U.PORTO

DOUTORAMENTO

CIÊNCIAS BIOMÉDICAS

The role of *FMR1* gene repetitive tract complexity in female fertility

Bárbara Rodrigues

D

2024

U.PORTO

Bárbara Rodrigues. **The role of *FMR1* gene repetitive tract complexity in female fertility**

D.ICBAS **2024**

The role of *FMR1* gene repetitive tract complexity in female fertility

Bárbara Luísa Cerqueira Rodrigues

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

Bárbara Luísa Cerqueira Rodrigues

The role of *FMR1* gene repetitive tract complexity in female fertility

Tese de candidatura ao grau de Doutor em Ciências Biomédicas;
Programa Doutoral da Universidade do Porto
(Instituto de Ciências Biomédicas Abel Salazar)

Orientadora – Paula Maria Vieira Jorge

Categoria – Professor Auxiliar

Afiliação – Laboratório de Citogenética, Departamento de Microscopia, Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto

Coorientador – António José Arsénia Nogueira Categoria

Categoria – Professor Catedrático

Afiliação – Centro de Estudos Ambientais e Marinhos (CESAM), Departamento de Biologia, Universidade de Aveiro

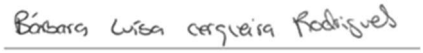
Coorientador – Carlos Pedro Fontes Oliveira

Categoria – Professor Associado

Afiliação – Departamento de Química & Laboratório Associado para a Química Verde (LAQV-REQUIMTE), Universidade de Aveiro

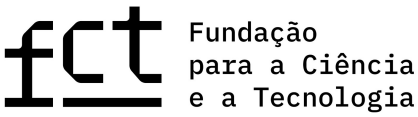
Declaração de Honra

Declaro que a presente tese é de minha autoria e não foi utilizada previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (afirmações, ideias, pensamentos) respeitam escrupulosamente as regras da atribuição, e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referenciação. Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.



Bárbara Luísa Cerqueira Rodrigues

This doctoral project was funded by the Portuguese Foundation for Science and Technology (FCT) under grants SFRH/BD/136398/2018 and COVID/BD/153204/2023.



doi.org/10.54499/COVID/BD/153204/2023

Para a estrela mais brilhante que o céu já viu,

o meu querido pai.

Acknowledgments

I would like to express my sincere gratitude to everyone who has supported and guided me throughout my PhD. Your collective encouragement and help have been fundamental in helping me reach this significant milestone.

To my supervisor, **Prof. Drª. Paula Jorge**, I would like to express my gratitude for the opportunity to be part of your team. Your guidance, critical and scientific vision, encouragement and constant support have been fundamental. Your dedication and passion for your work is truly inspiring and I am extremely grateful for everything you have taught and shared with me. Thank you for believing in me every step of the way.

To my co-supervisor, **Prof. António J. A. Nogueira**, I would like to express my gratitude for your generosity and for sharing your vast knowledge with me over the past six years. Your critical and scientific vision, input, unconditional support, and words of encouragement have been extremely important, making a significant difference to this journey. Thank you for always reassuring me that things would work out.

To my co-supervisor, **Prof. Pedro Fontes Oliveira**, I would like to express my deepest gratitude for the enlightening discussions, valuable teachings, unwavering support, and constant availability.

I would like to extend my heartfelt gratitude to the **collaborators** of the **Molecular Genetics Unit** at the Laboratory Genetics Service, Genetics and Pathology Clinic of the Santo António Local Health Unit (ULSSA) for their unconditional support, constant encouragement, invaluable teachings, and for always being there for me. You have been my home for the last seven years. Special thanks to **Drª. Rosário Santos**, the head of the laboratory, for the enlightening discussions, scientific contributions, and constant support. I am also deeply grateful to **Emília Vieira** for her immense support, especially for her comfort during one of the most vulnerable periods of my life.

To **Drª. Isabel Marques**, I am very grateful for the enlightening scientific discussions, unwavering support, and continued availability.

I would like to thank the collaborators of the Centre for Assisted Reproduction and the Public Gamete Bank at the Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), ULSSA, especially **Dr. Emídio Vale-Fernandes** and the embryologists **Carla Leal**, **Daniela Sousa**, and **Raquel Brandão**, for their help, availability, scientific input, and enlightening discussions. Thank you for being an integral part of this project.

I would like to thank the Instituto Nacional de Saúde, Dr. Ricardo Jorge (INSA), and especially **Drª. Solange Costa**, **Filipa Esteves**, and **Joana Pires** for their warm welcome, sharing of knowledge, and support. Your positiveness is truly inspiring.

To my **family** for their constant concern and encouragement, and especially to my Aunt **Aurora** for her support in making this possible. To my family friend, **Sr. Clemente**, thank you for your wise words that gradually showed me I was on the right path. Your guidance gave me the strength and motivation to keep pursuing my dreams.

To my **friends**, thank you for your unconditional support and for patiently enduring all the times you've been dragged into scientific debates - I'm grateful to have such great listeners. A special thank you to **Sara Nadine**, who always believed in me, lifted me up, and reminded me that I was capable.

To my first student, **Vanessa Sousa**, who arrived full of dreams and quickly became a close ally: I hope I was up to the challenge and demonstrated the value of perseverance. Your constant support, positive attitude, and willingness to help me have been invaluable. You have also taught me a lot.

To my research partner, **Nuno Maia**, words cannot fully express how important you have been to me during this journey. From the beginning, our connection was immediate and strong. Thank you for everything you have taught me, for your unwavering support, motivation, strength, friendship, and companionship. Thank you for always believing in me. We will always have each other's backs.

To my best friend, **Tânia Pereira**, thank you for your unwavering support, friendship, encouragement, and presence. You have been my anchor, giving me strength in the most difficult moments and constantly reminding me of my resilience and courage. I am deeply grateful to have you in my life and for the extraordinary friendship we share.

And finally, to my parents, **Carlos Henrique** and **Maria Luísa**, who have always been by my side, encouraging me and helping me make all my dreams come true. A special hug to my father, who isn't here to share this moment with me, but I know he was with me every step of the way. This achievement is also his, and I know he would be so proud. I miss him very much. A special thank you to my mother, who, over the last two years, has been exceptionally strong and has done everything possible - and even impossible - to make everything work. To my little dog, **Balu**, who brought so much joy into my life: thank you for being my faithful companion during the long days of writing. To my boyfriend, **Luís Filipe**, thank you for your unwavering support and for always believing in me. You've been my partner through life's ups and downs, encouraging me every day to follow my dreams. Your

ability to make me smile, even on the toughest days, is extraordinary. Your constant presence and boundless generosity make you my safe haven. I feel incredibly lucky to have you in my life. These four are the loves of my life.

Abstract

Infertility is defined as the inability to achieve a clinical pregnancy after twelve months of regular, unprotected sexual intercourse. Affecting approximately 13% of reproductive age couples, it underscores the imperative need for effective preventive measures. Recent research has increasingly highlighted the role of genetic factors in reproductive challenges. The *FMR1* gene, a pivotal player in female reproductive health, has been implicated in primary ovarian insufficiency. Specifically, the *FMR1* premutation (PM) is a recognized cause of fragile X-associated primary ovarian insufficiency (FXPOI). Moreover, variations in the CGG repeat length within the 5' UTR of *FMR1* gene may be associated with diverse infertility phenotypes.

This thesis investigates the association between the *FMR1* gene and female fertility through a detailed characterization of its repetitive tract. A multifaceted approach, encompassing genetic analysis, clinical evaluations, and statistical modeling, was employed to explore how variations in the repetitive tract, such as total CGG repeat length and the number and pattern of AGG interruptions, may correlate with various fertility outcomes.

The present study yielded significant findings: i) The *FMR1* repetitive tract complexity was successfully quantified using the novel measure, *allelic score*, providing an accurate measure of its variation; ii) The methylation status of the *FMR1* gene was shown to be a reliable indicator of X-chromosome inactivation patterns; iii) The observed high prevalence of *FMR1* PM in females undergoing intracytoplasmic sperm injection (ICSI) highlights need for targeted genetic screening in the Portuguese population to identify potential carriers and to inform reproductive planning; iv) The combinations of allelic complexity were proposed as potential predictors of the risk of FXPOI development in *FMR1* PM carriers, allowing for early identification and intervention; v) The allelic complexity of the *FMR1* gene was demonstrated to be a predictor of fertilization success; vi) DNA damage in cumulus cells was established as potential marker of oocyte competence, providing a valuable tool for assessing oocyte quality and predicting fertilization potential; vii) The establishment of a biorepository and associated database represents a substantial achievement, enabling long-term research into the genetic factors influencing infertility.

These findings emphasize the importance of genetic screening in addressing infertility. By implementing preventive measures and tailoring therapeutic approaches based on genetic risk we can potentially improve outcomes for females affected by *FMR1*-related infertility. This research advances our understanding of reproductive health and provides a valuable foundation for future investigations in reproductive genetics.

Keywords: *FMR1* gene, CGG/AGG repetitive tract, *allelic score*, *FMR1* premutation, oocyte competence, fertilization success

Resumo

A infertilidade é definida como a incapacidade de alcançar uma gravidez clínica após doze meses de relações sexuais regulares e desprotegidas. Este problema afeta aproximadamente 13% dos casais em idade reprodutiva, evidenciando a necessidade imperativa de medidas preventivas eficazes. Investigações recentes têm destacado, cada vez mais, o papel dos fatores genéticos nos desafios reprodutivos. O gene *FMR1*, um elemento fundamental na saúde reprodutiva, tem sido implicado na insuficiência ovárica primária. Especificamente, a premutação (PM) neste gene é uma causa conhecida de insuficiência ovárica primária associada ao X frágil (FXPOI). Além disso, variações no tamanho das repetições CGG, localizadas na região 5' não traduzida do gene *FMR1,* podem estar associadas a diversos fenótipos de infertilidade.

Esta tese aprofundou a compreensão da relação entre o gene *FMR1* e a fertilidade feminina, com foco na caracterização detalhada da sua região repetitiva. Através de uma abordagem multidisciplinar que combinou análise genética, avaliações clínicas e modelagem estatística, foram exploradas as variações na região repetitiva, como o comprimento total das repetições CGG e o número e padrão das interrupções AGG, e a sua relação com diferentes resultados de fertilidade.

As principais conclusões deste trabalho são as seguintes: i) A complexidade da região repetitiva foi quantificada com sucesso através de uma nova medida, a pontuação alélica, que reflete com precisão a variabilidade desta região; ii) O estado de metilação do gene *FMR1* demonstrou ser um marcador confiável do padrão de inativação do cromossoma-X; iii) A elevada prevalência de portadoras de PM observada sublinhou a necessidade de um rastreio genético direcionado na população portuguesa, para identificar potenciais portadoras e orientar o planeamento reprodutivo; iv) As combinações de complexidade alélica foram sugeridas como potenciais preditores do risco de desenvolvimento de FXPOI em portadoras de PM no gene *FMR1*; v) A complexidade alélica do gene *FMR1* demonstrou ser um preditor do sucesso da fertilização; vi) O dano no DNA nas células do cumulus foi estabelecido como um potencial marcador da competência ovocitária, podendo constitutir uma ferramenta valiosa para avaliar a qualidade dos ovócitos e prever o sucesso da fertilização; vii) A criação de um repositório de amostras e respetivos dados representa uma conquista significativa, permitindo a continuidade da investigação dos fatores genéticos que influenciam a infertilidade a longo prazo.

Estas descobertas salientam a importância do rastreio genético no tratamento da infertilidade. A implementação de medidas preventivas e a adaptação das abordagens

terapêuticas, baseadas no risco genético, podem potencialmente melhorar os resultados para as mulheres afetadas pela infertilidade relacionada com o gene *FMR1.* Esta investigação aprofundou a nossa compreensão da saúde reprodutiva e fornece uma base valiosa para futuras investigações em genética reprodutiva.

Palavras-chave: gene *FMR1*, região repetitiva CGG/AGG, pontuação alélica, premutação no gene *FMR1*, competência ovocitária, sucesso da fertilização

List of contents

List of figures

List of tables

Chapter V. Use of the *FMR1* gene methylation status to assess the X-Chromosome Inactivation pattern: a stepwise analysis

List of abbreviations and acronyms

0-9

0PN – Zero pronuclei

1PN – One pronuclei

2PN –Two pronuclei

3PN – Three pronuclei

%TDNA – Levels of DNA damage

A

A1 – Allele 1

ACMG – American College of Medical Genetics

AFC – Antral follicle count

AGG – Adenine-Guanine-Guanine

AMH – Anti-Müllerian hormone

ANCOVA – Analysis of covariance

AR – Human androgen receptor gene

ART – Assisted reproductive technologies

ASRM – American Society for Reproductive Medicine

B

BMI – Body mass index

bp – base pairs

C

CA – California

CAG – Cytosine-Adenine-Guanine

cAMP – cyclic adenosine monophosphate

CCs – Granulosa cumulus cells

CE – Capillary electrophoresis

CGG – Cytosine-Guanine-Guanine

CGMJM – Centro de Genética Médica Doutor Jacinto Magalhães

cGMP – cyclic guanosine monophosphate

CHUP– Centro Hospitalar Universitário do Porto

CMIN – Centro Materno-Infantil do Norte Dr. Albino Aroso

COC – Cumulus-oocyte complex

D

dATP – deoxyadenosine triphosphate

dCTG - deoxycytidine triphosphate

dGTP – deoxyguanosine triphosphate

DHE-S – Dehydroepiandrosterone sulfate

DHT – Dihydrotestosterone

DMEM/F-12 – Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12

DMSO – Dimethyl sulfoxide

DOR – Diminished ovarian reserve

dTTP – 2′- deoxythymidine 5′-triphosphate

E

E_2 – Estradiol

EDTA – Ethylenediaminetetraacetic acid

EMQE – European Molecular Genetics Quality Network

F

F – Forward

FCT – Fundação para a Ciência e Tecnologia

FF – Follicular fluid

FMR1 – Fragile X messenger Ribonucleoprotein

FMRP – Fragile X mental retardation protein

FSH – Follicle-stimulating hormone

FXAND – Fragile X-associated neuropsychiatric disorders

FXDOR – Fragile-X-associated diminished ovarian reserve

FXPOI – Fragile X-associated Primary Ovarian Insufficiency

FXS – Fragile X syndrome

FXTAS – Fragile X-associated tremor/ataxia syndrome

G

GC – Guanine-Cytosine

GCs – Granulosa cells

gDNA – genomic Deoxyribonucleic acid

H

HA – Hyaluronic acid

HUMARA – Human androgen-receptor assay

I

ICBAS – Instituto de Ciências Biomédicas Abel Salazar

ICSI – Intracytoplasmic sperm injection

IL – Illinois

ITR – Laboratory for Integrative and Translational Research in Population Health

IVF – *In vitro* fertilization

K

Kb – Kilobases

$KHCO_3$ – Potassium bicarbonate

L

LA – Long and accurate

LH – Luteinizing hormone

LMA – Low-melting point

M

Me – Percentage of methylation

MGs – Mural granulosa cells

MII – Metaphase II

mPCR – methylation PCR

mRNA – messenger RNA

N

NCBI – National Center for Biotechnology Information

O

OECD – Organisation for Economic Co-operation and Development

OHP – 17-Hydroxyprogesterone

P

P1 – Primer 1

P2 – Primer 2

PBS – Phosphate buffered saline

PC1 – Principal component

PC2 – Second principal component

PCA – Principal component analysis

PCOS – Polycystic ovary syndrome

PCR – Polymerase Chain reaction

PGT – Preimplantation genetic testing

PM – Premuation

POF – Premature ovarian failure

POI – Primary ovarian insufficiency

R

R – Reverse

RFU – Relative fluorescence units

RNA – Ribonucleic acid

S

S.D. – Standard deviation

SDS – Sodium dodecyl sulfate

SPSS – Statistical package for social science for windows

T

T3L – Free triiodothyronine

T4L – Free thyroxine

TLE – Erythrocyte lysis buffer

TLN – Nucleus lysis buffer

TP-PCR – Triplet-primed polymerase chain reaction

Tris – Tris(hydroxymethyl)aminomethane

TSH – Stimulating hormone

TX – Texas

U

UK – United Kingdom

ULSSA – Unidade local de saúde de Santo António

UMIB – Unity for Multidisciplinary Research in Biomedicine

UP – Universidade do Porto

USA – United States of America

UTR – Untranslated region

W

WI – Wisconsin

X

XCI – X-chromosome inactivation

Z

ZP – Zona pellucida

List of publications

A – Publications resulting from work performed during the PhD thesis

A1. International peer-reviewed journals

1. Bárbara Rodrigues, Emídio Vale-Fernandes, Vanessa Sousa, Isabel Marques, Rosário Santos, António J A Nogueira, Paula Jorge. Exploring the predictive value of the *FMR1* gene allelic complexity for *in vitro* fertilization success. Submitted for publication in Journal of Assisted Reproduction and Genetics.

- BR conducted the statistical analysis and drafted the entire manuscript, including the interpretation and presentation of the results.

2. Bárbara Rodrigues*, Vanessa Sousa*, Filipa Esteves, Emídio Vale-Fernandes, Solange Costa, Daniela Sousa, Raquel Brandão, Carla Leal, Isabel Gaivão, João Paulo Teixeira, António J A Nogueira, Paula Jorge. Cumulus cell DNA damage linked to fertilization success in females with an ovulatory dysfunction phenotype. Accepted for publication in Frontiers in Cell and Developmental Biology, October 24th, 2024. doi: 10.3389/fcell.2024.1448733

- BR was responsible for analyzing the percentage of DNA in the comet tails of individual cells, conducting statistical tests to assess the significance of the data, and participating in the discussion of the results.

3. Bárbara Rodrigues, Vanessa Sousa, Carolyn M. Yrigollen, Flora Tassone, Olatz Villate, Emily G. Allen, Anne Glicksman, Nicole Tortora, Sarah L. Nolin, António J. A. Nogueira, Paula Jorge. *FMR1* allelic complexity in premutation carriers provides no evidence for a correlation with age at amenorrhea. Reprod Biol Endocrinol. 2024 Jun 21;22(1):71. doi: 10.1186/s12958-024-01227-5

- BR's key responsibilities included drafting the entire manuscript and collaborating with co-authors through extensive discussions.

4. Bárbara Rodrigues, Ana Gonçalves, Vanessa Sousa, Nuno Maia, Isabel Marques, Emídio Vale-Fernandes, Rosário Santos, António J A Nogueira, Paula Jorge. Use of the *FMR1* Gene Methylation Status to Assess the X-Chromosome Inactivation Pattern: A Stepwise Analysis. Genes (Basel). 2022 Feb 25;13(3):419. doi: 10.3390/genes1303041

- BR performed laboratory experiments and data analysis on *FMR1* methylation, analyzing the resulting data to identify potential genetic markers.

5. Bárbara Rodrigues, Emídio Vale-Fernandes, Nuno Maia, Flávia Santos, Isabel Marques, Rosário Santos, António J A Nogueira, Paula Jorge. Development and Validation of a Mathematical Model to Predict the Complexity of *FMR1* Allele Combinations. Front Genet. 2020 Nov 13:11:557147. doi: 10.3389/fgene.2020.557147

- BR contributed to the development of a formula for calculating the allelic complexity of the *FMR1* gene's repetitive tract by preparing and organizing the required data.

* Authors contributed equally

Thesis outline

This thesis is structured into four main sections. Section 3 is comprised of six chapters, five of which (Chapters I to V) are based on published, accepted or submitted scientific papers in peer-reviewed journals.

Section 1 provides a general overview of the *FMR1* gene's role in female fertility. It delves into commonly employed assisted reproductive technologies, including conventional *in vitro* fertilization (IVF) and intracytoplasmic sperm injection (ICSI). Additionally, this section underscores the significance of cumulus cells (CCs) in oocyte development and maturation, highlighting the utility of their DNA integrity as a marker for assessing oocyte quality. Furthermore, an overview of the *FMR1* gene's function within granulosa cells (GCs) is presented.

Section 2 outlines the overarching goals and specific objectives of this doctoral research project.

Section 3 is divided into six chapters: **Chapter I** introduces a novel approach for analyzing the repetitive tract of the *FMR1* gene through the development of a formula that quantifies this tract in an *allelic score*, reflecting the allelic complexity of the repetitive region; **Chapter II** explores allelic complexity combinations as potential predictors in the risk assessment of developing fragile X-associated primary ovarian insufficiency (FXPOI); **Chapter III** investigates the potential of allelic complexity as a predictor of ovarian reserve and IVF success; **Chapter IV** examines the association between cumulus cell DNA damage and oocyte competence, with a focus on IVF outcomes. Additionally, the study explores blood samples as a surrogate marker for DNA damage in CCs, aiming to establish a reliable and accessible method for assessing DNA damage in a broader population; **Chapter V** investigates the use of *FMR1* gene methylation status to determine X-chromosome inactivation (XCI) patterns; **Chapter VI** provides a comprehensive account of the establishment of a biorepository and its associated database, including participant recruitment, sample collection, and preservation protocols.

Section 4 offers a comprehensive discussion of the thesis's findings, concluding remarks, and potential avenues for future research in the field.

1

Introduction

1. Overview of the fragile X messenger ribonucleoprotein 1 gene

The fragile X messenger ribonucleoprotein 1 (*FMR1*) gene is located at locus Xq27.3 on the long arm of the X chromosome. Spanning 38 kilobases (Kb), the gene contains 17 exons [1]. Within the 5' untranslated region (UTR), adjacent to the promoter-associated CpG island, resides a polymorphic CGG trinucleotide repeat, an example of a prototypical human dynamic mutation (Figure 1). This gene encodes fragile X mental retardation protein (FMRP), a multifunctional RNA-binding protein ubiquitously expressed. Within the central nervous system, FMRP, is pivotal for synaptic plasticity and architectural integrity [2–4]. Although its ovarian function remains to be elucidated, emerging evidence suggests a potential role in follicular development [5].

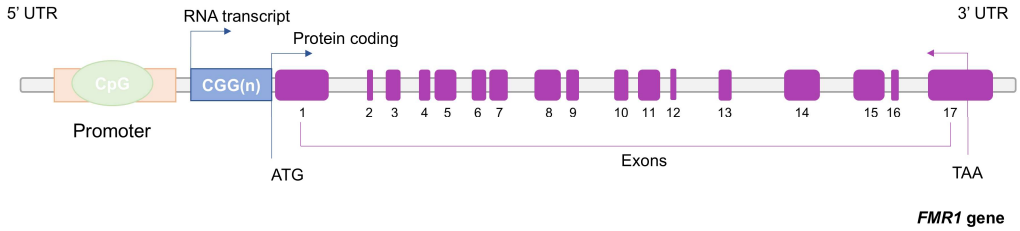


Figure 1. Schematic representation of the *FMR1* gene structure. The *FMR1* gene consists of a coding sequence encompassing 17 exons (NM.002024.6 mane select transcript), encoding the FMRP protein. Within the 5' UTR of the *FMR1* transcript lies a polymorphic trinucleotide repeat sequence composed of CGG units. Pathological genic alterations in *FMR1* are primarily associated with transcriptional silencing mediated by epigenetic mechanisms or dysregulation of *FMR1* mRNA levels. UTR – Untranslated region. Figure based on data obtained from the NCBI (National Center for Biotechnology Information), accessed on September 29th, 2024, from https://www.ncbi.nlm.nih.gov/.

1.1. Classification of *FMR1* alleles and associated diseases

Depending on the number of CGG repeats, *FMR1* alleles can be classified into four categories: normal (CGG < 45), intermediate or “gray zone” (45 ≤ CGG ≤ 54), premutation (PM, 55 ≤ CGG ≤ 200), and full mutation (CGG > 200) [6]. Full mutation alleles exhibit hypermethylation, resulting in transcriptional silencing and a consequence absence of FMRP. This underlies fragile X syndrome (FXS, OMIM #300624), the most prevalent inherited form of intellectual disability. In contrast, PM led to *FMR1* mRNA overexpression and FMRP reduction, manifesting as distinct phenotypes: fragile X–associated tremor/ataxia syndrome (FXTAS, OMIM #300623) and fragile X–associated primary ovarian insufficiency (FXPOI, OMIM #311360) (Figure 2) [2,7].

