[CGG]8	[CGG]8AGG[CGG]6AGG[CGG]9AGG[CGG]9
	D43	25	6	4,3	5,3	10,6	13,1	19	37	[CGG]9AGG[CGG]9	[CGG]8AGG[CGG]8AGG[CGG]9AGG[CGG]9
	D41	26	10	4,6	3,3	36,0	10,6	19	37	[CGG]10AGG[CGG]8	[CGG]8AGG[CGG]8AGG[CGG]9AGG[CGG]9

PO - Polycystic Ovary; NA- Not Available; (D45) - sample removed from statistical study; normal/normal; normal/low; low/high; normal/high; low/low; high/high.