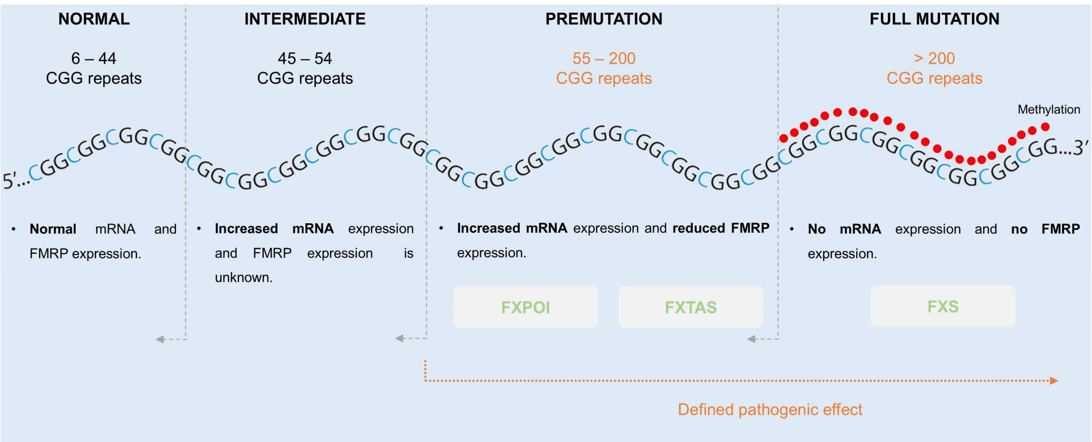


Figure 2. Schematic representation of *FMR1* allele classes based on CGG repeat number and associated clinical phonotypes. FXPOI – Fragile X-associated primary ovarian insufficiency; FXTAS – Fragile X-associated tremor/ataxia syndrome; FXS – Fragile X syndrome. Figure based on data from Biancalana et al. (2015) [6].

2. Importance of studying *FMR1* in the context of female reproductive health

Infertility is a globally prevalent reproductive health condition characterized by the failure to achieve a clinical pregnancy after twelve months or more of regular, unprotected sexual intercourse [8,9]. Approximately 2% of females experience primary infertility, defined as the inability to conceive, while 10% encounter secondary infertility, which occurs after a previous pregnancy [8,10,11]. It is estimated that infertility affects approximately 13% of couples worldwide, with nearly half seeking assisted reproductive technology (ART) to address this reproductive challenge [10,12,13].

According to the American Society for Reproductive Medicine (ASRM), infertility etiology can be attributed to male factors in approximately one third of cases, female factors in another third, and a combination for both in the remaining cases. However, in roughly 25–30% of couples, the underlying cause remains unidentified, a condition referred to as idiopathic infertility. A significant proportion of these idiopathic cases is suspected to have a genetic basis, encompassing chromosomal abnormalities, single-gene mutations, or complex multifactorial inheritance [14,15]. Moreover, lifestyle and environmental factors, such as tobacco use and obesity, have been implicated in compromised fertility [16,17].

Female factors account for approximately 70% of infertility cases either exclusively or in conjunction with male factors etiologies. Female infertility encompasses a spectrum of conditions, including genetic, hormonal and anatomical abnormalities. Predisposing factors such as pelvic inflammatory disease, uterine fibroids, and age-related decline in

reproductive function further contribute to infertility. In these conditions, compromised oocyte quality, impaired fertilization capacity, and implantation failures are commonly implicated [18,19]. Genetic influences, namely large alterations in the X chromosome, have been associated with ovarian dysgenesis, underscoring the complex interplay between genetic and environmental determinants of female fertility [20,21].

Ovulatory dysfunctions constitute a primary etiology of female infertility. Conditions such as polycystic ovary syndrome (PCOS), hypothalamic disorders, and primary ovarian insufficiency (POI), are prominent contributors [22–24]. POI, characterized by the loss of ovarian function before the age of 40, manifests as irregular or absent menstrual cycles (oligomenorrhea or amenorrhea), elevated levels of follicle-stimulating hormone (FSH), and diminished estradiol [25–27]. Affecting approximately 1% of females under 40 [28]. POI has emerged as a significant reproductive concern. Advances in next-generation sequencing have identified numerous genetic variants associated with POI [25,28,29], with over 137 genes reported to date in the WikiPathways database (ID: WP5461), although only a few X-linked genes have been described. The *FMR1* PM, linked to fragile X-associated primary ovarian insufficiency (FXPOI, OMIM #311360) represents the most prevalent monogenic cause of POI, accounting for approximately 6% of cases [30,31]. The *FMR1* gene's influence extends beyond FXPOI and has been implicated in other infertility phenotypes, such as diminished ovarian reserve (DOR) [32].

2.1. The impact of *FMR1* premutation on female fertility

The prevalence of the PM in the general population is estimated to be between 1 in 150 and 1 in 300 females, with approximately 20% of carries developing fragile X-associated primary ovarian insufficiency (FXPOI) [33,34]. FXPOI, also known as premature ovarian failure (POF), is defined by hypergonadotropic hypogonadism and the absence or irregularity of menstrual cycles prior to the age of 40 [30,35]. Diagnostic criteria for POF include amenorrhea for at least four months, in females under 40, accompanied by elevated follicle-stimulating hormone (FSH) levels (> 40 IU/L), decreased estradiol levels (< 30 - 50 pg/ mL) and reduced anti-Müllerian hormone (AMH) concentrations, a reliable marker of follicular depletion [31,36,37].

Females carrying the *FMR1* PM exhibit a heightened susceptibility to a range of health issues. Reproductive challenges, including infertility, are prevalent, and these individuals are also at increased risk for comorbidities, such as autoimmune, chronic pain, endocrine and cardiovascular disorders, as well as osteoporosis and mental health problems [38,39]. The spectrum of medical conditions associated with the *FMR1* PM has been

comprehensively termed fragile X–associated neuropsychiatric disorders (FXAND) [40]. Moreover, females PM carriers face the risk of transmitting fragile X syndrome (FXS) to their offspring [41,42]. Reproductive outcomes are often compromised in females with PM, characterized by reduced responsiveness to controlled ovarian hyperstimulation (COH), diminished fertilization rates, decreased fertility and DOR [43–45]. The observation of shorter menstrual cycles in PM carriers compared to age-matched controls further supports the notion that fragile-X-associated diminished ovarian reserve (FXDOR) is a precursor to FXPOI. Notably, females with 70 to 90 CGG repeats within the *FMR1* gene carry the highest risk for developing DOR [46]. While the precise mechanisms underlying the pathogenesis of FXPOI remain elusive, it is established that carriers of the *FMR1* PM exhibit elevated *FMR1* mRNA levels and concomitant reductions in FMRP protein expression as the CGG repeat length expands [47]. Conversely, females with a *FMR1* full mutation, characterized by complete *FMR1* silencing, do not demonstrate ovarian dysfunction. This observation refutes the hypothesis that FMRP deficiency is the primary etiological factor in FXPOI [48]. Emerging evidence suggests a model wherein elevated *FMR1* mRNA levels exert toxic effects within granulosa cells, culminating in follicular atresia and compromised follicle development [49].

2.1.1. CGG repetitive tract and X-chromosome inactivation

The relationship between CGG repeat length and FXPOI is not linear. While females with 70 to 100 CGGs exhibit a heightened risk of FXPOI compared to those with more than 100 repeats the penetrance of this phenotype is incomplete [50,51]. Allen et al. (2021) further refined this association, demonstrating that a CGG repeat length between 85 and 89 CGG carries a particularly elevated risk [52].

Skewed X-chromosome inactivation (XCI) has been proposed as a potential modifier of FXPOI risk and severity [53]. García-Alegría et al. (2007) observed a complex relationship between CGG repeat length and *FMR1* mRNA levels, suggesting that skewed XCI might influence mRNA expression [54]. However, a direct causal link between skewed XCI and FXPOI manifestation remains elusive [55]. Additional factors, such as AGG interruptions, may also contribute to the variability in FXPOI risk among PM carriers.

2.1.2. Mechanisms of expansion of the CGG repetitive tract

The CGG repetitive tract is usually interrupted by one or more AGG trinucleotide repeats, predominantly positioned every nine or ten CGG triplets, particularly in alleles of normal

length [56]. These AGG interruptions function as stabilizing elements preventing strand slippage during DNA replication and thus inhibiting the expansion of the repeat into pathogenic ranges [57,58]. The frequency of AGG interruptions within PM alleles inversely correlates with the risk of transmitting a full mutation to offspring in female carriers. A higher number of AGG interruptions within a PM allele is associated with a reduced likelihood of repeat expansion from the PM to the full mutation during oogenesis [59–63]. Conversely, a paucity of AGG interruptions leads to extended, uninterrupted CGG tracts at the 3' end of the allele, rendering the repeat more susceptible to instability. Generally, the length of these uninterrupted CGG segments is directly proportional to the risk of further expansion in subsequent generations [64].

2.1.3. Effect of repetitive tract variations on ovarian function

In addition to their well-established role in stabilizing CGG repeats, AGG interruptions have emerged as potential modulators of female fertility in PM carriers. Lekovich et al. (2018) reported an increased risk of DOR in carriers with fewer AGG interruptions (from two to one or none), and a correlation between longer uninterrupted CGG repeat tracts and decreased ovarian reserve. They proposed a mechanism whereby AGG interruptions prevent the formation of toxic *FMR1* mRNA secondary structures, which can sequester cellular proteins and impair ovarian function [46]. However, the influence of AGG interruptions on ovarian function remains controversial. Allen et al. (2018) found no association between AGG interruptions and the age of secondary amenorrhea in PM carriers, while Friedman-Gohas et al. (2020) failed to demonstrate a protective effect of AGG interruptions on ovarian response to controlled ovarian hyperstimulation [65,66] These conflicting findings underscore the complexity of the relationship between AGG interruptions and ovarian function, necessitating further investigation to elucidate their precise role in female fertility.

2.2. Other CGG repeat lengths and phenotypic outcomes

While a decline in oocyte quantity and quality is a natural physiological process associated with aging, some females experience POF, leading to premature infertility [67]. A decrease in the number and/or quality of remaining oocytes within the ovaries, characterizes a DOR, despite regular menstrual cycles, accompanied by elevated-yet-not-menopausal basal FSH levels, low AMH, and low antral follicle count (AFC) [68–70]. The prevalence of DOR among young females is estimated to be approximately 10% [68]. Females with DOR, indicative of premature ovarian aging, may exhibit reduced fecundity or a diminished response to ovarian stimulation, although this does not necessarily preclude the possibility of conception [30].

Although aging is the most common cause of DOR, a multitude of factors including chemotherapy, radiotherapy, autoimmune diseases, and genetic conditions can influence ovarian reserve [68]. Traditionally, the risk of DOR was primarily attributed to PM carriers; nevertheless, recent evidence suggests an association with carriers of normal-sized alleles as well.

Given the established association between the *FMR1* gene and ovarian reserve, specifically concerning ovarian aging, Gleicher et al. (2010) proposed a revised normal range of 26 to 34 CGG repeats [71]. Building on this, the authors categorize *FMR1* alleles into three groups based on CGG repeat number: normal (26-34 CGG repeats), low (<26 CGG repeats), and high (>34 CGG repeats) (Figure 3) [71,72].

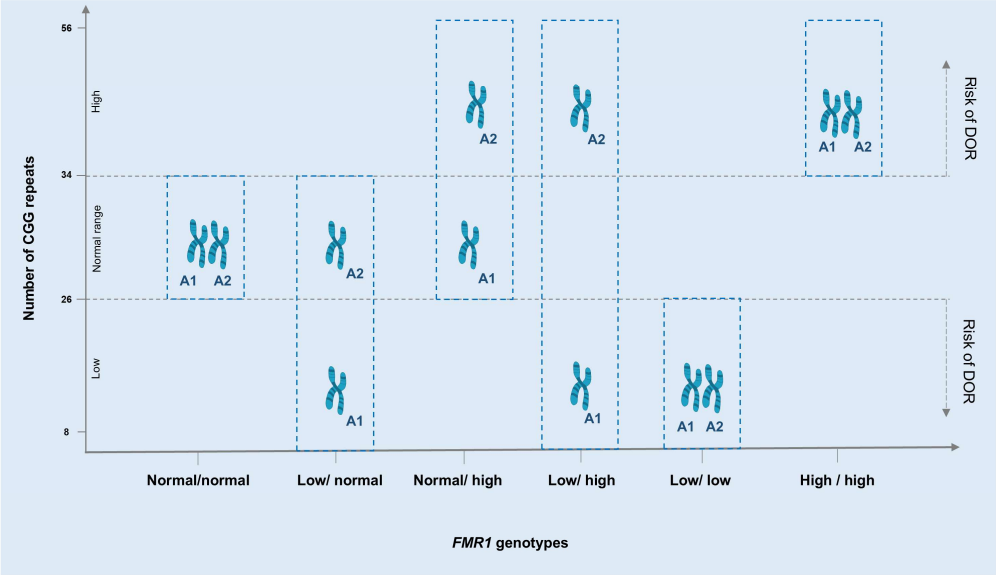


Figure 3. Schematic representation of *FMR1* genotypes and associated risk of diminished ovarian reserve (DOR). Females carrying *FMR1* alleles within the revised normal range (normal/normal genotype) were found to have a lower risk of DOR compared to those with one or both alleles exhibiting repeat numbers outside this range. A1 – Allele with shorter CGG repeat length; A2 – Allele with longer CGG repeat length. Figure based on data from Gleicher et al. (2010) [71].

A study reported a more pronounced decline in AMH levels in healthy females carrying one allele with fewer than 26 CGG repeats [73]. In a cohort of infertile females, lower AMH concentrations, indicative of DOR, were associated with *FMR1* alleles containing fewer than 26 or more than 34 CGG repeats. Notably, alleles with fewer than 26 CGG repeats were suggested to confer a slightly elevated risk of DOR compared to those with more than 34 CGG repeats [74]. This reduction in AMH levels may reflect impaired granulosa cell function and/or a diminished pool of developing follicles [46]. Wang et al. (2017) demonstrated decreased *FMR1* mRNA expression in females with low *FMR1*

alleles. This alteration can disrupt the regulation of steroidogenic enzymes and hormone receptors, ultimately contributing to ovarian dysfunction and infertility [75].

Some studies have implicated intermediate alleles of the *FMR1* gene in ovarian dysfunction and occult POI, i.e. menstrual cycles, but impaired ovarian response [76,77]. Nonetheless, the literature presents conflicting evidence regarding the impact of these alleles on ovarian function, with some investigations yielding contradictory results [78–80].

2.2.1. The impact of other CGG repeat lengths in reproductive outcomes

In addition to DOR/ premature ovarian senescence, *FMR1* genotypes have been implicated in a range of reproductive challenges, including decreased *in vitro* fertilization (IVF) pregnancy rates, suboptimal embryo quality, poor ovarian stimulation response, and occult POI [32,39,73,74,81–86]. Notably, alleles with fewer than 26 CGG repeats have been associated with the most pronounced adverse reproductive outcomes [32]. However, the literature regarding these alleles is inconsistent, and their impact on fertility remains inconclusive [87–95].

The controversial association between normal-sized alleles, particularly those with fewer than 26 CGG repeats, and DOR or other infertility phenotypes suggests that important attributes of the *FMR1* gene need further consideration. While it is established that normal alleles typically contain AGG interruptions, often positioned at the 9^{th} or 10^{th} CGG repeat [56]. The functional implications of these interruptions on fertility are unclear. Although the influence of AGG interruptions on fertility has been explored in females with the *FMR1* PM, their impact on carriers of normal-sized alleles has yet to be elucidated.

2.3. The importance of early screening of the *FMR1* repetitive tract

Given the potential health implications for female PM carriers, early identification can facilitate informed decision-making [96]. Female PM carriers may benefit from access to specialized care, including genetic counselling regarding reproductive risks, and the exploration of fertility preservation options such as oocyte retrieval, *in vitro* fertilization (IVF), and preimplantation genetic testing (PGT) [40]. Moreover, the observed association between alleles containing fewer than 26 CGG repeats (low alleles) and conditions like occult POI and reduced IVF success rates underscores the potential advantages of early screening for these individuals, enabling them to consider similar reproductive strategies as those offered to female PM carriers [32].

3. Overview of fertility treatments: conventional IVF and ICSI

Females undergoing ART for pregnancy, fertility preservation or oocyte donation typically follow standardized protocols encompassing controlled ovarian stimulation, oocyte retrieval, and subsequent fertilization or cryopreservation. Conventional IVF and Intracytoplasmic sperm injection (ICSI) are the primary ART procedures employed in fertility clinics. In recent years, the use of ICSI has become more prevalent compared to conventional IVF. While IVF has historically been the predominant method, ICSI has witnessed a substantial increase in utilization in recent years. Both techniques involve sequential steps including controlled ovarian hyperstimulation, follicular aspiration, oocyte retrieval, fertilization, embryo cultivation, and ultimately, embryo transfer. Preimplantation genetic testing (PGT) may be integrated into this process [97–99].

Both conventional IVF and ICSI exclusively utilize mature oocytes at the metaphase II (MII) stage [100]. The oocyte is encased by a layer of cumulus granulosa cells, collectively termed the cumulus-oocyte complex (COC) [101,102]. In conventional IVF, motile spermatozoa surround the COC and undergo acrosomal reactions, releasing hyaluronidase enzymes to penetrate the cumulus granulosa cells and corona radiata, ultimately accessing the zona pellucida (ZP) for fertilization. This process mimics natural fertilization [99,103]. Conversely, ICSI involves the removal of cumulus granulosa cells surrounding the oocyte, a procedure known as oocyte denudation (see Figure 1, Section 3, Chapter VI) [104]. Subsequently, a single spermatozoon is immobilized and injected directly into the ooplasm of the denuded oocyte to induce fertilization (Figure 4) [99,103].

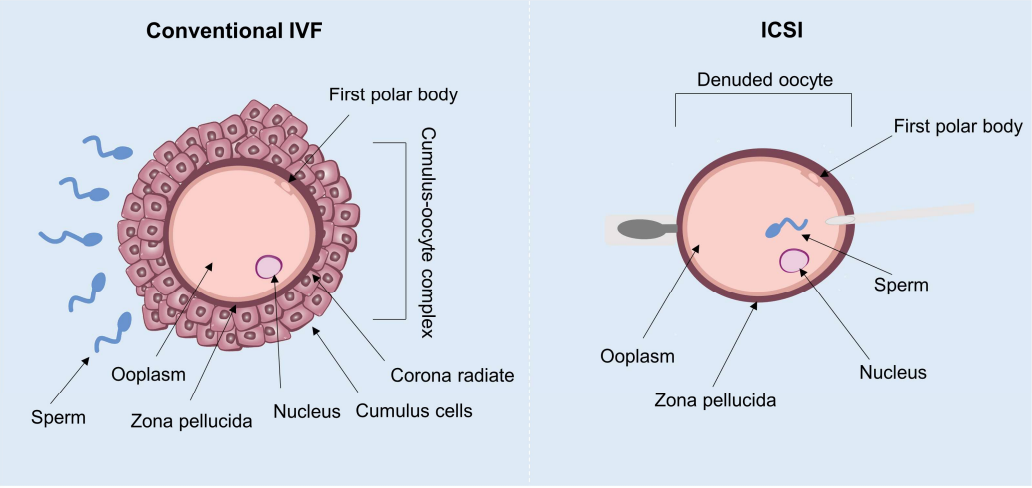


Figure 4. Schematic representation of *in vitro* fertilization (IVF) processes. A comparative illustration of fertilization processes in conventional IVF and intracytoplasmic sperm injection (ICSI). Figure based on data from Haas et al. (2021) [103].

Typically, successful fertilization is confirmed by the visualization of two distinct pronuclei (2PN) within the zygote, accompanied by the presence of two polar bodies, approximately 16 to 18 hours post-insemination [105,106]. Embryo transfer is conventionally conducted on the fifth day of development, coinciding with the blastocyst stage [107].

3.1. Granulosa cells

Cumulus granulosa cells (CCs) originated from undifferentiated granulosa cells (GCs) surround the oocyte in primordial and preantral follicles. As primary somatic cell type within the follicle, GCs provide critical physical support and establish the microenvironment essential for oocyte development [108]. During the preovulatory phase, GCs undergo differentiation into mural granulosa cells (MGCs) and CCs [101]. While MGs primarily contribute to the follicle wall structure and endocrine function, CCs establish direct contact with the oocyte, supporting its growth, and producing hyaluronic acid (HA) in response to follicle-stimulating hormone (FSH)-induced expansion (Figure 4) [109,110].

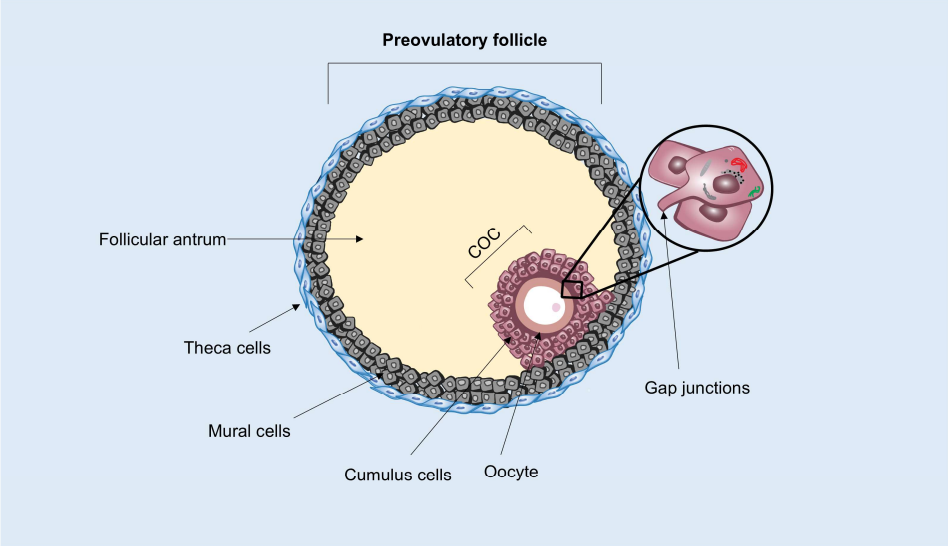


Figure 5. Schematic representation of the preovulatory follicle. A diagram depicting the structural organization of a preovulatory follicle. The follicle is characterized by distinct layers of granulosa cells, including mural granulosa cells forming the follicular wall and cumulus granulosa cells surrounding the oocyte. A fluid-filled antrum separates these two cell types. The follicular aspiration process involves the retrieval of the cumulus-oocyte complex (COC) along with the follicular fluid. Figure adapted from Xie et al. (2023) [101].

3.1.1. Cumulus cells and oocyte competence

CCs are intimately associated with the oocyte, forming the COC. Essential for oocyte growth and maturation within the follicular microenvironment, CCs also contribute to

meiotic completion, ovulation, and fertilization [109,111]. Bidirectional communications between the oocyte and CCs facilitated by gap junctions, is pivotal for normal oocyte development [112] (Figure 5). These gap junctions are cytoplasmic extensions that traverse the ZP, connecting the oocyte's ooplasm with the surrounding CCs [101,113].

3.1.1.1. Oocyte growth and maturation

The role of CCs in oocyte growth includes synchronizing nuclear and cytoplasmic maturation. They synchronize nuclear and cytoplasmic development, facilitating meiotic resumption through the provision of key signaling molecules, including cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) [109]. Gap junctional communication enables the transfer of essential signals and nutrients, from CCs to the oocyte, with CCs serving as a primary energy source via robust glycolytic activity [110–111]. Oocyte competence, reflecting its capacity for successful fertilization and embryonic development, is intricately linked to its maturation status [114,115]. Both nuclear and cytoplasmic maturation are prerequisites for successful fertilization [115–117].

Although nuclear and cytoplasmic maturation processes are temporally linked, cytoplasmic development generally lags behind nuclear progression. The morphological hallmarks of cytoplasmic maturation remain largely undefined. Conversely, the extrusion of the first polar body serves as a well-established marker of completed nuclear maturation [118–121]. Nuclear maturation comprises the resumption of meiosis from prophase I and advancement to metaphase II. Cytoplasmic maturation encompasses the developmental changes that render the oocyte competent for fertilization and subsequent embryonic development [121–123].

3.1.1.2. Cumulus-oocyte complex expansion, meiotic maturation, ovulation and fertilization

During oocyte maturation, the COC expansion, which involves the disruption of gap junctions between cells, is crucial for the progression of meiotic maturation. This process is driven the secretion of hyaluronic acid (HA) by CCs, which accumulates within the extracellular matrix surrounding the oocyte [109,124,125]. COC expansion is essential for ovulation and subsequent fertilization within the oviductal ampulla [124,126]. CCs contribute to the fertilization process by releasing factors, such as prostaglandins E1, E2, and F2α, which enhance sperm motility and the acrosome reaction. Additionally, CCs play a supportive role during early embryonic development, including cleavage and blastocyst formation [109]. Although sperm are indispensable for fertilization and

embryogenesis, the oocyte intrinsically possesses many of the cellular and molecular mechanisms required for these processes [115].

3.1.2. Cumulus cells and DNA damage for oocyte quality

The unique relationship between CCs and oocytes renders CCs invaluable for assessing oocyte competence. Compromised CCs integrity due to various pathogenic factors has been implicated in reduced fertilization rates and subsequent pregnancy failures [127–129].

Traditional methods of assessing oocyte maturity, such as COC expansion or follicular fluid hormone levels, have demonstrated limited reliability. In the context of ICSI, oocyte selection is primarily based on morphological criteria, which do not necessarily correlate with oocyte competence [130]. Oocyte competence requires the completion of both nuclear and cytoplasmic maturation, enabling subsequent fertilization and pronuclear formation [117]. Thus, it is imperative to emphasize that successful fertilization is contingent upon the preceding completion of oocyte maturation. During COC expansion, the CCs detach from the oocyte and undergo apoptosis, characterized by DNA fragmentation [131–134]. The integrity of CCs DNA is thought to influence oocyte maturation and subsequent fertilization. Previous investigations have explored the correlation between CCs DNA damage and oocyte quality, though these studies have been limited by the challenges associated with human CCs acquisition and inconsistent findings. Raman et al. (2001) reported a positive correlation between tail moment, a measure of DNA damage, in CCs and fertilization rates following ICSI in a cohort of infertile females. The authors suggested that apoptotic CCs, detached from the oocyte, could serve as a biomarker for oocyte competence. Conversely, immature oocytes are often surrounded by CCs exhibiting active growth and intact nuclear DNA [135]. In line with these results, Lourenço et al. (2014) demonstrated that CCs surrounding fertilized oocytes had elevated caspase activity compared with those surrounding oocytes that did not fertilize, suggesting a potential link to oocyte quality and oocyte fertilization competence [136]. Conversely, Lee at al. (2001) demonstrated a lower incidence of apoptosis in cumulus cells from fertilized oocytes compared to unfertilized ones [137]. A discrepant body of literature exists concerning the association between CCs apoptosis and oocyte developmental competence. Another investigation reported a negative correlation between CCs apoptosis and oocyte maturation, fertilization rate, and embryo development, suggesting that elevated levels of CCs apoptosis may compromise oocyte quality [138]. In contrast, other studies have failed to identify a significant relationship between CCs apoptosis and these reproductive outcomes [139,140]. The precise role of

CCs apoptosis in modulating oocyte quality remains to be fully elucidated, necessitating further research to clarify these inconsistencies.

3.1.3. *FMR1* function in granulosa cells

Elucidating the role of the *FMR1* gene in CCs function is imperative. In humans, *FMR1* is expressed in GCs. While the precise mechanisms remain to be fully elucidated, the pronounced expression of FMRP during ovarian folliculogenesis suggests a regulatory role in this process [141]. Female carrying *FMR1* PM, exhibit elevated *FMR1* mRNA levels within GCs, correlated with impaired follicular development and diminished ovarian reserve. A potential RNA toxic gain-of-function mechanism has been proposed, whereby increased *FMR1* mRNA may induce the formation of non-canonical RNA structures. These structures could sequester essential RNA-binding proteins, leading to compromised cellular integrity and accelerated follicular atresia [49]. Furthermore, elevated *FMR1* mRNA levels have been implicated in increased R-loop formation, a non-B DNA structure associated with genomic instability. This phenomenon may contribute to premature follicular degeneration in *FMR1* PM carriers [30,142]. The accumulation of *FMR1* mRNA-containing nuclear inclusions, termed FMRpolyG aggregates, within GCs of PM carriers has been reported [143]. Prior studies have linked the presence of such intranuclear inclusions in ovarian stromal cells to follicular atresia and accelerated ovarian reserve depletion [144].

A study investigating females with low/normal *FMR1* genotypes and poor ovarian response revealed elevated *FMR1* mRNA levels in GCs compared to other genotype groups. These increased mRNA levels, resembling the *FMR1* expression pattern observed in PM carriers, may contribute to altered FMRP protein levels and potentially exert toxic effects. The authors proposed that diminished ovarian responsiveness might reflect pre-existing impairments in folliculogenesis [82]. In contrast, Rehnitz et al. (2021) reported decreased *FMR1* expression in females with poor ovarian response relative to those with normal ovarian function. Additionally, the study identified significant differences in CpG94 methylation patterns in primary GCs between females with normal and compromised folliculogenesis, suggesting an epigenetic regulatory mechanism influencing *FMR1* expression and ovarian function [145]. Wang et al. (2018) demonstrated a correlation between reduced *FMR1* mRNA expression and the presence of low *FMR1* CGG repeat alleles in females. These females exhibited decreased expression of multiple *FMR1* transcript isoforms. The authors proposed that alterations in *FMR1* mRNA levels and isoform ratios might influence the translation and/or cellular localization of FMRP, consequently impacting the expression of steroidogenic enzymes

and hormone receptors or engaging in epigenetic mechanisms, thereby contributing to ovarian dysfunction and infertility [75].

While the precise mechanisms remain to be fully elucidated, it is evident that aberrant mRNA expression within GCs can perturb the regulatory networks governing follicular development and oocyte maturation. These dysregulated processes subsequently compromise oocyte competence, potentially leading to adverse reproductive outcomes. Consistent with this notion Rehnitz et al. (2021) demonstrated a correlation between *FMR1* expression and the AKT/mTOR signaling pathway, suggesting that *FMR1*-mediated alterations in gene expression may contribute to follicular maturation disturbances [146].

References

1. Eichler EE, Richards S, Gibbs RA, Nelson DL. Fine structure of the human FMR1 gene. Hum Mol Genet. 1993;2:1147–53.

2. Malecki C, Hambly BD, Jeremy RW, Robertson EN. The RNA-binding fragile-X mental retardation protein and its role beyond the brain. Biophys Rev. 2020;12:903–16.

3. Richter JD, Zhao X. The molecular biology of FMRP: new insights into fragile X syndrome. Nat Rev Neurosci. 2021;22:209–22.

4. Sidorov MS, Auerbach BD, Bear MF. Fragile X mental retardation protein and synaptic plasticity. Mol Brain. 2013;6:1.

5. Noto V, Harrity C, Walsh D, Marron K. The impact of FMR1 gene mutations on human reproduction and development: a systematic review. J Assist Reprod Genet. 2016;33:1135–47.

6. 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;23:417–25.

7. Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. Hum Mol Genet. 2000;9:901–8.

8. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Fertil Steril. 2017;108:393–406.

9. Gurunath S, Pandian Z, Anderson RA, Bhattacharya S. Defining infertility-a systematic review of prevalence studies. Hum Reprod Update. 2011;17:575–88.

10. WHO - World Health Organization. Infertility prevalence estimates 1990-2021. 2023.

11. Vander Borght M, Wyns C. Fertility and infertility: Definition and epidemiology. Clin Biochem. 2018;62:2–10.

12. W. Ombelet. WHO fact sheet on infertility gives hope to millions of infertile couples worldwide. Facts Views Vis Obgyn. 2020;12:249–51.

13. 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;22:1506–12.

14. Quaas A, Dokras A. Diagnosis and Treatment of Unexplained Infertility. Rev Obs Gynecol. 2008;1:69–76.

15. Shah K, Sivapalan G, Gibbons N, Tempest H, Griffin DK. The genetic basis of infertility. Reproduction. 2003;126:13–25.

16. Sharma R, Biedenharn KR, Fedor JM, Agarwal A. Lifestyle factors and reproductive health: Taking control of your fertility. Reprod Biol Endocrinol. 2013;11:1.

17. Bala R, Singh V, Rajender S. Environment, Lifestyle, and Female Infertility. Reprod Sci. 2021;28:617–38.

18. Gaware V., Parjane SK, Abhijit Mereka N, Pattan SR, Dighe NS, Kuchekar BS, et al. Female infertility and its treatment by alternative medicine: A review. J Chem Pharm Res. 2009;1:148–62.

19. Eniola W, Adetola A, Abayomi T. A review of Female Infertility; Important Etiological Factors and Management. J Microbiol Biotechnol Res. 2012;2:379–85.

20. Yatsenko SA, Rajkovic A. Genetics of human female infertility. Biol Reprod. 2019;101:549–66.

21. Venkatesh T, Suresh S P, Tsutsumi R. New insights into the genetic basis of infertility. Appl Clin Genet. 2014;1:235–43.

22. Carson SA, Kallen AN. Diagnosis and Management of Infertility: A Review. JAMA - J Am Med Assoc. 2021;326:65–76.

23. Mallepaly R, Butler PR, Herati AS, Lamb DJ. Genetic Basis of Male and Female Infertility. Monogr Hum Genet. 2017;21:1–16.

24. Munro MG, Balen AH, Cho S, Critchley HOD, Díaz I, Ferriani R, et al. The FIGO Ovulatory Disorders Classification System. Russ J Hum Reprod. 2023;29:116–36.

25. Luo W, Ke H, Tang S, Jiao X, Li Z, Zhao S, et al. Next-generation sequencing of 500 POI patients identified novel responsible monogenic and oligogenic variants. J Ovarian Res. 2023;16:1–13.

26. Chon SJ, Umair Z, Yoon MS. Premature Ovarian Insufficiency: Past, Present, and Future. Front Cell Dev Biol. 2021;9:1–13.

27. Fortuño C, Labarta E. Genetics of primary ovarian insufficiency: a review. J Assist Reprod Genet. 2014;31:1573–85.

28. Huhtaniemi I, Hovatta O, La Marca A, Livera G, Monniaux D, Persani L, et al. Advances in the Molecular Pathophysiology, Genetics, and Treatment of Primary Ovarian Insufficiency. Trends Endocrinol Metab. 2018;29:400–19.

29. Quilichini J, Perol S, Cuisset L, Grotto S, Fouveaut C, Barbot JC, et al. Stratification of the risk of ovarian dysfunction by studying the complexity of intermediate and premutation alleles of the FMR1 gene. Am J Med Genet Part A. 2024;194:1–10.

30. 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.

31. Barasoain M, Barrenetxea G, Huerta I, Télez M, Criado B, Arrieta I. Study of the genetic etiology of primary ovarian insufficiency: FMR1 gene. Genes (Basel). 2016;7:1–18.

32. 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.

33. Seltzer MM, Baker MW, Hong J, Maenner M, Greenberg J, Mandel D. Prevalence of CGG expansions of the FMR1 gene in a US population-based sample. Am J Med Genet Part B Neuropsychiatr Genet. 2012;159 B:589–97.

34. Ma Y, Wei X, Pan H, Wang S, Wang X, Liu X, et al. The prevalence of CGG repeat expansion mutation in FMR1 gene in the northern Chinese women of reproductive age. BMC Med Genet. 2019;20:1–5.

35. Sherman SL. Premature ovarian failure in the fragile X syndrome. Am J Med Genet - Semin Med Genet. 2000;97:189–94.

36. 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;10:1–6.

37. Welt CK. Primary ovarian insufficiency: A more accurate term for premature ovarian failure. Clin Endocrinol (Oxf). 2008;68:499–509.

38. Allen EG, Charen K, Hipp HS, Shubeck L, Amin A, He W, et al. Clustering of comorbid conditions among women who carry an FMR1 premutation. Genet Med. 2020;22:758–66.

39. Fink DA, Nelson LM, Pyeritz R, Johnson J, Sherman SL, Cohen Y, et al. Fragile X Associated Primary Ovarian Insufficiency (FXPOI): Case Report and Literature Review. Front Genet. 2018;9:1–12.

40. Tassone F, Protic D, Allen EG, Archibald AD, Baud A, Brown TW, et al. Insight and Recommendations for Fragile X-Premutation-Associated Conditions from the Fifth International Conference on FMR1 Premutation. Cells. 2023;12:2330.

41. Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, et al. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med. 2012;14:729–36.

42. Stoyanova M, Hachmeriyan M, Levkova M, Bichev S, Georgieva M, Mladenov V, et al. Molecular Screening for Fragile X Syndrome in Children with Unexplained Intellectual Disability and/or Autistic Behaviour. Folia Med (Plovdiv). 2022;64:27–32.

43. Bibi G, Malcov M, Yuval Y, Reches A, Ben-Yosef D, Almog B, et al. The effect of CGG repeat number on ovarian response among fragile X premutation carriers undergoing preimplantation genetic diagnosis. Fertil Steril. 2010;94:869–74.

44. Allen EG, Sullivan AK, Marcus M, Small C, Dominguez C, Epstein MP, et al. Examination of reproductive aging milestones among women who carry the FMR1 premutation. Hum Reprod. 2007;22:2142–52.

45. Rohr J, Allen EG, Charen K, Giles J, He W, Dominguez C, et al. Anti-Mullerian hormone indicates early ovarian decline in fragile X mental retardation (FMR1) premutation carriers: A preliminary study. Hum Reprod. 2008;23:1220–5.

46. Lekovich J, Man L, Xu K, Canon C, Lilienthal D, Stewart JD, et al. CGG repeat length and AGG interruptions as indicators of fragile X–associated diminished ovarian reserve. Genet Med. 2018;20:957–64.

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

48. Tassone F, Beilina A, Carosi C, Albertosi S, Bagni C, Li L, et al. Elevated FMR1 mRNA in premutation carriers is due to increased transcription. Rna. 2007;13:555–62.

49. Elizur SE, Lebovitz O, Derech-Haim S, Dratviman-Storobinsky O, Feldman B, Dor J, et al. Elevated Levels of FMR1 mRNA in Granulosa Cells Are Associated with Low Ovarian Reserve in FMR1 Premutation Carriers. Bardoni B, editor. PLoS One. 2014;9:e105121–e105121.

50. Mailick MR, Hong J, Greenberg J, Smith L, Sherman S. Curvilinear Association of CGG Repeats and Age at Menopause in Women with FMR1 Premutation Expansions. Am J Med Genet B Neuropsychiatr Genet. 2014;0:705–11.

51. Ennis S, Ward D, Murray A. Nonlinear association between CGG repeat number and age of menopause in FMR1 premutation carriers. Eur J Hum Genet. 2006;14:253–5.

52. Allen EG, Charen K, Hipp HS, Shubeck L, Amin A, He W, et al. Refining the risk for fragile X–associated primary ovarian insufficiency (FXPOI) by FMR1 CGG repeat size. Genet Med [. 2021;23:1648–55.

53. Sherman SL, Curnow EC, Easley CA, Jin P, Hukema RK, Tejada MI, et al. Use of model systems to understand the etiology of fragile X-associated primary ovarian insufficiency (FXPOI). J Neurodev Disord. 2014;6:1–12.

54. García-Alegría E, Ibáñez B, Mínguez M, Poch M, Valiente A, Sanz-Parra A, et al. Analysis of FMR1 gene expression in female premutation carriers using robust segmented linear regression models. RNA. 2007;13:756–62.

55. Spath MA, Nillesen WN, Smits APT, Feuth TB, Braat DDM, Van Kessel AG, et al. X chromosome inactivation does not define the development of premature ovarian failure in fragile X premutation carriers. Am J Med Genet Part A. 2010;152:387–93.

56. Nolin SL, Glicksman A, Tortora N, Allen E, Macpherson J, Mila M, et al. Expansions and contractions of the FMR1 CGG repeat in 5,508 transmissions of normal, intermediate, and premutation alleles. Am J Med Genet Part A. 2019;179:1148–56.

57. McGinty RJ, Mirkin SM. Cis- and Trans-Modifiers of Repeat Expansions: Blending Model Systems with Human Genetics. Trends Genet. 2018;34:448–65.

58. Pearson CE, Eichler EE, Lorenzetti D, Kramer SF, Zoghbi HY, Nelson DL, et al. Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. Biochemistry. 1998;37:2701–8.

59. Yrigollen CM, Martorell L, Durbin-Johnson B, Naudo M, Genoves J, Murgia A, et al. AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. J Neurodev Disord. 2014;6:1–11.

60. Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, et al. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med. 2012;14:729–36.

61. Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, et al. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med. 2012;14:729–36.

62. Villate O, Ibarluzea N, Maortua H, de la Hoz AB, Rodriguez-Revenga L, Izquierdo-Álvarez S, et al. Effect of AGG Interruptions on FMR1 Maternal Transmissions. Front Mol Biosci. 2020;7:1–6.

63. Nolin SL, Glicksman A, Ersalesi N, Dobkin C, Brown WT, Cao R, et al. Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers. Genet Med. 2015;17:358–64.

64. Eichler EE, Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeu SN, et al. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. Nat Genet. 1994;8:88–94.

65. Allen EG, Glicksman A, Tortora N, Charen K, He W, Amin A, et al. FXPOI: Pattern of AGG interruptions does not show an association with age at amenorrhea among women with a premutation. Front Genet. 2018;9:1–7.

66. Friedman-Gohas M, Kirshenbaum M, Michaeli A, Domniz N, Elizur S, Raanani H, et al. Does the presence of AGG interruptions within the CGG repeat tract have a protective effect on the fertility phenotype of female FMR1 premutation carriers? J Assist Reprod Genet. 2020;37:849–54.

67. 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;31:1295–301.

68. Eslami H, Eslami A, Favaedi R, Asadpour U, Moradi SZ, Eftekhari-Yazdi P, et al. Epigenetic Aberration of FMR1 Gene in Infertile Women with Diminished Ovarian Reserve. Cell J. 2018;20:78–83.

69. Pastore LM, Christianson MS, Stelling J, Kearns WG, Segars JH. Reproductive ovarian testing and the alphabet soup of diagnoses: DOR, POI, POF, POR, and FOR. J Assist Reprod Genet. 2018;35:17–23.

70. Gleicher N, Weghofer A, Barad DH. Defining ovarian reserve to better understand ovarian aging. Reprod Biol Endocrinol. 2011;9:23.

71. Gleicher N, Weghofer A, Barad DH. Ovarian reserve determinations suggest new function of FMR1 (fragile X gene) in regulating ovarian ageing. Reprod Biomed Online. 2010;20:768–75.

72. Gleicher N, Weghofer A, Lee IH, Barad DH. FMR1 genotype with autoimmunity-associated polycystic ovary-like phenotype and decreased pregnancy chance. PLoS One. 2010;5:1–6.

73. 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;166:502–7.

74. 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;88:1024–30.

75. Wang Q, Kushnir VA, Darmon S, Barad DH, Wu Y, Zhang L, et al. Reduced RNA expression of the FMR1 gene in women with low (CGGn<26) repeats. Fertil Steril. 2017;108:e143.

76. Barasoain M, Barrenetxea G, Huerta I, Télez M, Carrillo A, Pérez C, et al. Study of FMR1 gene association with ovarian dysfunction in a sample from the Basque Country. Gene. 2013;521:145–9.

77. 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;92:464–70.

78. 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;25:1335–8.

79. 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;21:740–8.

80. Lledo B, Guerrero J, Ortiz JA, Morales R, Ten J, Llacer J, et al. Intermediate and normal sized CGG repeat on the FMR1 gene does not negatively affect donor ovarian response. Hum Reprod. 2012;27:609–14.

81. Kushnir VA, Yu Y, Barad DH, Weghofer A, Himaya E, Lee HJ, et al. Utilizing FMR1 gene mutations as predictors of treatment success in human in vitro fertilization. PLoS One. 2014;9:1–9.

82. Rehnitz J, Alcoba DD, Brum IS, Dietrich JE, Youness B, Hinderhofer K, et al. FMR1 expression in human granulosa cells increases with exon 1 CGG repeat length depending on ovarian reserve. Reprod Biol Endocrinol. 2018;16:1–9.

83. Gleicher N, Weghofer A, Oktay K, Barad DH. Can the FMR1 (Fragile X) Gene Serve As Predictor of Response to Ovarian Stimulation? Reprod Sci. 2009;16:462–7.

84. Lu CL, Li R, Chen XN, Xu YY, Yan LY, Yan J, et al. The 'normal' range of FMR1 triple CGG repeats may be associated with primary ovarian insufficiency in China. Reprod Biomed Online. 2017;34:175–80.

85. Jin X, Zeng W, Xu Y, Jin P, Dong M. Cytosine–guanine–guanine repeats of FMR1 gene negatively affect ovarian reserve and response in Chinese women. Reprod Biomed Online. 2023;49:103779.

86. 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;117:376–82.

87. Maslow BSL, Davis S, Engmann L, Nulsen JC, Benadiva CA. Correlation of normal-range FMR1 repeat length or genotypes and reproductive parameters. J Assist Reprod Genet. 2016;33:1149–55.

88. 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;31:2396–403.

89. Morin SJ, Tiegs AW, Franasiak JM, Juneau CR, Hong KH, Werner MD, et al. FMR1 gene CGG repeat variation within the normal range is not predictive of ovarian response in IVF cycles. Reprod Biomed Online. 2016;32:496–502.

90. Cogendez E, Ozkaya E, Eser AÇ, Eken M, Karaman A. Can FMR1 CGG repeat lengths predict the outcome in ICSI cycles? Ginekol Pol. 2022;93:735–41.

91. 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;02:10–3.

92. Schufreider A, McQueen DB, Lee SM, Allon R, Uhler ML, Davie J, et al. 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;30:2686–92.

93. Mailick MR, Hong J, Rathouz P, Baker MW, Greenberg JS, Smith L, et al. Low-normal FMR1 CGG repeat length : phenotypic associations. Front Genet. 2014;5:1–9.

94. Nunes ACV, Trevisan CM, Peluso C, Loureiro FA, Dias AT, Rincon D, et al. Low and High-Normal FMR1 Triplet Cytosine, Guanine Guanine Repeats Affect Ovarian Reserve

and Fertility in Women Who Underwent In Vitro Fertilization Treatment? Results from a Cross-Sectional Study . DNA Cell Biol. 2024;43:414–24.

95. Pastore LM, Christianson MS, McGuinness B, Vaught KC, Maher JY, Kearns WG. Does theFMR1 gene affect IVF success? Reprod Biomed Online. 2019;38:560–9.

96. Le Poulennec T, Dubreuil S, Grynberg M, Chabbert-Buffet N, Sermondade N, Fourati S, et al. Ovarian reserve in patients with FMR1 gene premutation and the role of fertility preservation. Ann Endocrinol (Paris). 2024;85:269–75.

97. Farquhar C, Marjoribanks J. Assisted reproductive technology: An overview of Cochrane Reviews. Cochrane Database Syst Rev. 2018;2018.

98. Esteves SC, Humaidan P, Roque M, Agarwal A. Female infertility and assisted reproductive technology. Panminerva Med. 2019;61:1–2.

99. Chamayou S, Giacone F, Cannarella R, Guglielmino A. What Does Intracytoplasmic Sperm Injection Change in Embryonic Development? The Spermatozoon Contribution. J Clin Med. 2023;12.

100. Carles M, Lefranc E, Bosquet D, Capelle S, Scheffler F, Copin H, et al. In vitro maturation of oocytes from stimulated IVF-ICSI cycles using autologous cumulus cell co-culture: A preliminary study. Morphologie. 2023;107:28–37.

101. Xie J, Xu X, Liu S. Intercellular communication in the cumulus–oocyte complex during folliculogenesis: A review. Front Cell Dev Biol. 2023;11:1–13.

102. Marchais M, Gilbert I, Bastien A, Macaulay A, Robert C. Mammalian cumulus-oocyte complex communication: a dialog through long and short distance messaging. J Assist Reprod Genet. 2022;39:1011–25.

103. Haas J, Miller TE, Nahum R, Aizer A, Kirshenbaum M, Zilberberg E, et al. The role of ICSI vs. conventional IVF for patients with advanced maternal age—a randomized controlled trial. J Assist Reprod Genet. 2021;38:95–100.

104. de Moura BRL, Gurgel MCA, Machado SPP, Marques PA, Rolim JR, de Lima MC, et al. Low concentration of hyaluronidase for oocyte denudation can improve fertilization rates and embryo quality. J Bras Reprod Assist. 2017;21:27–30.

105. Zhu M, Dong Q, Zhu Y, Le Y, Wang T, Zhou Y, et al. Developmental potential of non- and mono-pronuclear zygotes and associated clinical outcomes in IVF cycles. Front Endocrinol (Lausanne). 2024;15:1–8.

106. Holubcová Z, Kyjovská D, Martonová M, Páralová D, Klenková T, Otevřel P, et al. Egg maturity assessment prior to ICSI prevents premature fertilization of late-maturing oocytes. J Assist Reprod Genet. 2019;36:445–52.

107. Zhang G li, Sun T yi, Li S, Jiang M xi, Guo L. The pregnancy outcomes of day-5 poor-quality and day-6 high-quality blastocysts in single blastocyst transfer cycles. Clin Exp Reprod Med. 2023;50:63–8.

108. Huang Z, Wells D. The human oocyte and cumulus cells relationship: New insights from the cumulus cell transcriptome. Mol Hum Reprod. 2010;16:715–25.

109. Turathum B, Gao EM, Chian RC. The function of cumulus cells in oocyte growth and maturation and in subsequent ovulation and fertilization. Cells. 2021;10:1–18.

110. Yokoo M, Sato E. Cumulus-oocyte complex interactions during oocyte maturation. Int Rev Cytol. 2004;235:251–91.

111. Martinez CA, Rizos D, Rodriguez-Martinez H, Funahashi H. Oocyte-cumulus cells crosstalk: New comparative insights. Theriogenology. 2023;205:87–93.

112. Baena V, Terasaki M. Three-dimensional organization of transzonal projections and other cytoplasmic extensions in the mouse ovarian follicle. Sci Rep. 2019;9:1–13.

113. Kidder GM, Mhawi AA. Gap junctions and ovarian folliculogenesis. Reproduction. 2002;123:613–20.

114. Su YQ, Sugiura K, Eppig JJ. Mouse oocyte control of granulosa cell development and function: Paracrine regulation of cumulus cell metabolism. Semin Reprod Med. 2009;27:32–42.

115. Reader KL, Stanton JAL, Juengel JL. The role of oocyte organelles in determining developmental competence. Biology (Basel). 2017;6:1–22.

116. Mobarak H, Heidarpour M, Tsai PSJ, Rezabakhsh A, Rahbarghazi R, Nouri M, et al. Autologous mitochondrial microinjection; A strategy to improve the oocyte quality and subsequent reproductive outcome during aging. Cell Biosci. 2019;9:1–15.

117. Eppig JJ, O'Brien M, Wigglesworth K. Mammalian oocyte growth and development in vitro. Mol Reprod Dev. 1996;44:260–73.

118. Rienzi L, Balaban B, Ebner T, Mandelbaum J. The oocyte. Hum Reprod. 2012;27:2–22.

119. Swain JE, Pool TB. ART failure: Oocyte contributions to unsuccessful fertilization. Hum Reprod Update. 2008;14:431–46.

120. Lu C, Zhang Y, Zheng X, Song X, Yang R, Yan J, et al. Current perspectives on in vitro maturation and its effects on oocyte genetic and epigenetic profiles. Sci China Life Sci. 2018;61:633–43.

121. Trebichalská Z, Kyjovská D, Kloudová S, Otevřel P, Hampl A, Holubcová Z. Cytoplasmic maturation in human oocytes: An ultrastructural study. Biol Reprod. 2021;104:106–16.

122. Zhou CJ, Wu SN, Shen JP, Wang DH, Kong XW, Lu A, et al. The beneficial effects of cumulus cells and oocyte-cumulus cell gap junctions depends on oocyte maturation and fertilization methods in mice. PeerJ. 2016;2016.

123. Sirait B, Wiweko B, Jusuf AA, Iftitah D, Muharam R. Oocyte Competence Biomarkers Associated With Oocyte Maturation: A Review. Front Cell Dev Biol. 2021;9.

124. Richards JS. Ovulation: New factors that prepare the oocyte for fertilization. Mol Cell Endocrinol. 2005;234:75–9.

125. Liu Y-X. Interaction and signal transduction between oocyte and samatic cells in the ovary. Front Biosci. 2005;3172–82.

126. Zhuo L, Kimata K. Cumulus oophorus extracellular matrix: Its construction and regulation. Cell Struct Funct. 2001;26:189–96.

127. Dumesic DA, Meldrum DR, Katz-Jaffe MG, Krisher RL, Schoolcraft WB. Oocyte environment: Follicular fluid and cumulus cells are critical for oocyte health. Fertil Steril. 2015;103:303–16.

128. Da Broi MG, Giorgi VSI, Wang F, Keefe DL, Albertini D, Navarro PA. Influence of follicular fluid and cumulus cells on oocyte quality: Clinical implications. J Assist Reprod Genet. 2018;35:735–51.

129. Kong P, Yin M, Tang C, Zhu X, Bukulmez O, Chen M, et al. Effects of Early Cumulus Cell Removal on Treatment Outcomes in Patients Undergoing In Vitro Fertilization: A Retrospective Cohort Study. Front Endocrinol (Lausanne). 2021;12:1–8.

130. Rattanachaiyanont M, Leader A, Léveillé MC. Lack of correlation between oocyte-corona-cumulus complex morphology and nuclear maturity of oocytes collected in stimulated cycles for intracytoplasmic sperm injection. Fertil Steril. 1999;71:937–40.

131. Barcena P, López-Fernández C, García-Ochoa C, Obradors A, Vernaeve V, Gosálvez J, et al. Detection of DNA damage in cumulus cells using a chromatin dispersion assay. Syst Biol Reprod Med. 2015;61:277–85.

132. Di Giacomo M, Camaioni A, Klinger FG, Bonfiglio R, Salustri A. Cyclic AMP-elevating agents promote cumulus cell survival and hyaluronan matrix stability, thereby prolonging the time of mouse oocyte fertilizability. J Biol Chem. 2016;291:3821–36.

133. Luciano AM, Modina S, Vassena R, Milanesi E, Lauria A, Gandolfi F. Role of Intracellular Cyclic Adenosine 3′,5′-Monophosphate Concentration and Oocyte-Cumulus Cells Communications on the Acquisition of the Developmental Competence during In Vitro Maturation of Bovine Oocyte. Biol Reprod. 2004;70:465–72.

134. Majtnerová P, Roušar T. An overview of apoptosis assays detecting DNA fragmentation. Mol Biol Rep. 2018;45:1469–78.

135. Raman RS, Chan PJ, Corselli JU, Patton WC, Jacobson JD, Chan SR, et al. Comet assay of cumulus cell DNA status and the relationship to oocyte fertilization via intracytoplasmic sperm injection. Hum Reprod. 2001;16:831–5.

136. Lourenço B, Sousa AP, Almeida-Santos T, Ramalho-Santos J. Relation of cumulus cell status with single oocyte maturity, fertilization capability and patient age. J Reprod Infertil. 2014;15:15–21.

137. Lee KS, Joo BS, Na YJ, Yoon MS, Choi OH, Kim WW. Cumulus cells apoptosis as an indicator to predict the quality of oocytes and the outcome of IVF-ET. J Assist Reprod Genet. 2001;18:490–8.

138. Høst E, Mikkelsen AL, Lindenberg S, Smidt-Jensen S. Apoptosis in human cumulus cells in relation to maturation stage and cleavage of the corresponding oocyte. Acta Obstet Gynecol Scand. 2000;79:936–40.

139. Abu-Hassan D, Koester F, Shoepper B, Schultze-Mosgau A, Asimakopoulos B, Diedrich K, et al. Comet assay of cumulus cells and spermatozoa DNA status, and the relationship to oocyte fertilization and embryo quality following ICSI. Reprod Biomed Online. 2006;12:447–52.

140. Tola EN, Koşar PA, Karatopuk DU, Sancer O, Oral B. Effect of DNA damage of cumulus oophorus cells and lymphocytes analyzed by alkaline comet assay on oocyte quality and intracytoplasmic sperm injection success among patients with polycystic ovary syndrome. J Obstet Gynaecol Res. 2018;45:609–18.

141. Schuettler J, Peng Z, Zimmer J, Sinn P, Von Hagens C, Strowitzki T, et al. Variable expression of the Fragile X Mental Retardation 1 (FMR1) gene in patients with premature ovarian failure syndrome is not dependent on number of (CGG)n triplets in exon 1. Hum Reprod. 2011;26:1241–51.

142. Rosario R, Anderson R. The molecular mechanisms that underlie fragile X-associated premature ovarian insufficiency: is it RNA or protein based? Mol Hum Reprod. 2020;26:727–37.

143. Friedman-Gohas M, Elizur SE, Dratviman-Storobinsky O, Aizer A, Haas J, Raanani H, et al. FMRpolyG accumulates in FMR1 premutation granulosa cells. J Ovarian Res. 2020;13:1–10.

144. Buijsen RAM, Visser JA, Kramer P, Severijnen EAWFM, Gearing M, Charlet-Berguerand N, et al. Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency. Hum Reprod. 2016;31:158–68.

145. Rehnitz J, Youness B, Nguyen XP, Dietrich JE, Roesner S, Messmer B, et al. FMR1 expression in human granulosa cells and variable ovarian response: control by epigenetic mechanisms. Mol Hum Reprod. 2021;27:1–12.

146. Rehnitz J, Capp E, Messmer B, Nguyen XP, Germeyer A, Freis A, et al. Fmr1 and akt/mtor signaling in human granulosa cells: Functional interaction and impact on ovarian response. J Clin Med. 2021;10:1–12.

2

Purpose and research objectives

General objective:

This study aims to elucidate the underlying mechanisms by which the *FMR1* gene and its repetitive tract influence reproductive function. The findings from this research will contribute to a deeper understanding of female infertility and provide valuable clinical insights for diagnosis, prognosis, and potential therapeutic interventions.

Specific objectives:

1. Characterization of the *FMR1* CGG/AGG repetitive tract

- Develop a novel allelic complexity score that comprehensively assesses the *FMR1* repetitive tract, incorporating both total CGG repeat length and number and pattern AGG interruptions (Chapter I).
- Investigate the methylation status of the *FMR1* gene to elucidate X-chromosome inactivation (XCI) patterns by comparing the methylation status of the *FMR1* gene with that of the *AR* gene (Chapter V).

2. Identification of risk markers for female infertility

- Explore the association between specific allelic complexity combinations and the risk of developing fragile X-associated primary ovarian insufficiency (FXPOI). By investigating the relationship between *allelic scores* and the age of amenorrhea (Chapter II).
- Assess the predictive value of allelic complexity in determining ovarian reserve and *in vitro* fertilization (IVF) success by correlating *allelic scores* with ovarian reserve markers and with IVF outcomes in females undergoing intracytoplasmic sperm injection (ICSI) (Chapter III).
- Evaluate the relationship between DNA damage in cumulus cells (CCs) and oocyte competence for fertilization by exploring the correlation between levels of DNA damage and fertility outcomes and investigating the use of whole blood as a surrogate marker for DNA damage in cumulus cells (chapter IV).

3. Establishment of a biorepository and associated data (chapter VI)

- Create a comprehensive biorepository encompassing a diverse range of biological samples, including peripheral blood, plasma, follicular fluid, and cumulus cells, obtained from consenting participants.
- Establish robust protocols for the collection, processing, and cryopreservation of biological samples. Employ optimized DNA and RNA extraction methods to ensure high-quality genetic material for downstream analyses.

- Develop a robust database integrating clinical, laboratory, and sample processing data.

3

Results

Chapter I. Development and validation of a mathematical model predict the complexity of *FMR1* allele combinations

Chapter II. *FMR1* allelic complexity in premutation carriers provides no evidence for a correlation with age at amenorrhea

Chapter III. Exploring of the predictive value of the *FMR1* gene allelic complexity for *in vitro* fertilization success

Chapter IV. Cumulus cell DNA damage linked to fertilization success in females with an ovulatory dysfunction phenotype

Chapter V. Use of the *FMR1* gene methylation status to assess the X-Chromosome Inactivation pattern: a stepwise analysis

Chapter VI. Experimental and dataset resources

Chapter I. Development and validation of a mathematical model predict the complexity of *FMR1* allele combinations[1]

Abstract

The polymorphic trinucleotide repetitive region in the *FMR1* gene 5'UTR contains AGG interspersions within the CGG repeat tract, particularly in normal-sized alleles (CGG < 45). In this range repetitive stretches are typically interrupted once or twice, although alleles without or with three or more AGG interspersions can also be observed. AGG interspersions together with the total length of the repetitive region confer stability and hinder expansion to pathogenic ranges: either premutation (55 ≤ CGG ≤ 200) or full mutation (CGG > 200). The AGG interspersions have long been identified as one of the most important features of *FMR1* repeat stability, being particularly important to determine expansion risk estimates in female premutation carriers. We sought to compute the combined AGG interspersion numbers and patterns, aiming to define *FMR1* repetitive tract complexity combinations. A mathematical model, the first to compute this cumulative effect, was developed and validated using data from 131 young and healthy females. Plotting of their allelic complexity enabled the identification of two statistically distinct groups – *equivalent* and *dissimilar* allelic combinations. The outcome, a numerical parameter designated *allelic score*, depicts the repeat substructure of each allele, measuring the allelic complexity of the *FMR1* gene including the AGGs burden, thus allowing new behavioral scrutiny of normal-sized alleles in females.

Keywords: *FMR1* gene, CGG repeats, AGG interspersion pattern, modelling allelic complexity, *allelic score*

[1]Rodrigues *et al*., Development and Validation of a Mathematical Model to Predict the Complexity of *FMR1* Allele Combinations. Front Genet. 2020 Nov 13:11:557147.

1. Introduction

The fragile X-related disorders result from the expansion of a CGG-repeat tract in the 5' untranslated region of the *FMR1* gene (Xq27.3), coding for the fragile X mental retardation protein (FMRP), an RNA-binding protein that regulates expression of several genes [1]. Depending on the number of CGG repeats, *FMR1* alleles can be categorized into four classes: normal (CGG < 45), intermediate or "gray zone" (45 ≤ CGG ≤ 54), premutation (55 ≤ CGG ≤ 200), and full mutation (CGG > 200) [2]. Premutations causing *FMR1* mRNA overexpression and reduced FMRP synthesis underly both fragile X-associated tremor/ataxia syndrome (FXTAS, OMIM #300623) and fragile X-associated primary ovarian insufficiency (FXPOI, OMIM #311360). The full mutation alleles undergo hypermethylation, leading to gene silencing and absence of FMRP, causing fragile X syndrome (FXS, OMIM #300624), the most common heritable cause of intellectual disability [1]. Due to the repeat tract instability above a threshold, expansions and contractions can be observed both in the germline and in the somatic cells. Some rare contraction events can originate mosaicism with mutated and normal alleles in clinically typical fragile-X phenotypes [3]. In the normal population, the vast majority of the alleles contain one or more AGG interspersions within the repetitive tract, usually at every 9 or 10 CGG repeat intervals, being highly stable. In higher repeat ranges, the number of AGGs tends to be progressively smaller as the size of the repetitive tract increases [4–6]. The AGG interspersions together with the repetitive region's total length confer stability and hinder expansion to pathogenic size-ranges [7–9]. In premutation female carriers, the risk of having a child with FXS depends on both the repeat length and AGG interspersions [10]. The incidence of normal pure alleles (without interspersions) is low and their origin as well as the phenotypic impact in females, are still debatable. It has been proposed that "low zone" alleles, with a (CGG ≤ 26 or CGG ≤ 23 according to different studies, are associated with different phenotypic outcomes [11–13]. Some studies show that they are associated with decreased ovarian reserve and fertility issues, due to a mechanism not yet elucidated, possibly different from that involved in premutated alleles [12,14], although such negative effects were not corroborated by others [15,16]. These contradictory assumptions require further studies to elucidate the clinical impact of "low zone" alleles.

Few studies focus on AGG interspersion patterns to assess allele stability, within the normal range. Given the importance of understanding the cumulative effect of the CGG repeat tract length and its AGG interspersions, we developed a mathematical model that

considers these patterns and produces a functional model predicting the complexity of allele combinations (*allelic score*).

2. Material and Methods

2.1. Study Population

Young and potentially fertile females were recruited among candidates for oocyte donation at the Portuguese Public Gamete Bank, Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), Centro Hospitalar Universitário do Porto (CHUP). The donor population, originating from the entire national territory, includes actively recruited students from major Portuguese universities, with a wide range of nationalities. Around 10% of the donor candidates were of foreign nationality, 95% were Caucasian and about 30% of those who donated at our center lived outside Porto [17]. Two independent cohorts were used for development (cohort 1) and for validation (cohort 2) studies. Cohort 1, n = 50, mean age 25.4 ± 3.93 years (range 18 to 33), recruited between 2016 and 2017. Cohort 2, n = 81, mean age 26.5 ± 3.86 years (range 19 to 33), collected between 2018 and 2019. All participants provided written informed consent, and this project was approved by the hospital's Ethics Committee (2018.231/201-DEFI/200-CES).

2.2. *FMR1* repeat region substructure profile

Sizing of *FMR1* alleles had been previously obtained as part of the routine oocyte donor's protocol, on blood samples. Categorizing the respective genotype followed the ACMG/EMQN guidelines: normal (CGG < 45), intermediate or "gray zone" (45 ≤ CGG ≤ 54), premutation (55 ≤ CGG ≤ 200), and full mutation (CGG > 200) [2,18]. AGG interspersion pattern was determined by Triplet Repeat Primed-PCR using FRAXA PCR kit LabGscan™ (Diagnostica Longwood, Zaragoza, Spain), according to the manufacturer's instructions. This method allowed the confirmation of the total repeat length and the characterization of the CGG/AGG substructure. Thirteen samples with different patterns were additionally verified by Sanger sequencing to confirm the previously determined CGG/AGG pattern.

2.3. Statistical analysis

Hierarchical Cluster Analysis using euclidean distance as a metric to evaluate similarity was used in statistical software SPSS® version 26 (IBM developer, 2019: SPSS Statistics version 26 - Armonk, New York, USA). Linear regression of the linearised form of an exponential model [i.e. regression of ln(score 2) against score 1] was used to obtain

a functional model to relate the complexity of both alleles in each sample. The analysis of covariance (ANCOVA), as outlined by Zar [19], was used to compare the regression models, and derive common regression lines, with *allelic scores* as variables [i.e. score 1 and ln(score 2)]. All statistical tests were carried out for a significance level of 0.05.

2.4. Determination of X-chromosome inactivation pattern and *FMR1* methylation status

X-chromosome inactivation (XCI) pattern was determined by the human androgen-receptor assay (HUMARA), resorting to the CAG trinucleotide repeat located in the first exon and two methylation-sensitive endonuclease sites located upstream of the gene [20]. The percentage of allele activity was determined using the peak heights and normalized to the corresponding undigested allele peak height. The *FMR1* methylation status was determined using AmplideX® mPCR *FMR1* kit (Asuragen, Inc., Austin, TX, USA), according to the manufacturer's instructions. The mPCR assay determines both the number of CGG repeats and the percentage promoter methylation of each *FMR1* allele.

3. Results

A similar *FMR1* CGG size distribution was obtained in both cohorts with normal alleles, ranging from 15 to 40 CGG in cohort 1 and from 15 to 44 CGG in cohort 2 (n = 127, 97%) and intermediate genotypes, one allele with 48 CGG in cohort 1 and three alleles with 45 CGG in cohort 2 (n = 4, 3%) (Tables 1 and 2). Homozygosity was observed in eleven samples (22%, cohort 1), of which nine shared the same CGG/AGG substructure, and in seventeen samples (21%, cohort 2), of which thirteen shared the same AGG pattern. In line with previous publications, the vast majority of the alleles (93%) showed one or two AGGs, 5% were pure (4, cohort 1 and 9, cohort 2) and the remaining 2% showed three AGG interspersions. The most common structure, $(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$, was identified in 29 (29%, cohort 1) and 40 alleles (25%, cohort 2). Similar to other worldwide populations, a highly polymorphic CGG/AGG substructure was observed: forty-one and fifty-five unique patterns were identified in cohorts 1 and 2, respectively (Tables 1 and 2) [21].

3. Results

Chapter I. Development and validation of a mathematical model predict the complexity of *FMR1* allele combinations

---

Table 1. Cohort 1 data used to calculate the *allelic scores* and identify *equivalent* (white background) and *dissimilar* (gray background) groups.

Allele 1			Allele 2		
CGG/AGG pattern	**Repeat lengh**	***Allelic score***	**CGG/AGG pattern**	**Repeat lengh**	***Allelic score***
$(CGG)_{8}AGG(CGG)_{9}$	$18^{\S}$	**41**	$(CGG)_{23}AGG(CGG)_{9}$	33	**101**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**
$(CGG)_{9}AGG(CGG)_{8}AGG(CGG)_{9}$	$20^{\S}$	**185**	$(CGG)_{10}AGG(CGG)_{8}AGG(CGG)_{9}$	29	**201**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{20}AGG(CGG)_{9}$	30	**89**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{13}AGG(CGG)_{9}$	$23^{\S}$	**61**
$(CGG)_{9}AGG(CGG)_{11}$	$21^{\S}$	**47**	$(CGG)_{12}AGG(CGG)_{16}$	29	**64**
$(CGG)_{10}AGG(CGG)_{11}$	$22^{\S}$	**51**	$(CGG)_{13}AGG(CGG)_{16}$	30	**68**
$(CGG)_{9}AGG(CGG)_{13}$	$23^{\S}$	**49**	$(CGG)_{12}AGG(CGG)_{25}$	38	**73**
$(CGG)_{9}AGG(CGG)_{15}$	$25^{\S}$	**51**	$(CGG)_{10}AGG(CGG)_{19}$	30	**59**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{10}$	31	**206**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{10}AGG(CGG)_{9}$	30	**193**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{14}$	35	**210**
$(CGG)_{9}AGG(CGG)_{10}AGG(CGG)_{9}$	30	**193**	$(CGG)_{10}AGG(CGG)_{11}AGG(CGG)_{9}$	32	**213**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{10}$	30	**190**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{8}AGG(CGG)_{9}$	39	**825**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{10}AGG(CGG)_{9}$	30	**193**	$(CGG)_{10}AGG(CGG)_{11}AGG(CGG)_{9}$	32	**213**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{12}$	32	**192**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{18}$	39	**214**
$(CGG)_{8}AGG(CGG)_{9}AGG(CGG)_{21}$	40	**185**	$(CGG)_{9}AGG(CGG)_{8}AGG(CGG)_{29}$	$48^{\#}$	**205**
$(CGG)_{15}$	$15^{\S}$	**15**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{7}AGG(CGG)_{9}$	$17^{\S}$	**37**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{10}$	31	**206**
(CGG)10AGG(CGG)9	$20^{\S}$	**49**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{10}$	32	**210**
$(CGG)_{9}AGG(CGG)_{10}$	$20^{\S}$	**46**	$(CGG)_{10}AGG(CGG)_{8}AGG(CGG)_{9}AGG(CGG)_{9}$	39	**813**
$(CGG)_{9}AGG(CGG)_{10}$	$20^{\S}$	**46**	$(CGG)_{10}AGG(CGG)_{8}AGG(CGG)_{9}$	29	**201**
$(CGG)_{9}AGG(CGG)_{10}$	$20^{\S}$	**46**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}$	$20^{\S}$	**49**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{8}AGG(CGG)_{9}$	39	**825**
$(CGG)_{9}AGG(CGG)_{10}$	$20^{\S}$	**46**	$(CGG)_{10}AGG(CGG)_{8}AGG(CGG)_{7}AGG(CGG)_{9}$	37	**805**
$(CGG)_{9}AGG(CGG)_{10}$	$20^{\S}$	**46**	$(CGG)_{9}AGG(CGG)_{12}AGG(CGG)_{9}$	32	**201**
$(CGG)_{25}$	$25^{\S}$	**25**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{10}$	31	**206**
$(CGG)_{15}AGG(CGG)_{9}$	$25^{\S}$	**69**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{10}$	31	**206**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{9}AGG(CGG)_{29}$	39	**65**
$(CGG)_{9}AGG(CGG)_{9}AGG(CGG)_{9}$	29	**189**	$(CGG)_{11}AGG(CGG)_{20}$	32	**64**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{22}$	33	**62**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{20}$	31	**60**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{22}$	33	**62**
$(CGG)_{30}$	30	**30**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**
$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**	$(CGG)_{10}AGG(CGG)_{20}$	31	**60**
$(CGG)_{30}$	30	**30**	$(CGG)_{10}AGG(CGG)_{9}AGG(CGG)_{9}$	30	**205**

Homoallelism for CGG-repeat length (black background) and homozygosity for both CGG-repeat length and AGG pattern (*allelic score* in green background); #intermediate size; §normal “low zone” alleles (see discussion).

Table 2. Cohort 2 data used to calculate the *allelic scores* and identify *equivalent* (white background) and *dissimilar* (gray background) groups.

Allele 1			Allele 2		
CGG/AGG pattern	**Repeat lengh**	***Allelic score***	**CGG/AGG pattern**	**Repeat lengh**	***Allelic score***
(CGG)$_{15}$	15§	**15**	(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**
(CGG)$_{18}$	18§	**18**	(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{25}$AGG(CGG)$_{9}$	35	**109**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{20}$	20§	**20**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{19}$	30	**59**
(CGG)$_{11}$AGG(CGG)$_{9}$	21§	**53**	(CGG)$_{12}$AGG(CGG)$_{10}$	23§	**58**
(CGG)$_{9}$AGG(CGG)$_{13}$	23§	**49**	(CGG)$_{9}$AGG(CGG)$_{19}$	29	**55**
(CGG)$_{9}$AGG(CGG)$_{13}$	23§	**49**	(CGG)$_{12}$AGG(CGG)$_{32}$	45$^{\#}$	**80**
(CGG)$_{13}$AGG(CGG)$_{9}$	23§	**61**	(CGG)$_{24}$	24§	**24**
(CGG)$_{10}$AGG(CGG)$_{13}$	24§	**53**	(CGG)$_{13}$AGG(CGG)$_{16}$	30	**68**
(CGG)$_{16}$AGG(CGG)$_{9}$	26§	**73**	(CGG)$_{29}$	29	**29**
(CGG)$_{9}$AGG(CGG)$_{18}$	28	**54**	(CGG)$_{9}$AGG(CGG)$_{28}$	38	**64**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{8}$AGG(CGG)$_{9}$	29	**201**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{8}$	29	**204**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{12}$AGG(CGG)$_{9}$	32	**201**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	30	**190**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{16}$	37	**212**
(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{10}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	32	**213**
(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{10}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	32	**213**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	30	**190**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{8}$	39	**828**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{10}$AGG(CGG)$_{10}$	32	**210**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{10}$AGG(CGG)$_{14}$AGG(CGG)$_{9}$	35	**225**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	30	**190**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{20}$	41	**216**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{10}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	32	**213**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	30	**190**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{19}$	40	**215**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	30	**190**	(CGG)$_{9}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	31	**197**
(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{10}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	32	**213**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	31	**209**
(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{10}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	32	**213**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{10}$AGG(CGG)$_{12}$AGG(CGG)$_{9}$	33	**217**
(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{11}$	31	**191**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{17}$	38	**213**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{19}$	39	**199**	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{15}$	45$^{\#}$	**771**
(CGG)$_{10}$AGG(CGG)$_{5}$	16§	**45**	(CGG)$_{10}$AGG(CGG$_{)9}$AGG(CGG)$_{10}$	31	**206**
(CGG)$_{10}$AGG(CGG$_{)9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{10}$	20§	**46**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{8}$	29	**204**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{10}$AGG(CGG)$_{10}$	32	**210**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{10}$	20§	**46**	(CGG)$_{10}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	32	**213**
(CGG)$_{9}$AGG(CGG)$_{10}$	20§	**46**	(CGG)$_{10}$AGG(CGG)$_{8}$AGG(CGG)$_{9}$	29	**201**
(CGG)$_{20}$	20§	**20**	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{10}$	20§	**46**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{10}$	20§	**46**	(CGG)$_{10}$AGG(CGG)$_{8}$AGG(CGG)$_{9}$	29	**201**
(CGG)$_{10}$AGG(CGG)$_{9}$	20§	**49**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{10}$AGG(CGG)$_{11}$	22§	**51**	(CGG)$_{12}$AGG(CGG)$_{7}$AGG(CGG)$_{9}$	30	**229**
(CGG)$_{13}$AGG(CGG)$_{9}$	23§	**61**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{12}$AGG(CGG)$_{10}$	23§	**58**	(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**
(CGG)$_{13}$AGG(CGG)$_{9}$	23§	**61**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{13}$AGG(CGG)$_{9}$	23§	**61**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{13}$AGG(CGG)$_{9}$	23§	**61**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{27}$	27	**27**	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**
(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{29}$AGG(CGG)$_{9}$	39	**125**

3. Results

Chapter I. Development and validation of a mathematical model predict the complexity of *FMR1* allele combinations

$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{38}$	38	**38**
$(CGG)_9AGG(CGG)_{19}$	29	**55**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**
$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{10}AGG(CGG)_{20}$	31	**60**
$(CGG)_9AGG(CGG)_{19}$	29	**55**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**
$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{22}AGG(CGG)_9$	32	**97**
$(CGG)_{10}AGG(CGG)_{19}$	30	**59**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**
$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	**193**	$(CGG)_{10}AGG(CGG)_{19}$	30	**59**
$(CGG)_{30}$	30	**30**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**
$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{33}AGG(CGG)_9$	43	**141**
$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{10}AGG(CGG)_{19}$	30	**59**
$(CGG)_{10}AGG(CGG)_{20}$	31	**60**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{23}$	44	**219**
$(CGG)_9AGG(CGG)_{21}$	31	**57**	$(CGG)_{10}AGG(CGG)_8AGG(CGG)_{25}$	$45^{\#}$	**217**
$(CGG)_{10}AGG(CGG)_{20}$	31	**60**	$(CGG)_{10}AGG(CGG)9AGG(CGG)_{10}$	31	**206**

Homoallelism for CGG-repeat length (black background) and homozygosity for both CGG-repeat length and AGG pattern (*allelic score* in green background); #intermediate size; §normal "low zone" alleles (see discussion).

3.1. Development of the mathematical model

A mathematical model was developed to integrate the AGG interspersion number and pattern and the total repeat length, reflecting the CGG/AGG substructure. The result score, named *allelic score*, was calculated separately for each allele as follows:

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

where,

R_i: number of CGG repeats before the first AGG interspersion of order i;

i: CGG repeat order number;

n: total number of AGG interspersions;

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

Base-4 numeral system was used to ensure that the *allelic score* is unique to each of the AGG interspersion patterns and sufficiently spaced.

For the purpose of addressing allelic complexity, two different aspects of the allelic structure are considered: number of AGG interspersions and number of CGG repeats between interspersions. Higher relevance is given to the number of interspersions as, for alleles with identical number of CGG repeats, higher number of AGG interspersions is usually linked with allelic stability [3,4]. As example, an allele with two AGGs shows an *allelic score* of 193 whereas an allele with a similar length but only one AGG has an *allelic score* of 59.

$$\textit{Allelic score}\ [(CGG)_9AGG(CGG)_{10}AGG(CGG)_9] =$$

$$[(9 \times 4^{1-1}) + (10 \times 4^{2-1})] + (9 \times 4^2) = 193$$

$$\textit{Allelic score}\ [(CGG)_{10}AGG(CGG)_{19}] =$$

$$(19 \times 4^{1-1}) + (10 \times 4^1) = 59$$

The *allelic score* calculation emphasizes the 5' end of the allele by assigning higher weights to CGG repeats closer to the 5' end and decreasing weights to repeats located further downstream towards the 3' end from the initial AGG interruptions.

This mathematical model is protected with a national patent (reference – 115244) and international patent application submitted on December 6, 2019 (reference - PCT/IB2019/060520).

3.2. Application and validation of the mathematical model

Allelic scores ranged from 15 to 825 (cohort 1) and 15 to 828 (cohort 2), with most samples scoring below 220 (95.4%) and six with a score in the order of 800, due to the presence of three AGG interspersions (Tables 1 and 2). Scores under 220 either represent zero, one or two AGG interspersions; above two AGG interspersions, the *allelic score* grows exponentially. An exploratory cluster analysis identified four major clusters, with observations within each quadrant separated in both axes by an *allelic score* of 150 (Supplementary Figures S1 and S2). Similar behaviors were observed among the two quadrants where *allelic scores* were both lower than 150 or both higher than 150, and the other two where alleles show low and high *allelic score*, allowing the definition of two groups. The *equivalent* group contains samples where both alleles show a similar complexity, and the *dissimilar* group with samples where alleles show a different complexity. These groups include samples with three AGGs as the behavior of their alleles fits that of other samples in the same quadrant (Supplementary Figures S1 and S2). In both groups, an exponential model was used to describe the correlation between the *allelic score* of each allele. Significant correlations were found: cohort 1 - *equivalent* group: r = 0.8092; df = 24; $p < 0.0001$ and *dissimilar* group: r = -0.7067; df = 22; $p < 0.0001$ (Supplementary Figure S3). To validate the mathematical models and their reproducibility, a covariance (ANCOVA) analysis was used to compare the models calculated for cohort 1 and the same models computed using cohort 2 data (*equivalent* group: r = 0.8603; df = 43; $p < 0.0001$ and *dissimilar* group: r = -0.8716; df = 33; $p < 0.0001$) (Supplementary Figure S4). There was no statistically significant difference

between cohort 1 (development cohort) and cohort 2 (validation cohort) with respect to the *equivalent* and *dissimilar* group's models, as demonstrated by the coincident regression lines (Supplementary Figure S5). A more robust model including all observations (both cohorts) was derived: *equivalent* group – $F_{(2, 68)} = 1.8048$; $p = 0.1723$: ln(score 2) = 3.6452 + 0.0088 x score 1 and *dissimilar* group – $F_{(2, 55)} = 0.9574$; $p = 0.3902$: ln(score 2) = 5.6944 – 0.0065 x score 1.

Seven samples from each group (cohort 2) were tested for XCI pattern (Supplementary Table S1). Interestingly, in a sample belonging to the *dissimilar* group, *FMR1* mPCR showed extreme skewing (85%) towards the smallest "low zone" allele.

4. Discussion

Our study focused on developing a tool to score and evaluate the complexity of the *FMR1* gene repetitive tract structure. To this end, a mathematical model was designed that computes the *FMR1* gene CGG repeat length, as well as the AGG interspersion number and pattern. The output, a number designated *allelic score*, deciphers a functional model to predict the complexity of allele combinations. Two cohorts of young, healthy, and potentially fertile females were used independently for development and validation studies. The fact that two statistically significant groups, *equivalent* and *dissimilar*, were identified in both cohorts, justified the pooling of data. Furthermore, the identification of two groups shows the model's ability to compare the complexity of the two alleles. Interestingly, the *dissimilar* group is enriched with "low zone" heterozygous samples (herein defined as CGG ≤ 26). It has been proposed that these "low zone" alleles may exert negative effects, although controversial [11,12,15,16]. Another study claims that normal *FMR1* repeat length outside 26 > CGG > 34 concur with a higher X-chromosome inactivation skew, a putative mechanism underlying the ovarian reserve impairment (as assessed by AMH), particularly in infertile older females [22]. Moreover, the AGG "protective" effect towards a decreased risk of ovarian malfunction was observed in females carrying premutated alleles with 2 or more interspersions [23]. According to our model, these alleles would show a high *allelic score*, which seems to suggest a correlation between the allelic complexity and a protective effect. Replication of these results is still required using larger control and patient cohorts. Nonetheless, with this mathematical model developed to calculate the *FMR1 allelic score*, further research can now be undertaken with a different perspective in terms of *FMR1* characterization.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committe of Centro Hospitalar Universitário do Porto. Written informed consent from the participants was obtained in accordance with the national legislation (lei 12/2005) and the institutional requirements. The participants provided their written informed consent to participate in this study.

Author Contributions

PJ conceived and designed the study together with BR. AN developed the mathematical model and performed the statistical analysis with BR who also carried out laboratory work, analyzed the data, and drafted the manuscript. FS performed methylation/inactivation studies. EV-F, NM, IM, and RS provided critical feedback, helped conduct the research, and contributed toward the manuscript. All authors discussed the final results and critically reviewed the manuscript.

Funding

This work was supported by national funds: FCT/MCTES (Fundação para a Ciência e a Tecnologia) – Project Reference SFRH/BD/136398/2018 to Bárbara Rodrigues, UMIB (Unidade Multidisciplinar de Investigação Biomédica) – Reference UIDP/00215/2020 and UIDB/00215/2020, DEFI (Departamento de Ensino, Formação e Investigação) – Reference 2015-DEFI/145/12, and CESAM (Centro de Estudos do Ambiente e do Mar) – Reference UIDP/50017/2020 and UIDB/50017/2020.

Acknowledgments

The authors gratefully acknowledge the Centre for Medically Assisted Procreation/Public Gamete Bank, Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), Centro Hospitalar Universitário do Porto (CHUP). Without the invaluable help of all collaborators (potential donors, clinicians, nurses and embryologists), our work would not have been possible. Special thanks to Isabel Sousa Pereira for recruiting cohort 1 participants. The authors are also grateful for important feedback from international reviewers.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

References

1. 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.

2. 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;23:417–25.

3. 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;62:1–7.

4. Manor E, Gonen R, Sarussi B, Keidar-Friedman D, Kumar J, Tang HT, et al. The role of AGG interruptions in the FMR1 gene stability: A survey in ethnic groups with low and high rate of consanguinity. Mol Genet Genomic Med. 2019;7:1–14.

5. Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, et al. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med. 2012;14:729–36.

6. Mila M, Alvarez-Mora MI, Madrigal I, Rodriguez-Revenga L. Fragile X syndrome: An overview and update of the FMR1 gene. Clin Genet. 2018;93:197–205.

7. McGinty RJ, Mirkin SM. Cis- and Trans-Modifiers of Repeat Expansions: Blending Model Systems with Human Genetics. Trends Genet. 2018;34:448–65.

8. Latham GJ, Coppinger J, Hadd AG, Nolin SL. The role of AGG interruptions in fragile X repeat expansions: A twenty-year perspective. Front Genet. 2014;5:1–6.

9. Domniz N, Ries-Levavi L, Cohen Y, Marom-Haham L, Berkenstadt M, Pras E, et al. Absence of AGG Interruptions Is a Risk Factor for Full Mutation Expansion Among Israeli FMR1 Premutation Carriers. Front Genet. 2018;9:1–8.

10. Ardui S, Race V, Ravel T de, Esch H Van, Devriendt K, Matthijs G, et al. Detecting AGG interruptions in females with a FMR1 premutation by long-read single-molecule sequencing: A 1 year clinical experience. Front Genet. 2018;9:1–6.

11. Mailick MR, Hong J, Rathouz P, Baker MW, Greenberg JS, Smith L, et al. Low-normal FMR1 CGG repeat length : phenotypic associations. Front Genet. 2014;5:1–9.

12. 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;166:502–7.

13. Rehnitz J, Alcoba DD, Brum IS, Dietrich JE, Youness B, Hinderhofer K, et al. FMR1 expression in human granulosa cells increases with exon 1 CGG repeat length depending on ovarian reserve. Reprod Biol Endocrinol. 2018;16:1–9.

14. Wang Q, Kushnir VA, Darmon S, Barad DH, Wu Y, Zhang L, et al. Reduced RNA expression of the FMR1 gene in women with low (CGGn<26) repeats. Fertil Steril. 2017;108:e143.

15. 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;31:2396–403.

16. 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;02:10–3.

17. Galvão A, Vale-Ferndandes E, Pereira IS, Fraga S, Lourenço C, Morgado A, et al. Applicants for oocyte donors from the Portuguese public gamete bank: who are they? C SUMMIT 17 - Inovações e Controv na Saúde da Mulh e da Criança. 2017.

18. Monaghan KG, Lyon E, Spector EB. ACMG standards and guidelines for fragile X testing: A revision to the disease-specific supplements to the standards and guidelines for Clinical Genetics Laboratories of the American College of Medical Genetics and Genomics. Genet Med. 2013;15:575–86.

19. Zar JH. Bioestatistical Analysis fifth edition. 2010.

20. Allen RC, Y.Zoghbi H, It Annemarie B. Moseley HMR, Belmont JW. Methylation of Hpall and Hhal Sites Near the Polymorphic CAG Repeat in the Human Androgen-

Receptor Gene Correlates with X Chromosome Inactivation. Am J Hum Genet. 1992;51:1229-12:1229–39.

21. Yrigollen CM, Sweha S, Durbin-Johnson B, Zhou L, Berry-Kravis E, Fernandez-Carvajal I, et al. Distribution of AGG interruption patterns within nine world populations. Intractable Rare Dis Res. 2014;3:153–61.

22. Barad DH, Darmon S, Weghofer A, Latham GJ, Wang Q, Kushnir VA, et al. Association of skewed X-chromosome inactivation with FMR1 CGG repeat length and anti-Mullerian hormone levels: A cohort study. Reprod Biol Endocrinol. 2017;15:1–8.

23. Lekovich J, Man L, Xu K, Canon C, Lilienthal D, Stewart JD, et al. CGG repeat length and AGG interruptions as indicators of fragile X–associated diminished ovarian reserve. Genet Med. 2018;20:957–64.

Supporting Information

1. Supplementary Figures

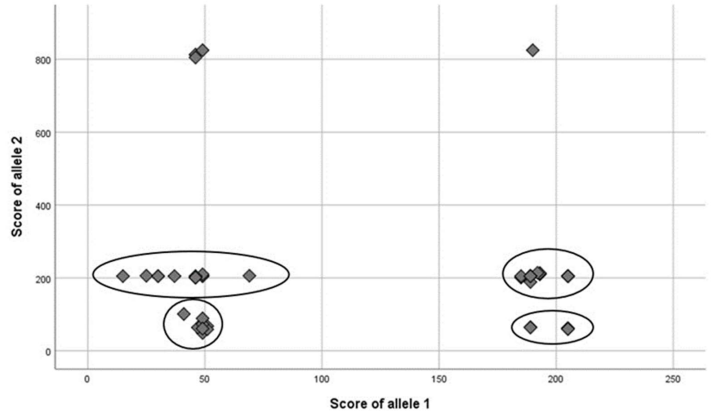


Figure S1. Cohort 1 samples clustering using the *allelic score* of each allele.

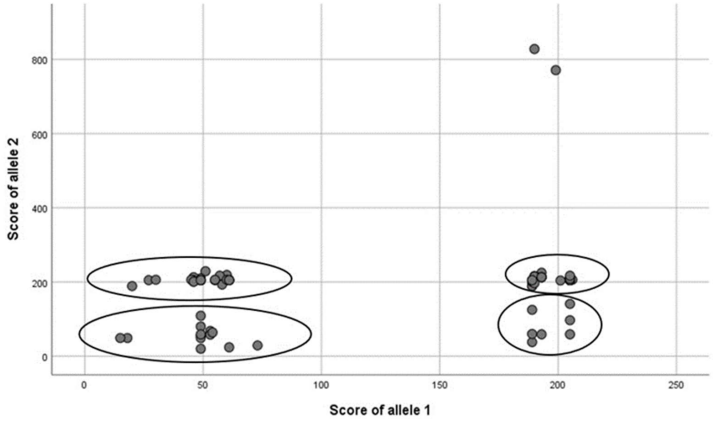


Figure S2. Cohort 2 samples clustering using the *allelic score* of each allele.

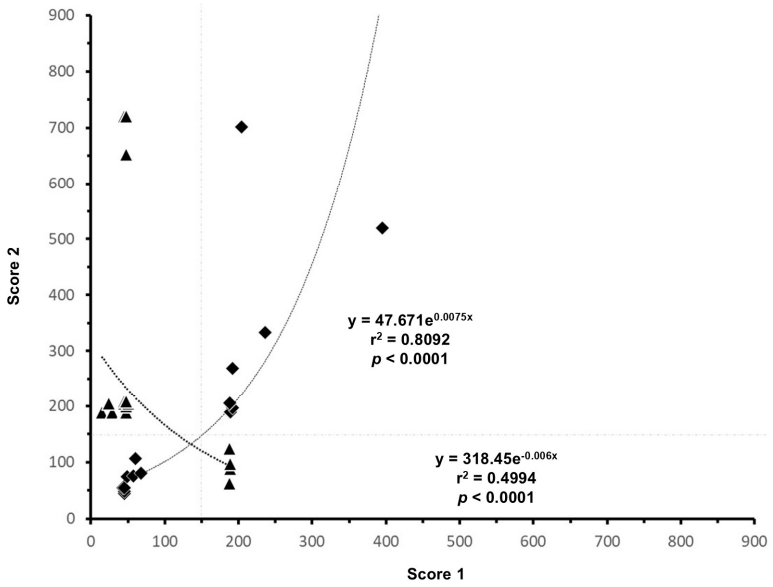


Figure S3. Correlation between *allelic score* of each allele in cohort 1 *equivalent* (diamond) and *dissimilar* (triangles) groups.

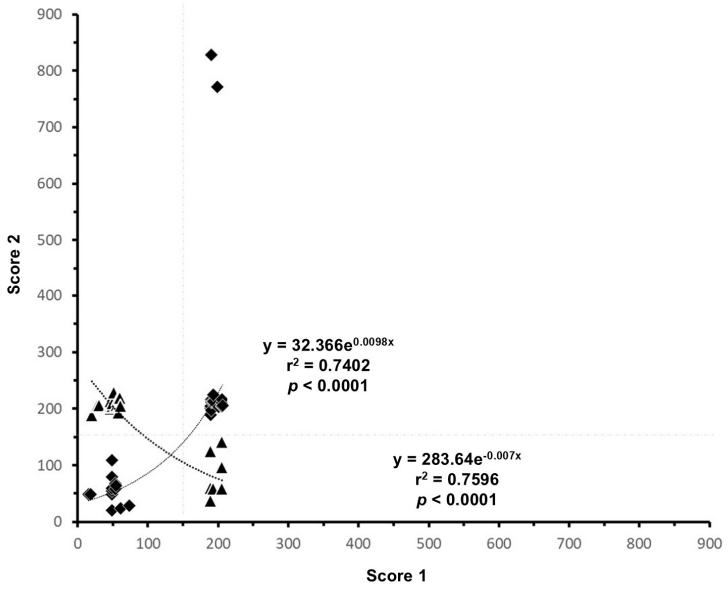


Figure S4. Correlation between *allelic score* of each allele in cohort 2 *equivalent* (diamond) and *dissimilar* (triangles) groups.

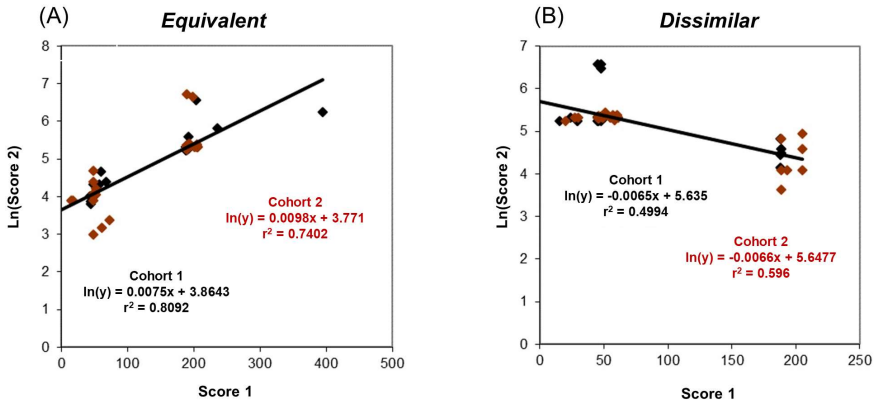


Figure S5. Comparison between cohort 1 (black) and cohort 2 (red) models in *equivalent* (A) and *dissimilar* (B) groups.

2. Supplementary Tables

Table S1. X-chromosome inactivation patterns observed using HUMARA.

Samples	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Allelic score allele 1-allele 2	53-58	189-205	193-213	205-205	193-225	205-205	205-205	46-204	49-205	20-189	**61-205**	205-141	59-206	57-217
Methylation% [allele 1:allele 2]	[21:79]	[44:56]	[53:47]	[48:52]	[19:81]	[71:29]	[67:33]	[48:52]	[66:34]	[75:25]	**[92:8]**	[62:38]	[41:59]	[56:44]

Results are shown according to the order of samples in table 2. White and gray background represent the samples from the *equivalent* and *dissimilar* groups, respectively. Sample 11 shows a non-random X-inactivation pattern highlighted in bold [genotype: $(CGG)_{13}AGG(CGG)_9$ and $(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$].

3. Results

Chapter I. Development and validation of a mathematical model predict the complexity of *FMR1* allele combinations

---

Chapter II. *FMR1* allelic complexity in premutation carriers provides no evidence for a correlation with age at amenorrhea[2]

Abstract

Background: Premutations in the Fragile X Messenger Ribonucleoprotein 1 (*FMR1*) gene, defined as between 55 to 200 CGGs, have been implicated in fragile X-associated primary ovarian insufficiency (FXPOI). Only 20% of female premutation carriers develop early ovulatory dysfunction, the reason for this incomplete penetrance is unknown. This study validated the mathematical model in premutation alleles, after assigning each allele a score representing allelic complexity. Subsequently, *allelic scores* were used to investigate the impact of allele complexity on age at amenorrhea for 58 premutation cases (116 alleles) previously published.

Methods: The *allelic score* was determined using a formula previously described by our group. The impact of each *allelic score* on age at amenorrhea was analyzed using Pearson's test and a contour plot generated to visualize the effect.

Results: Correlation of *allelic score* revealed two distinct complexity behaviors in premutation alleles. No significant correlation was observed between the *allelic score* of premutation alleles and age at amenorrhea. The same lack of significant correlation was observed regarding normal-sized alleles, despite a nearly significant trend.

Conclusions: Our results suggest that the use of *allelic scores* combination have the potential to explain female infertility, namely the development of FXPOI, or ovarian dysfunction, despite the lack of correlation with age at amenorrhea. Such a finding is of great clinical significance for early identification of females at risk of ovulatory dysfunction, enhancement of fertility preservation techniques, and increasing the probability for a successful pregnancy in females with premutations. Additional investigation is necessary to validate this hypothesis.

Keywords: *FMR1* gene premutation, age at amenorrhea, *FMR1* allelic complexity, Fragile X-associated Primary Ovarian Insufficiency, CGG repeats, AGG interspersion pattern

[2]Rodrigues *et al*., *FMR1* allelic complexity in premutation carriers provides no evidence for a correlation with age at amenorrhea. Reprod Biol Endocrinol. 2024 Jun 21;22(1):71.

1. Introduction

The Fragile X Messenger Ribonucleoprotein 1 (*FMR1*) gene, located on the X chromosome (Xq27.3), contains a polymorphic CGG repeat on its 5' untranslated region (UTR) implicated in three disorders depending on the repeat number: Fragile X Syndrome (FXS; OMIM #300624) when CGGs > 200, and Fragile X-associated Tremor/Ataxia Syndrome (FXTAS; OMIM #300623) and Fragile X-associated Primary Ovarian Insufficiency (FXPOI; OMIM #311360) in the premutation (PM) range of 55 ≤ CGG ≤ 200 [1–3]. The mechanism of FXPOI development is not fully understood but it is believed to be due to the toxic effect of elevated *FMR1* mRNA levels [2,4]. PM carriers with FXPOI show hypergonadotropic hypogonadism and absent or irregular menstrual cycles before 40 years of age [5]. The CGG repeat length correlates unevenly with FXPOI, as females carrying 70 to 100 CGGs have an increased risk of FXPOI when compared with those with more than 100 repeats [6,7]. Furthermore, FXPOI development is not fully penetrant. In *FMR1* normal-sized alleles (5 to 44 CGGs), the repetitive region is usually interrupted by one or more AGGs, typically occurring at every 9th or 10th CGG [8]. PM alleles are predominantly composed of pure CGGs; loss of AGG interruption(s) has been linked to the instability of the repetitive region and the increased risk of expansion [3,9–13]. A formula integrating the total repeat length, and the number and pattern of the AGGs was developed to calculate *FMR1 allelic score* [14]. *Allelic score* is a metric that reflects the complexity of the *FMR1* gene CGG repetitive region. Herein, we evaluate the association between the combination of normal and PM *allelic scores* and ovulatory dysfunction underlying FXPOI. Our formula was applied to calculate the *allelic scores* in *FMR1* PM carriers, validating its use in samples with distinct genotypic characteristics. It was hypothesized that the combination of the AGG number and pattern from both normal and PM alleles would associate indirectly with age at amenorrhea and hormone levels with a potential impact on FXPOI development.

2. Material and Methods

2.1. *FMR1 allelic scores* determination

Molecular data of both alleles regarding samples from PM carriers, previously published, were requested to the respective authors: Villate et al. (2020) [15] (Spain), Allen et al. (2018) [16] (United States of America) and Yrigollen et al. (2014) [17] (United States of America). Of all data provided by the authors, 577 results were retrieved: Villate et al. (2020) [15] (*n* = 20, designated by set 1), Allen et al. (2018) [16] (*n* = 59, designated by set 2) and Yrigollen et al. (2014) [17] (*n* = 498, designated by set 3). The *allelic score*,

which reflects the *FMR1* CGG/AGG substructure, was calculated separately for each allele (normal and PM), using the formula described in Rodrigues et al. (2020) [14]. The age at amenorrhea - defined by at least 4 months of secondary amenorrhea and menopausal levels of follicle-stimulating hormone (FSH) [16] - was reported in 58 observations from set 2 (mean age 38.7 ± 8.5 years, range 18 - 56) thus resulting in a slightly smaller dataset (58 observations instead of 59).

2.2. Reference set

The reference set, composed of one hundred and thirty-one female samples with normal (*n* = 127) and intermediate genotypes (*n* = 4), was previously described and characterized in Rodrigues et al. (2020) [14] (the summary of the results can be found in Supplementary Table 1).

2.3. Statistical analysis

A linearized form of a logarithmic model [i.e., regression of ln(score 1) against score 2] was used to obtain a functional model to relate the complexity of both alleles in each set. Covariance analysis (ANCOVA) compared the reference set with PM sample set regression models, following the methodology outlined by Zar [18]. SigmaPlot version 14.0 (Systat Software®Inc., Chicago, IL, USA) was used for One-Way ANOVA on ranks (Kruskal-Wallis test) to compare separately *allelic score* and the size of each allele (normal and PM). Dunn's method was used for multiple comparisons after conducting a Kruskal-Wallis test, comparing sets based on median allele size and *allelic score*. The relationship between the age at amenorrhea and *allelic score* was assessed by Pearson correlation coefficient. R sofware version 4.3.0 by R Core Team [19] with the ggplot2 package [20] was used to generate contour plots to display the relationship between independent variables normal and PM *allelic scores*, and the dependent variable, age at amenorrhea. All statistical tests were carried out for a significance level of 0.05.

3. Results

3.1. *FMR1* CGG repeat characterization

FMR1 molecular data of 1154 alleles are summarized in Supplementary Table 2. In a set 3 sample both alleles are in the PM range, a rare event previously reported in seven cases [21]. The most frequent repeat length among normal-sized alleles is 30 CGGs, despite the significant differences among allele sizes (Kruskal-Wallis test: H = 12.3; df = 2; p = 0.002) (Supplementary Fig. 1a). The majority of normal-sized alleles contained

one or two AGG interruptions (93.8%, *n* = 540) while pure alleles occurred in 4.7% of samples (*n* = 4, set 2, *n* = 23, set 3), and the remaining 1.5% of samples showed three AGGs (*n* = 9, set 3). In total, ninety-seven different AGG patterns (*n* = 8/ 20, set 1, *n* = 23/ 59, set 2 and *n* = 75/ 498, set 3) were identified. The most common AGG interspersion pattern in sets 1 and 3 is $(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$. On the contrary, a very rare pattern was identified as commonest among set 2 samples $(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$. Around half of the PM alleles had no AGGs (50.3%, *n* = 8, set 1, *n* = 26, set 2, *n* = 257, set 3), and approximately 49.7% showed one or two AGG interruptions (*n* = 287). Two hundred and eleven different patterns were identified (*n* = 13/ 20, set 1, *n* = 45/ 59, set 2 and *n* = 177/ 498, set 3) in PM alleles revealing very exclusive CGG/AGG structures.

3.2. Mathematical model validation

Descriptive statistics and frequency analyses of PM *allelic scores* are shown in Table 1. The median PM *allelic scores* did not show statistically significant differences between sets (Kruskal - Wallis test: H = 1.45; df = 2; $p = 0.484$) (Supplementary Fig. 1d); despite the sets having significantly different normal median *allelic scores* (Kruskal-Wallis test: H = 33.20; df = 2; $p < 0.001$) (Supplementary Fig. 1b). To compare these PM samples with previously published data using the same mathematical model, a reference set was built from that publication [14]. All sets distributed *allelic scores* into four quadrants, separated by a value of 150 as previously observed in the reference set, revealing similar compositions (Fig. 1): samples with alleles showing a similar complexity (*equivalent* group, quadrants 1 and 3) (Fig. 1a) and samples where alleles have a different complexity (*dissimilar* group, quadrants 2 and 4) (Fig. 1b). Thus, the correlation between the *allelic score* of each allele (Fig. 1a and b) was described following a logarithmic model. Significant correlations were found in both groups from all sets: reference set – *equivalent* group: r = -0.539; df = 72; $p < 0.0001$ and *dissimilar* group: r = -0.416; df = 55; $p < 0.0001$ (Fig. 1a and b, represented by circles); set 1 – *equivalent* group: r = 0.539; df = 9; $p < 0.0001$ and *dissimilar* group: r = -0.416; df = 7; $p < 0.0001$ (Fig. 1a and b, represented by squares); set 2 – *equivalent* group r = 0.539; df = 27; $p = 0.003$ and d*issimilar* group: r = -0.173; df = 28; $p = 0.02$ (Fig. 1a and b, represented by lozenges), and set 3 – *equivalent* group: r = 0.556; df = 192; $p < 0.0001$ and *dissimilar* group: r = -0.416; df = 302; $p < 0.0001$ (Fig. 1a and b, represented by triangles). An exponential growth of the *allelic score* was observed, particularly in alleles having more than two AGGs (Supplementary Table 3); due to the relevance attributed to the AGG number by

the formula. For instance, samples with three AGGs show scores above 700 (n = 6, reference set, and n = 8, set 3; represented by a gray circle in Fig. 1a and b). To validate the mathematical model in expanded alleles a covariance analysis between the reference and PM sample sets logarithmic models was performed separately for each group (Supplementary Fig. 2). Supplementary table 4 shows the individual models resulting from each set. Coincident regression lines demonstrate the absence of statistically significant differences in each *equivalent* and *dissimilar* groups from PM samples sets when compared with those of the reference set. This result supports a more robust model including observations from the four sets: *equivalent* group – $F_{(6, 300)} = 1.8278$; $p = 0.0934$: Score 2 = -238.3 + 87.4 × ln(score 1) and *dissimilar* group – $F_{(6, 392)} = 1.0679$; $p = 0.3812$: Score 2 = 573.9 – 88.4 × ln(score 1).

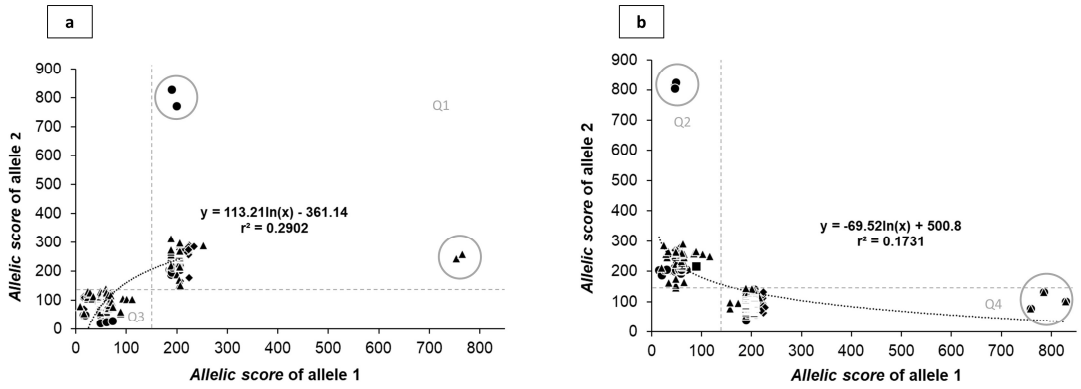


Fig. 1 Correlation between the *FMR1* allelic complexity (*allelic score*) of each allele in all sets, according to groups: *equivalent* (a) and *dissimilar* (b). The data are categorized into four quadrants (Q1, Q2, Q3, Q4) based on the distribution of *allelic scores*. The reference set is represented by circles, while set 1, set 2, and set 3 are represented by squares, lozenges, and triangles, respectively. Alleles with an *allelic score* greater than 700 are indicated by gray circles.

Table 1. Summary of the *FMR1* allelic complexity (*allelic score*) results.

		A1 - Shorter CGG repeat length allele			A2 - Longer CGG repeat length allele		
		Set 1	**Set 2**	**Set 3**	**Set 1**	**Set 2**	**Set 3**
Number of alleles		40	118	996	40	118	996
Allelic score	Mean (± S.D.)	159.0 ± 66.5	165.8 ± 79.8	159.7 ± 105.5	152.7 ± 81.8	150.2 ± 81.7	131.0 ± 61.6
	Median	205	207	193	217	214	109
	Range	49 - 206	16 - 234	9 - 829	56 - 242	63 - 288	55 - 313
	Most frequent (%, *n*)	205 (30%, *n* = 6) 206 (25%, *n* = 5)	223 (27.1%, *n* = 16) 207 (18.6%, *n* = 11)	205 (34.9%, *n* = 174) 189 (12.7%, *n* = 63)	231 (15%, *n* = 3) 217 (15%, *n* = 3)	133 (5.1%, *n* = 3) 83 (5.1%, *n* = 3)	103 (2.4%, *n* = 12) 100 (2.4%, *n* = 12)
Kruskal-Wallis Test			H = 33.20; df = 2; $p < 0.001$			H = 1.45; df = 2; $p = 0.484$	

S.D. = Standard deviation; % = Frequency; *n* = Number of alleles; *p* values represent significant levels between sets 1, 2 and 3 *allelic scores*; Multiple Comparison (Dunn's Method) results in Supplementary Fig. 1b and d;

Data published in Villate et al. (2020) [15] (set 1), Allen et al. (2018) [16] (set 2) and Yrigollen et al. (2014) [17] (set 3).

3.3. *FMR1 allelic scores* and age at amenorrhea association

To understand the impact of *FMR1 allelic score* on the age at amenorrhea, normal-sized (allele 1) and PM alleles (allele 2) from set 2 samples were analyzed separately (n = 58). No significant correlation was observed between A1 *allelic score* and age at amenorrhea ($p > 0.05$) (Supplementary Fig. 3a, c, e and g). The same was true when PM *allelic score* was used *(*$p > 0.05$) (Supplementary Fig. 3b, d, f, and h). A nearly significant trend (p = 0.058) is apparent between the A1 *allelic score* and age at amenorrhea in samples showing an *allelic score* between 206 - 234 (Supplementary Fig. 3a) and 16 - 68 (Supplementary Fig. 3e) (quadrants 1 and 3, respectively). Two distinct behaviors were observed: age at amenorrhea rise with increasing *allelic score* (above 200, quadrant 1, mean age at amenorrhea 40 ± 8.5 years, Supplementary Fig. 3a), and age at amenorrhea decrease with increasing *allelic score* (below 70, quadrant 3, mean age at amenorrhea 38 ± 8 years, Supplementary Fig. 3e). The majority of these samples have alleles with less than 26 CGGs (78.6%, n = 11), with one or no AGG interspersions (71.4%, n = 10, 28.6%, n = 4, respectively), whereas those with higher *allelic scores* have alleles ranging from 29 to 32 CGGs, with two AGG interspersions.

3.4. Age of amenorrhea assessment by *allelic scores* combination

PM alleles within the range 70-100 CGGs are known to have increased risk of developing FXPOI [6,7,22], however not all carriers develop disease and there is lack of knowledge on the underlying mechanisms. This led us to speculate if FXPOI development could be associated with a combined effect of *FMR1* allelic complexity. To analyze the joint effect of A1 and PM *allelic scores* in the age at amenorrhea, a contour plot was generated. Overall, different trends were observed: menopause age approaches normal (mean 51 years, range 40 to 60 years) when the *allelic score* of both alleles increases or decreases, showing that balanced *allelic scores* have minimal impact on early amenorrhea. Deeper analysis of samples with mean age at amenorrhea below 40 years and PM *allelic score* between 70 - 123 show that age decreases with increasing A1 *allelic score* (Fig. 2, A1 *allelic score* between 50 - 55).

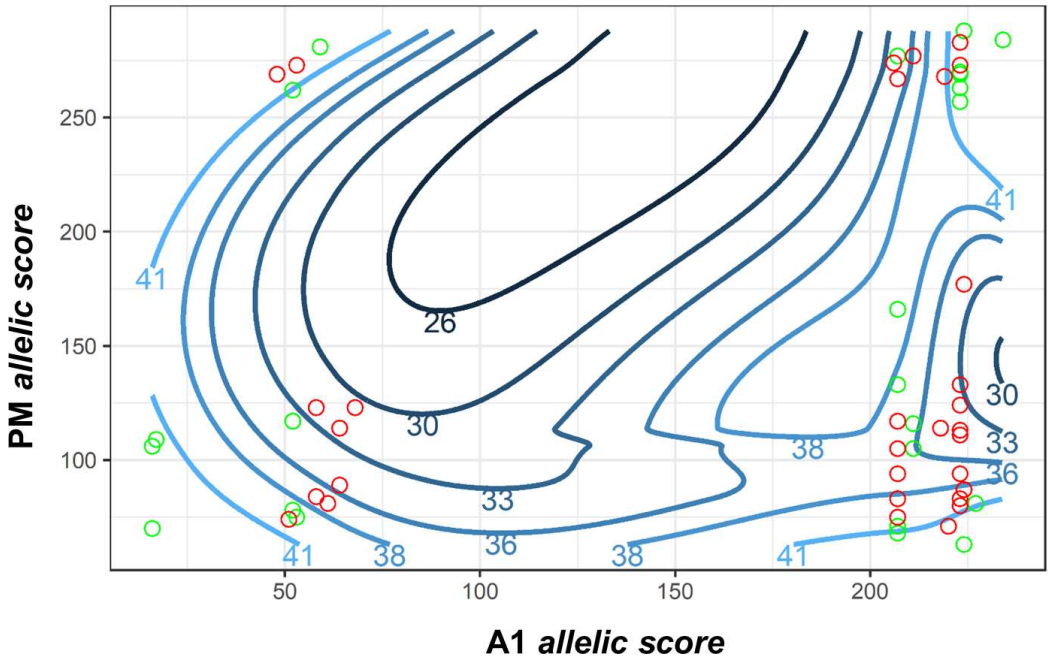


Fig. 2 Contour plot illustrating the interaction between *allelic scores* and age at amenorrhea. The contour plot visually represents the relationship between the *allelic scores* of both alleles (independent variables, x and y) and the age of amenorrhea (dependent variable, z). Red circles denote samples where the combination of allelic complexity is associated with amenorrhea occurring before the age 40 years, while green circles represent samples with amenorrhea onset after the age of 40. The varying colors of the contour lines indicate the age at amenorrhea as a function of the *allelic scores* of both alleles. A1 – Normal-sized alleles; PM – Premutation allele.

4. Discussion

In our study, aiming to validate the previously published mathematical model ascertained in normal and intermediate alleles, we compared three distinct datasets with PM carriers and subsequently explored the relationship between allelic complexity of the *FMR1* gene and age at amenorrhea – a clinical manifestation associated with development of FXPOI.

A comparative analysis of the CGG/AGG substructure across the sets revealed that the normal-sized allele with 30 repeats was the most frequent with the more prevalent AGG interspersion pattern being $(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ in sets 1 and 3, consistent with findings in other populations [14,23,24]. A notably rare pattern, $(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$, emerged as the most prevalent in set 2, which can be attributed to intrinsic characteristics of this subpopulation such as complexity and

heterogeneity of genetic traits. Expectedly, approximately half of the PM alleles showed one or two AGG interruptions with an overall average repeat length of 90.3 (S.D. = 20.8), since PM alleles with longer CGG lengths tend to demonstrate a lower incidence of AGG interruptions [11]. Notably, many different AGG interspersion patterns were found in PM alleles, most probably due to the inherent instability of these alleles [8].

No statistically significant difference was observed among the *allelic score* of PM alleles, but we found a difference for the *allelic score* of normal-sized alleles. A similar result was observed when comparing the total CGG repeat length of normal alleles, revealing great variability in the complexity of the CGG repetitive region of normal-sized alleles. The combination of *allelic scores* revealed the emergence of two groups with distinct characteristics: *equivalent* and *dissimilar*, both exhibiting significant correlations. Similar outcomes had been reported by Rodrigues et al. (2020) [14]. The validation of our mathematical model in females with *FMR1* expansions showed that this model can be applied in populations that exhibit varied genotypic characteristics, namely expanded alleles such as PM.

Several studies have sought to comprehend the impact of the CGG repetitive region on the development of FXPOI. However, the majority focus on examining the influence of the CGGs and AGGs independently, as exemplified by Friedman-Gohas et al. (2020) study [25]. Here, we employed our formula, which integrate the total CGG repeat length, the number of AGGs, and the AGG interspersion pattern. No statistically significant correlations were found between A1 *allelic scores* and age at amenorrhea, nor PM *allelic scores* and age at amenorrhea. The lack of statistical significance might be due to the reduced number of observations in each grouping [26]. Nevertheless, a significant trend was observed with normal *allelic scores* between 206 and 234 and 16 and 68 and age at amenorrhea. The influence of *FMR1* gene alleles within normal size in ovarian reserve is controversial. Gleicher et al. (2015, 2009) demonstrated a negative effect in ovarian reserve of alleles with less than 26 CGGs, evidenced by low levels of anti-Müllerian hormone (AMH) [27–28]. In contrast, Maslow et al. (2016) demonstrated no association between normal-range *FMR1* repeat lengths and reproductive parameters [29]. Wang et al. (2017) demonstrated reduction in *FMR1* mRNA levels in granulosa cells from females carrying alleles with CGGs < 26 and simultaneously a misregulation steroidogenic enzymes and hormone receptors, leading to ovarian dysfunction and ultimately infertility [30]. Rechnitz et al. (2018) illustrated a poor response to ovarian stimulation and elevated expression of *FMR1* mRNA in granulosa cells when compared to samples with different

FMR1 gene genotypes [31]. Interestingly, the majority of our samples with alleles < 26 CGGs show one or no AGGs, while alleles with a repeat size between 29 to 32 CGGs show two AGG interruptions. Lekovich et al. (2018) demonstrated that PM with none or one AGG showed poorer ovarian reserve than those with two, suggesting AGG interspersions have a protective effect [32]. It is thus tempting to speculate that by a similar mechanism the absence of AGGs in normal alleles correlates with ovarian dysfunction.

A minimal effect of *FMR1* allelic complexity with age at amenorrhea is observed in balanced *allelic scores*. Moreover, it appears that the age of amenorrhea decreased with increasing A1 *allelic score* when the PM allele had a score between 70 and 123. Despite the absence of statistical significance, a trend towards a correlation with the *allelic score* of the A1 allele suggests the need for larger-scale investigations to assess the impact of the combined *allelic scores* on the age at amenorrhea. It is likely that the age at amenorrhea may not provide a comprehensive assessment of FXPOI development. Therefore, it is important to test other clinical parameters, such as AMH levels, to gain a deeper understanding of the impact of combining *allelic scores* on disease development.

5. Conclusion

This is the first report investigating the combined effect of normal and PM *allelic scores* on FXPOI development impacted by age at amenorrhea. In our analysis, the presence of the correlation trend indicates the need for further studies and additional samples to explore the complex relationship between *allelic score* combinations and the development of FXPOI. Skewed X-chromosome inactivation and hormonal deregulation were not considered and might impact age at amenorrhea. Nevertheless, the use *allelic scores* combination may pave the way to the identification of an ovulatory dysfunction biomarker. This is of major clinical importance to improve fertility in PM carriers, to make choices about preservation strategies such as oocyte cryopreservation, increasing chances of a successful pregnancy.

List of abbreviations

AGG: Adenine-Guanine-Guanine; AMH: anti-Müllerian hormone; CGG: Cytosine-Guanine-Guanine; *FMR1*: fragile x messenger ribonucleoprotein 1 gene; FSH: Follicle-stimulating hormone; FXPOI: Fragile X-associated Primary Ovarian Insufficiency; FXS:

Fragile X Syndrome; FXTAS: Fragile X-associated Tremor/Ataxia Syndrome; PM: Premutation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12958-024-01227-5.

Acknowledgments

We gratefully acknowledge the Molecular Genetics Laboratory, Laboratory Genetics Service, Genetics and Pathology Clinic, Unidade Local de Saúde de Santo António (ULSSA), without the invaluable help of all collaborators, our work would not have been possible. A special thanks to Rosário Santos, Ph.D., the head of the lab, for providing an environment conducive to the progress of my project, and for the continuous support, availability, and insightful discussions about the research. Thanks, are also due to the Portuguese Foundation for Science and Technology (FCT) and Ministry of Science, Technology and Higher Education (MCTES) for the financial support to Ph.D. project (SFRH/BD/136398/2018 & COVID/BD/153204/2023) and CESAM (UIDP/50017/2020, UIDB/50017/2020 & LA/P/0094/2020).

Authors' contributions

P.J. and A.J.A.N. conceived and designed the study together with B.R. A.J.A.N. performed the statistical analysis with B.R. analysed all data and drafted the manuscript. C.M.Y., F.T., O.V.B., E.G.A., A.G., N.T. and S.L.N. worked on samples and provided the necessary data. C.M.Y., F.T., O.V.B., E.G.A., and V.S. provided critical feedback, and contributed towards the manuscript. All authors discussed the results and critically reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by national funds: FCT/FSE (Fundação para a Ciência e a Tecnologia) – Project Reference SFRH/BD/136398/2018 and COVID/BD/153204/2023 to B. Rodrigues and FCT/FSE – Project Reference EXPL/BIA-REP/0423/2021—X-EPIFERTILITY (10.54499/EXPL/BIA-REP/0423/2021). UMIB - Unit for Multidisciplinary Research in Biomedicine is funded by the FCT Portugal (grant numbers

UIDB/00215/2020, and UIDP/00215/2020), and ITR - Laboratory for Integrative and Translational Research in Population Health (LA/P/0064/2020). DEFI (Departamento de Ensino, Formação e Investigação) – Reference 2015-DEFI/145/12. Thanks are due to FCT/MCTES for the financial support to CESAM (UIDP/50017/2020+UIDB/50017/2020), through national funds.

Data availability

Data is contained within the article or supplementary material.

Declarations

Ethics approval and consent to participate

Informed consent was obtained from all participants within previously published studies. Participants of set 1 the informed consent approved by the clinical ethical committee of Cruces Hospital was obtained in all cases prior to genetic testing. Protocols and consent forms of participants in set 2 were conducted in concordance with the Belmont Report and approved by the Institutional Review Board at Emory University, and written informed consent was obtained from all subjects (IRB00045808). Participants of the set 3 were recruited through the Fragile X Research and Treatment Center at the MIND Institute – University of California Davis, USA. Whole blood was collected according to protocols approved by the Institutional Review Board at the University of California, Davis, and informed consent was obtained from all patients. This research was approved from the Ethics Committee of the Unidade Local de Saúde de Santo António (2020.119 /097-DEFI/099-CE) as part of B. Rodrigues's Ph.D. studies.

Consent for publication

Not Applicable

Availability of data and materials

Data is contained within the article or supplementary material.

Competing Interests

The authors declare no conflict of interest.

References

1. Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. Hum Mol Genet. 2000;9:901–8.

2. 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.

3. Latham GJ, Coppinger J, Hadd AG, Nolin SL. The role of AGG interruptions in fragile X repeat expansions: A twenty-year perspective. Front Genet. 2014;5:1–6.

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

5. Sherman SL. Premature ovarian failure in the fragile X syndrome. Am J Med Genet - Semin Med Genet. 2000;97:189–94.

6. Mailick MR, Hong J, Greenberg J, Smith L, Sherman S. Curvilinear Association of CGG Repeats and Age at Menopause in Women with FMR1 Premutation Expansions. Am J Med Genet B Neuropsychiatr Genet. 2014;0:705–11.

7. Ennis S, Ward D, Murray A. Nonlinear association between CGG repeat number and age of menopause in FMR1 premutation carriers. Eur J Hum Genet. 2006;14:253–5.

8. Tabolacci E, Nobile V, Pucci C, Chiurazzi P. Mechanisms of the FMR1 Repeat Instability: How Does the CGG Sequence Expand? Int J Mol Sci. 2022;23:1–17.

9. Manor E, Gonen R, Sarussi B, Keidar-Friedman D, Kumar J, Tang HT, et al. The role of AGG interruptions in the FMR1 gene stability: A survey in ethnic groups with low and high rate of consanguinity. Mol Genet Genomic Med. 2019;7:1–14.

10. Nolin SL, Glicksman A, Ersalesi N, Dobkin C, Brown WT, Cao R, et al. Fragile X full mutation expansions are inhibited by one or more AGG interruptions in premutation carriers. Genet Med. 2015;17:358–64.

11. Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, et al. AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med. 2012;14:729–36.

12. Domniz N, Ries-Levavi L, Cohen Y, Marom-Haham L, Berkenstadt M, Pras E, et al. Absence of AGG Interruptions Is a Risk Factor for Full Mutation Expansion Among Israeli FMR1 Premutation Carriers. Front Genet. 2018;9:1–8.

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

14. Rodrigues B, Vale-Fernandes E, Maia N, Santos F, Marques I, Santos R, et al. Development and Validation of a Mathematical Model to Predict the Complexity of FMR1 Allele Combinations. Front Genet. 2020;11:1–8.

15. Villate O, Ibarluzea N, Maortua H, de la Hoz AB, Rodriguez-Revenga L, Izquierdo-Álvarez S, et al. Effect of AGG Interruptions on FMR1 Maternal Transmissions. Front Mol Biosci. 2020;7:1–6.

16. Allen EG, Glicksman A, Tortora N, Charen K, He W, Amin A, et al. FXPOI: Pattern of AGG interruptions does not show an association with age at amenorrhea among women with a premutation. Front Genet. 2018;9:1–7.

17. Yrigollen CM, Martorell L, Durbin-Johnson B, Naudo M, Genoves J, Murgia A, et al. AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. J Neurodev Disord. 2014;6:1–11.

18. Zar JH. Bioestatistical Analysis fifth edition. 2010.

19. Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/. 2023.

20. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016.

21. Basuta K, Lozano R, Schneider A, Yrigollen CM, Hessl D, Randi J. Hagerman A, et al. A family with two female compound heterozygous for the FMR1 premutation alleles. Clin Genet. 2014;285:458–63.

22. Allen EG, Charen K, Hipp HS, Shubeck L, Amin A, He W, et al. Refining the risk for fragile X–associated primary ovarian insufficiency (FXPOI) by FMR1 CGG repeat size. Genet Med. 2021;23:1648–55.

23. Yrigollen CM, Sweha S, Durbin-Johnson B, Zhou L, Berry-Kravis E, Fernandez-Carvajal I, et al. Distribution of AGG interruption patterns within nine world populations. Intractable Rare Dis Res. 2014;3:153–61.

24. Nolin SL, Glicksman A, Tortora N, Allen E, Macpherson J, Mila M, et al. Expansions and contractions of the FMR1 CGG repeat in 5,508 transmissions of normal, intermediate, and premutation alleles. Am J Med Genet Part A. 2019;179:1148–56.

25. Friedman-Gohas M, Kirshenbaum M, Michaeli A, Domniz N, Elizur S, Raanani H, et al. Does the presence of AGG interruptions within the CGG repeat tract have a protective effect on the fertility phenotype of female FMR1 premutation carriers? J Assist Reprod Genet. 2020;37:849–54.

26. Bonett DG, Wright TA. Sample size requirements for estimating Pearson, Kendall and Spearman correlations. Psychometrika. 2000;65:23–8.

27. 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;166:502–7.

28. 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;88:1024–30.

29. Maslow BSL, Davis S, Engmann L, Nulsen JC, Benadiva CA. Correlation of normal-range FMR1 repeat length or genotypes and reproductive parameters. J Assist Reprod Genet. 2016;33:1149–55.

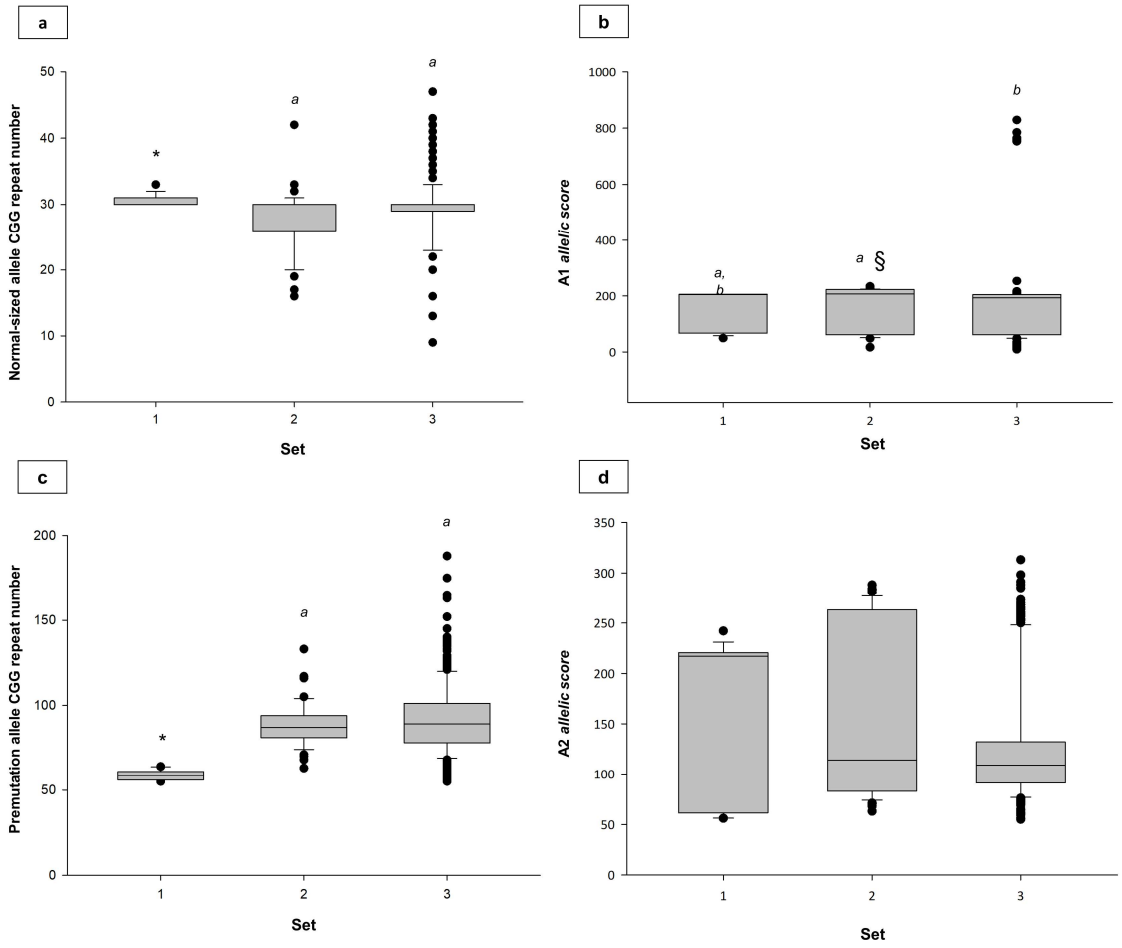
30. Wang Q, Kushnir VA, Darmon S, Barad DH, Wu Y, Zhang L, et al. Reduced RNA expression of the FMR1 gene in women with low (CGGn<26) repeats. Fertil Steril. 2017;108:e143.

31. Rehnitz J, Alcoba DD, Brum IS, Dietrich JE, Youness B, Hinderhofer K, et al. FMR1 expression in human granulosa cells increases with exon 1 CGG repeat length depending on ovarian reserve. Reprod Biol Endocrinol. 2018;16:1–9.

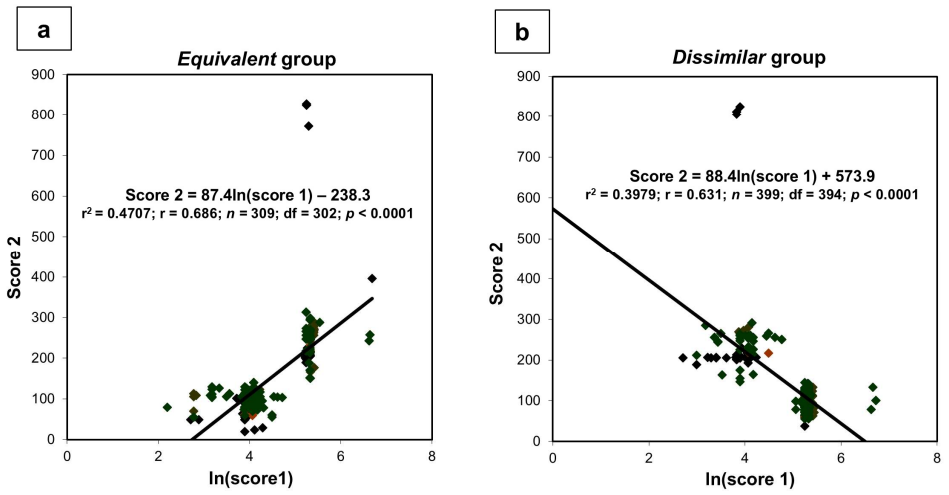
32. Lekovich J, Man L, Xu K, Canon C, Lilienthal D, Stewart JD, et al. CGG repeat length and AGG interruptions as indicators of fragile X–associated diminished ovarian reserve. Genet Med. 2018;20:957–64.

Supporting Information

1. Supplementary figures



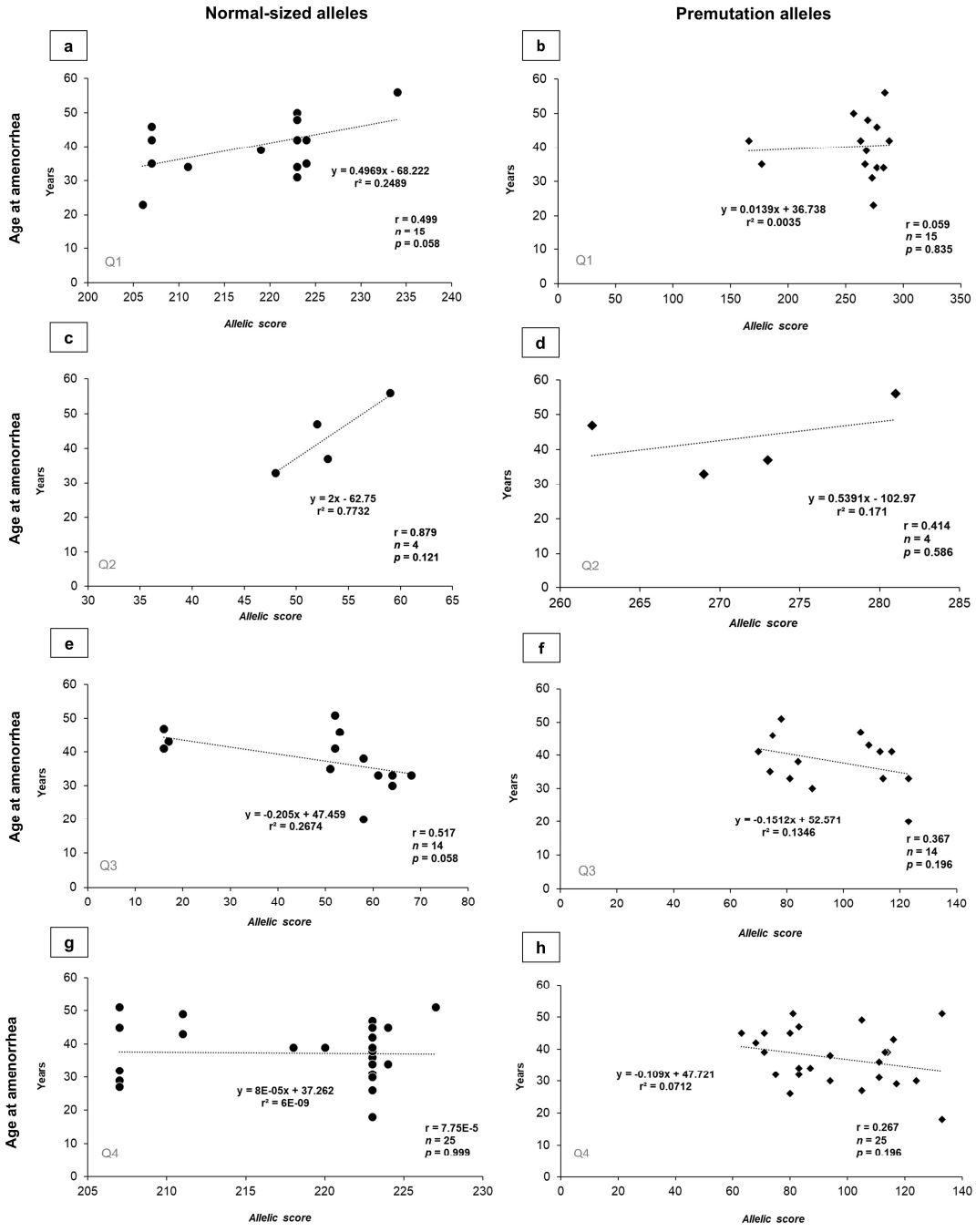
Supplementary Fig. 1 CGG repeat number (a), (c) and *allelic scores* (b), (d) distribution. A1 - Shorter CGG repeat length allele; A2 - Longer CGG repeat length allele; *Statistically significant differences were found in both alleles from set 1 when compared with sets 2 and 3 (Dunn's Method, normal-sized alleles: $p = 0.02$ and $p = 0.016$ respectively; premutation alleles: $p < 0.001$); §Set 2 *allelic score* is statistically different from set 3 (Dunn's Method: $p < 0.001$); a, bNo statistically significant differences (Dunn's Method: $p > 0.05$). See Supplementary Table 2.



Supplementary Fig. 2 Logarithmic models comparison between reference set and sets 1, 2 and 3 in *equivalent* (a) and *dissimilar* (b) groups. The reference set is shown in black, set 1 in brown, set 2 in red, and set 3 in green. Supplementary Table 4 shows the individual linear regression models.

3. Results

Chapter II. *FMR1* allelic complexity in premutation carriers provides no evidence for a correlation with age at amenorrhea



Supplementary Fig. 3 Correlation between the *allelic score* of normal-sized alleles (3a, c, e and g) and premutation (3b, d, f and h) allele with age at amenorrhea in samples from set 2. Qn – represents the four quadrants obtained after *allelic scores* combination (see Fig. 1 - lozenges). No statistically significant correlations were found (all $p > 0.05$).

2. Supplementary tables

Supplementary Table 1. Summary of the data of the reference set.

		A1 - Shorter CGG repeat length allele	A2 - Longer CGG repeat length allele
Number of alleles		131	131
CGG repeat number	Mean (± S.D.)	26.1 ± 5.0	31.4 ± 4.5
	Median	29	30
	Range	15 - 40	20 - 48
Most frequent allele sizes (%, *n*)		30 (35.1%, *n* = 46) 20 (23.7%, *n* = 31)	30 (39.7%, *n* = 52) 31, 32 (11.5%, *n* = 15 in all)
Most common $(CGG)_xAGG$ patterns (%, *n*)		$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (19.8%, *n* = 26) $(CGG)_{10}AGG(CGG)_9$ (14.5%, *n* = 19)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (32.8%, *n* = 43) $(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$ (7.6%, *n* = 10)
Allelic score	Mean (± S.D.)	126.6 ± 75.5	197.5 ± 149.0
	Median	185	205
	Range	15 - 206	20 - 828
	Most frequent (%, *n*)	205 (19.8%, *n* = 26) 49 (16.8%, *n* = 22)	205 (33.6%, *n* = 44) 206 (7.6%, *n* = 10)

S.D. = Standard deviation; % = Percentage; *n* = Number of alleles; x - Number of CGGs; Data obtained from reference set previously published in Rodrigues et al. (2020) [14].

Supplementary Table 2. Summary of the *FMR1* molecular data.

		A1 - Shorter CGG repeat length allele			A2 - Longer CGG repeat length allele		
		Set 1	**Set 2**	**Set 3**	**Set 1**	**Set 2**	**Set 3**
Number of alleles		40	118	995	40	118	997[#]
CGG repeat number	Mean (± S.D.)	30.7 ± 0.9	27.7 ± 5.0	29.2 ± 5.5	58.8 ± 2.9	88.0 ± 13.0	91.8 ± 20.9
	Median	30	30	30	58.5	87	89
	Range	30 - 33	16 - 42	9 - 47	55 - 64	63 - 133	55 - 188
Most frequent allele sizes (%, *n*)		30 (55%, *n* = 11) 31 (25%, *n* = 5)	30 (35.6%, *n* = 21) 29 (16.9%, *n* = 10)	30 (38.8%, *n* = 193) 31 (9.4%, *n* = 47)	56 (25%, *n* = 5) 59 (20%, *n* = 4)	83 (10.2%, *n* = 6) 84, 81, 76 (6.8%, *n* = 4 in all)	101 (3.6%, *n* = 19) 77, 83, 84, 88, 89 (3.4%, *n* = 17 in all)
Most common $(CGG)_xAGG$ patterns (%, *n*)		$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (30%, *n* = 6) $(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$ (25%, *n* = 5)	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$ (27.1%, *n* = 16) $(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$ (18.6%, *n* = 11)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (34.7%, *n* = 173) $(CGG)_9AGG(CGG)_9AGG(CGG)_9$ (12.9%, *n* = 64)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{35}$ (15%, *n* = 3) $(CGG)_9AGG(CGG)_9AGG(CGG)_{37}$ (15%, *n* = 3)	$(CGG)_{83}$ (5.1%, *n* = 3)*	$(CGG)_9AGG(CGG)_{67}$ (2.2%, *n* = 11) $(CGG)_{89}$ (2.2%, *n* = 11)
Allele sizes Kruskal-Wallis Test		H = 12.3; df = 2; *p* = 0.002			H = 51.3; df = 2; *p* < 0.001		

S.D. = Standard deviation; % = Percentage; *n* = Number of alleles; x - Number of CGGs; *Twelve AGG interruption patterns had a frequency of 3.4% (*n* = 2); #One sample with two premutation alleles; *p* values represent significant levels between sets 1, 2 and 3 allele sizes; Multiple Comparison (Dunn's Method) results in Supplementary Fig. 1a and c; Data published in Villate et al. (2020) [15] (set 1), Allen et al. (2018) [16] (set 2) and Yrigollen et al. (2014) [17] (set 3).

Supplementary Table 3. *FMR1* CGG repeat detailed data.

	A1 - Shorter CGG repeat length allele			A2 - Longer CGG repeat length allele		
	$(CGG)_xAGG$ Pattern	Repeat length	*Allelic score*	$(CGG)_xAGG$ Pattern	Repeat length	*Allelic score*
SET 1	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{35}$	56	231
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{35}$	56	231
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{37}$	57	217
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{37}$	57	217
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{38}$	58	218
	$(CGG)_9AGG(CGG)_{22}$	32	58	$(CGG)_{59}$	59	59
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{39}$	59	219
	$(CGG)_{10}AGG(CGG)_9$	30	49	$(CGG)_{61}$	61	61
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{41}$	61	221
	$(CGG)_9AGG(CGG)_{22}$	32	58	$(CGG)_{63}$	63	63
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{15}AGG(CGG)_{38}$	64	242
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{10}$	30	**190**	$(CGG)_{19}AGG(CGG)_{35}$	55	**111**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{56}$	56	**56**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{56}$	56	**56**
	$(CGG)_9AGG(CGG)_{23}$	33	**59**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{35}$	56	**231**
	$(CGG)_{20}AGG(CGG)_9$	30	**89**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{37}$	57	**217**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{59}$	59	**59**
	$(CGG)_9AGG(CGG)_{22}$	32	**58**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{39}$	59	**219**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	**189**	$(CGG)_{62}$	62	**62**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	**189**	$(CGG)_{64}$	64	**64**
SET 2	$(CGG)_{16}$	16	16	$(CGG)_{10}AGG(CGG)_{73}$	84	113
	$(CGG)_{16}$	16	16	$(CGG)_{10}AGG(CGG)_{66}$	76	106
	$(CGG)_{12}AGG(CGG)_{20}$	33	68	$(CGG)_{10}AGG(CGG)_{83}$	94	123
	$(CGG)_{14}AGG(CGG)_8$	23	64	$(CGG)_{89}$	89	89
	$(CGG)_{14}AGG(CGG)_8$	23	64	$(CGG)_{10}AGG(CGG)_{74}$	84	114
	$(CGG)_{11}AGG(CGG)_9$	21	53	$(CGG)_{75}$	75	75
	$(CGG)_{13}AGG(CGG)_6$	20	58	$(CGG)_{10}AGG(CGG)_{83}$	93	123
	$(CGG)_{11}AGG(CGG)_9$	20	52	$(CGG)_{78}$	78	78
	$(CGG)_{10}AGG(CGG)_{21}$	32	61	$(CGG)_{81}$	81	81
	$(CGG)_{17}$	17	17	$(CGG)_{11}AGG(CGG)_{65}$	76	109
	(CGG)10AGG(CGG11	22	51	$(CGG)_{74}$	74	74
	(CGG)11AGG(CGG)8	20	52	$(CGG)_{117}$	117	117
	$(CGG)_{16}$	16	16	$(CGG)_{70}$	70	70
	$(CGG)_{11}AGG(CGG)_{14}$	26	58	$(CGG)_{84}$	84	84
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	223	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{70}$	90	270
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	207	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{77}$	97	277
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	223	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{63}$	83	263
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	223	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{57}$	77	257
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	223	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{69}$	89	269
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	207	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{67}$	89	267
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	207	$(CGG)_{116}$	116	166
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_6$	28	206	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{74}$	94	274
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_8$	31	224	$(CGG)_{177}$	117	177
	$(CGG)_{10}AGG(CGG)_{13}AGG(CGG)_7$	32	219	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{68}$	88	268
	$(CGG)_{10}AGG(CGG)_{11}AGG(CGG)_7$	30	211	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{77}$	97	277
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_8$	31	224	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_{72}$	93	288
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	223	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{83}$	103	283
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	223	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{73}$	93	273
	$(CGG)_{11}AGG(CGG)_{13}AGG(CGG)_6$	32	234	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{84}$	104	284
	$(CGG)_{10}AGG(CGG)_8$	19	**48**	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{89}$	89	**269**
	$(CGG)_{11}AGG(CGG)_8$	20	**52**	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{62}$	82	**262**
	$(CGG)_{10}AGG(CGG)_{13}$	24	**53**	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{73}$	93	**273**
	$(CGG)_{11}AGG(CGG)_{15}$	27	**59**	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{81}$	101	**281**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{10}AGG(CGG)_{77}$	87	**117**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{105}$	105	**105**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{133}$	133	**133**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{71}$	71	**71**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{94}$	94	**94**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{68}$	68	**68**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{83}$	83	**83**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_7$	29	**207**	$(CGG)_{75}$	75	**75**
	$(CGG)_{10}AGG(CGG)_{11}AGG(CGG)_7$	30	**211**	$(CGG)_{10}AGG(CGG)_{65}$	75	**105**
	$(CGG)_{10}AGG(CGG)_{11}AGG(CGG)_7$	30	**211**	$(CGG)_{11}AGG(CGG)_{72}$	83	**116**
	$(CGG)_{10}AGG(CGG)_{13}AGG(CGG)_6$	31	**218**	$(CGG)_{10}AGG(CGG)_{74}$	84	**114**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{10}AGG(CGG)_{71}$	81	**111**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{10}AGG(CGG)_{71}$	81	**111**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{94}$	94	**94**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{80}$	80	**80**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{83}$	83	**83**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{10}AGG(CGG)_{73}$	83	**113**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{80}$	90	**80**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{12}AGG(CGG)_{76}$	88	**124**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{83}$	83	**83**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_7$	30	**223**	$(CGG)_{10}AGG(CGG)_{93}$	103	**133**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_8$	31	**224**	$(CGG)_{13}AGG(CGG)_{81}$	95	**133**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_8$	31	**224**	$(CGG)_{87}$	87	**87**
	$(CGG)_{11}AGG(CGG)_{10}AGG(CGG)_8$	30	**224**	$(CGG)_{63}$	63	**63**
	$(CGG)_{11}AGG(CGG)_{11}AGG(CGG)_7$	31	**227**	$(CGG)_{81}$	81	**81**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_{20}$	42	**220**	$(CGG)_{71}$	71	**71**

	$(CGG)_9$	9	9	$(CGG)_{79}$	79	79
	$(CGG)_{16}$	16	16	$(CGG)_{55}$	55	55
	$(CGG)_{16}$	16	16	$(CGG)_{55}$	55	55
	$(CGG)_{23}$	23	23	$(CGG)_9AGG(CGG)_{73}$	83	109
	$(CGG)_{24}$	24	24	$(CGG)_{103}$	103	103
	$(CGG)_{24}$	24	24	$(CGG)_9AGG(CGG)_{70}$	80	106
	$(CGG)_{24}$	24	24	$(CGG)_{112}$	112	112
	$(CGG)_{24}$	24	24	$(CGG)_{11}AGG(CGG)_{79}$	91	123
	$(CGG)_{24}$	24	24	$(CGG)_9AGG(CGG)_{94}$	104	130
	$(CGG)_{28}$	28	28	$(CGG)_9AGG(CGG)_{90}$	100	126
	$(CGG)_{33}$	33	33	$(CGG)_9AGG(CGG)_{70}$	80	106
	$(CGG)_{35}$	34	35	$(CGG)_{113}$	113	113
	$(CGG)_{35}$	35	35	$(CGG)_{113}$	113	113
	$(CGG)_9AGG(CGG)_{10}$	20	46	$(CGG)_{106}$	106	106
	$(CGG)_9AGG(CGG)_{10}$	20	46	$(CGG)_{106}$	106	106
	$(CGG)_9AGG(CGG)_{12}$	22	48	$(CGG)_{74}$	74	74
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{59}$	59	59
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{79}$	79	79
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{79}$	79	79
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{57}$	67	93
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{60}$	70	96
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{61}$	71	97
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{63}$	73	99
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{63}$	73	99
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{100}$	100	100
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{66}$	76	102
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{67}$	77	103
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{67}$	77	103
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{69}$	79	105
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{108}$	108	108
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_{68}$	79	108
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{111}$	111	111
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{78}$	88	114
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{78}$	88	114
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_{82}$	93	122
	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_{90}$	101	130
	$(CGG)_{10}AGG(CGG)_{11}$	22	51	$(CGG)_{120}$	120	120
S	$(CGG)_{10}AGG(CGG)_{11}$	22	51	$(CGG)_{126}$	126	126
E	$(CGG)_{10}AGG(CGG)_{12}$	23	52	$(CGG)_{79}$	79	79
T	$(CGG)_9AGG(CGG)_{19}$	29	55	$(CGG)_{77}$	74	77
3	$(CGG)_9AGG(CGG)_{19}$	29	55	$(CGG)_{77}$	74	77
	$(CGG)_9AGG(CGG)_{19}$	29	55	$(CGG)_{94}$	94	94
	$(CGG)_9AGG(CGG)_{19}$	29	55	$(CGG)_{94}$	94	94
	$(CGG)_{10}AGG(CGG)_{15}$	26	55	$(CGG)_9AGG(CGG)_{85}$	95	121
	$(CGG)_9AGG(CGG)_{20}$	30	56	$(CGG)_{10}AGG(CGG)_{78}$	89	118
	$(CGG)_9AGG(CGG)_{21}$	31	57	$(CGG)_{77}$	77	77
	$(CGG)_{12}AGG(CGG)_9$	22	57	$(CGG)_{87}$	87	87
	$(CGG)_{12}AGG(CGG)_9$	22	57	$(CGG)_{87}$	87	87
	$(CGG)_{12}AGG(CGG)_9$	22	57	$(CGG)_{87}$	87	87
	$(CGG)_{12}AGG(CGG)_9$	22	57	$(CGG)_{88}$	88	88
	$(CGG)_{12}AGG(CGG)_9$	22	57	$(CGG)_{88}$	88	88
	$(CGG)_{12}AGG(CGG)_9$	22	57	$(CGG)_{88}$	88	88
	$(CGG)_9AGG(CGG)_{22}$	22	58	$(CGG)_{70}$	70	70
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_{73}$	73	73
	$(CGG)_9AGG(CGG)_{22}$	32	58	$(CGG)_{91}$	91	91
	$(CGG)_9AGG(CGG)_{22}$	32	58	$(CGG)_{91}$	91	91
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_{93}$	93	93
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_{100}$	100	100
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_{100}$	100	100
	$(CGG)_9AGG(CGG)_{22}$	32	58	$(CGG)_9AGG(CGG)_{64}$	74	100
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_{74}$	84	110
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_{74}$	84	110
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_{74}$	84	110
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_{74}$	84	110
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_{76}$	86	112
	$(CGG)_9AGG(CGG)_{22}$	32	58	$(CGG)_{122}$	122	122
	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_{10}AGG(CGG)_{82}$	93	122
	$(CGG)_{10}AGG(CGG)_{19}$	30	59	$(CGG)_9AGG(CGG)_{71}$	81	107
	$(CGG)_{10}AGG(CGG)_{19}$	30	59	$(CGG)_{108}$	108	108
	$(CGG)_{10}AGG(CGG)_{19}$	30	59	$(CGG)_9AGG(CGG)_{91}$	101	127
	$(CGG)_{10}AGG(CGG)_{19}$	30	59	$(CGG)_9AGG(CGG)_{91}$	101	127
	$(CGG)_{10}AGG(CGG)_{20}$	31	60	$(CGG)_{73}$	73	73
	$(CGG)_{10}AGG(CGG)_{20}$	31	60	$(CGG)_{73}$	73	73
	$(CGG)_{10}AGG(CGG)_{20}$	31	60	$(CGG)_9AGG(CGG)_{104}$	114	140
	$(CGG)_{10}AGG(CGG)_{20}$	31	60	$(CGG)_9AGG(CGG)_{104}$	114	140
	$(CGG)_{10}AGG(CGG)_{20}$	31	60	$(CGG)_9AGG(CGG)_{104}$	114	140
	$(CGG)_{10}AGG(CGG)_{20}$	31	60	$(CGG)_9AGG(CGG)_{104}$	114	140
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{75}$	75	75
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{75}$	75	75

SET 3	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{78}$	78	78
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{78}$	78	78
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{94}$	94	94
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{94}$	94	94
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{99}$	99	99
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{99}$	99	99
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{104}$	104	104
	$(CGG)_{10}AGG(CGG)_{21}$	23	61	$(CGG)_9AGG(CGG)_{69}$	79	105
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{106}$	106	106
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_9AGG(CGG)_{72}$	82	108
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{110}$	110	110
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{110}$	110	110
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{10}AGG(CGG)_{73}$	84	113
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_9AGG(CGG)_{79}$	89	115
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{120}$	120	120
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{120}$	120	120
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{120}$	120	120
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{123}$	123	123
	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{123}$	123	123
	$(CGG)_{10}AGG(CGG)_{21}$	32	61	$(CGG)_9AGG(CGG)_{90}$	100	126
	$(CGG)_{10}AGG(CGG)_{21}$	32	61	$(CGG)_9AGG(CGG)_{90}$	100	126
	$(CGG)_{10}AGG(CGG)_{21}$	32	61	$(CGG)_9AGG(CGG)_{91}$	101	127
	$(CGG)_{10}AGG(CGG)_{21}$	32	61	$(CGG)_9AGG(CGG)_{91}$	101	127
	$(CGG)_9AGG(CGG)_{27}$	37	63	$(CGG)_9AGG(CGG)_{66}$	76	102
	$(CGG)_9AGG(CGG)_{27}$	37	63	$(CGG)_9AGG(CGG)_{73}$	83	109
	$(CGG)_9AGG(CGG)_{27}$	37	63	$(CGG)_9AGG(CGG)_{73}$	83	109
	$(CGG)_{11}AGG(CGG)_{20}$	32	64	$(CGG)_{81}$	81	81
	$(CGG)_9AGG(CGG)_{28}$	38	64	$(CGG)_{87}$	87	87
	$(CGG)_{10}AGG(CGG)_{24}$	35	64	$(CGG)_{97}$	97	97
	$(CGG)_{11}AGG(CGG)_{21}$	33	65	$(CGG)_{73}$	73	73
	$(CGG)_{14}AGG(CGG)_9$	24	65	$(CGG)_9AGG(CGG)_{61}$	71	97
	$(CGG)_{14}AGG(CGG)_9$	24	65	$(CGG)_{98}$	98	98
	$(CGG)_{14}AGG(CGG)_9$	24	65	$(CGG)_9AGG(CGG)_{83}$	93	119
	$(CGG)_{11}AGG(CGG)_{22}$	34	66	$(CGG)_9AGG(CGG)_{47}$	57	83
	$(CGG)_9AGG(CGG)_{31}$	41	67	$(CGG)_{100}$	100	100
	$(CGG)_9AGG(CGG)_{31}$	41	67	$(CGG)_{100}$	100	100
	$(CGG)_{15}AGG(CGG)_9$	25	69	$(CGG)_9AGG(CGG)_{69}$	79	105
	$(CGG)_9AGG(CGG)_{33}$	43	69	$(CGG)_9AGG(CGG)_{69}$	79	105
	$(CGG)_9AGG(CGG)_{33}$	43	69	$(CGG)_9AGG(CGG)_{69}$	79	105
	$(CGG)_9AGG(CGG)_{33}$	43	69	$(CGG)_9AGG(CGG)_{81}$	91	117
	$(CGG)_{12}AGG(CGG)_{24}$	37	72	$(CGG)_9AGG(CGG)_{59}$	69	95
	$(CGG)_{12}AGG(CGG)_{24}$	37	72	$(CGG)_9AGG(CGG)_{59}$	69	95
	$(CGG)_{13}AGG(CGG)_{22}$	36	74	$(CGG)_{78}$	78	78
	$(CGG)_{13}AGG(CGG)_{22}$	36	74	$(CGG)_{78}$	78	78
	$(CGG)_{20}AGG(CGG)_9$	30	89	$(CGG)_{55}$	55	55
	$(CGG)_{20}AGG(CGG)_9$	30	89	$(CGG)_{60}$	60	60
	$(CGG)_{20}AGG(CGG)_9$	30	89	$(CGG)_{60}$	60	60
	$(CGG)_{21}AGG(CGG)_9$	31	93	$(CGG)_{106}$	106	106
	$(CGG)_{23}AGG(CGG)_9$	33	101	$(CGG)_9AGG(CGG)_{68}$	78	104
	$(CGG)_1AGG(CGG)_2AGG(CGG)_3AGG(CGG)_4$	13	112	$(CGG)_9AGG(CGG)_{67}$	77	103
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{38}$	58	218
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{38}$	58	218
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	$(CGG)_{188}$	188	188
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	$(CGG)_{188}$	188	188
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	$(CGG)_9AGG(CGG)_9AGG(CGG)_{38}$	58	218
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{152}$	152	152
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{152}$	152	152
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{170}$	175	170
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_4AGG(CGG)_{54}$	69	214
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_4AGG(CGG)_{54}$	69	214
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{35}$	55	215
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{35}$	55	215
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{35}$	55	215
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{37}$	57	217
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{37}$	57	217
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{42}$	62	222
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{43}$	63	223
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{43}$	63	223
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{43}$	63	223
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{46}$	66	226
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{35}$	55	215
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{27}$	47	207	$(CGG)_{10}AGG(CGG)_{111}$	120	151
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{64}$	84	244
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{64}$	84	244
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{66}$	86	246
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{66}$	86	246
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{67}$	87	247
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{67}$	87	247

	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{67}$	87	247
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{67}$	87	247
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{68}$	88	248
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{68}$	88	248
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{68}$	88	248
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{68}$	88	248
	$(CGG)_9AGG(CGG)_{11}AGG(CGG)_9$	31	197	$(CGG)_9AGG(CGG)_9AGG(CGG)_{70}$	90	250
	$(CGG)_9AGG(CGG)_{11}AGG(CGG)_9$	31	197	$(CGG)_9AGG(CGG)_9AGG(CGG)_{70}$	90	250
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{13}$	33	193	$(CGG)_9AGG(CGG)_9AGG(CGG)_{72}$	92	252
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{72}$	92	252
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{72}$	92	252
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_{70}$	91	254
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{76}$	96	256
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{76}$	96	256
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{77}$	97	257
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{83}$	103	263
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_{79}$	100	263
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{68}$	89	264
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{86}$	106	266
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{86}$	106	266
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{86}$	106	266
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{86}$	106	266
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{88}$	108	268
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_9AGG(CGG)_{88}$	108	268
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{90}$	110	270
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{90}$	110	270
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{93}$	113	273
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{73}$	93	253	$(CGG)_9AGG(CGG)_9AGG(CGG)_{108}$	128	288
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	$(CGG)_9AGG(CGG)_9AGG(CGG)_{118}$	138	298
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{117}$	138	313
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{117}$	138	313
	$(CGG)_9AGG(CGG)_9AGG(CGG)_6AGG(CGG)_9$	36	753[§]	$(CGG)_9AGG(CGG)_9AGG(CGG)_{63}$	83	243
	$(CGG)_9AGG(CGG)_9AGG(CGG)_6AGG(CGG)_9$	36	753[§]	$(CGG)_9AGG(CGG)_9AGG(CGG)_{63}$	83	243
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9AGG(CGG)_9$	39	765[§]	$(CGG)_9AGG(CGG)_{10}AGG(CGG)74$	95	258
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9AGG(CGG)_9$	39	765[§]	$(CGG)_9AGG(CGG)_{10}AGG(CGG)74$	95	258
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{62}$	62	**62**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{55}$	55	**55**
S	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{55}$	55	**55**
E	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{56}$	56	**56**
T	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{59}$	59	**59**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{59}$	59	**59**
3	$(CGG)_9AGG(CGG)_{11}AGG(CGG)_9$	31	**197**	$(CGG)_{61}$	61	**61**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{62}$	62	**62**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{64}$	64	**64**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{64}$	64	**64**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{64}$	64	**64**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{65}$	65	**65**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{69}$	69	**69**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{70}$	70	**70**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{71}$	71	**71**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{71}$	71	**71**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{71}$	71	**71**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{71}$	71	**71**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{71}$	71	**71**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{72}$	72	**72**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{11}$	32	**207**	$(CGG)_{72}$	72	**72**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{10}$	30	**190**	$(CGG)_{73}$	73	**73**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{73}$	73	**73**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{13}$	33	**193**	$(CGG)_{74}$	74	**74**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{13}$	33	**193**	$(CGG)_{74}$	74	**74**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{74}$	74	**74**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{17}$	37	**197**	$(CGG)_{75}$	75	**75**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{17}$	37	**197**	$(CGG)_{75}$	75	**75**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	42	**201**	$(CGG)_{75}$	75	**75**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{75}$	75	**75**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{75}$	75	**75**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{75}$	75	**75**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{76}$	76	**76**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{76}$	76	**76**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{76}$	76	**76**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{77}$	77	**77**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{77}$	77	**77**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_{77}$	77	**77**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{77}$	77	**77**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{77}$	77	**77**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{21}$	32	**201**	$(CGG)_{78}$	78	**78**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{78}$	78	**78**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{78}$	78	**78**
	$(CGG)_7AGG(CGG)_9AGG(CGG)_9$	27	**157**	$(CGG)_{79}$	79	**79**

SET 3	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{79}$	79	**79**
	(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{80}$	80	**80**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{21}$	41	**201**	(CGG)$_{80}$	80	**80**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{82}$	82	**82**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	9AGG46	56	**82**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{83}$	83	**83**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{83}$	83	**83**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{83}$	83	**83**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{83}$	83	**83**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{83}$	83	**83**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{83}$	83	**83**
	(CGG)$_{10}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	31	**209**	(CGG)$_{84}$	84	**84**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{85}$	85	**85**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{85}$	85	**85**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{85}$	85	**85**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{85}$	85	**85**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{85}$	85	**85**
	(CGG)$_{10}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	31	**209**	(CGG)$_{85}$	85	**85**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{86}$	86	**86**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{86}$	86	**86**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{86}$	86	**86**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{86}$	86	**86**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{87}$	87	**87**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{87}$	87	**87**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{87}$	87	**87**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{88}$	88	**88**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{88}$	88	**88**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{88}$	88	**88**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{9}$AGG(CGG)$_{52}$	62	**88**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{88}$	88	**88**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{89}$	89	**89**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{89}$	89	**89**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{89}$	89	**89**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{89}$	89	**89**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{10}$	31	**206**	(CGG)$_{89}$	89	**89**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{9}$AGG(CGG)$_{55}$	65	**91**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{91}$	91	**91**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{91}$	91	**91**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{91}$	91	**91**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{92}$	92	**92**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{93}$	93	**93**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{58}$	68	**94**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{58}$	68	**94**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{94}$	94	**94**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{60}$	70	**96**
	(CGG)$_{9}$AGG(CGG)$_{10}$AGG(CGG)$_{9}$	30	**193**	(CGG)$_{96}$	96	**96**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{22}$	42	**202**	(CGG)$_{96}$	98	**96**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{22}$	42	**202**	(CGG)$_{96}$	98	**96**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{96}$	96	**96**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{96}$	96	**96**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{96}$	96	**96**
	(CGG)$_{8}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	28	**173**	(CGG)$_{9}$AGG(CGG)$_{61}$	71	**97**
	(CGG)$_{8}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	28	**173**	(CGG)$_{9}$AGG(CGG)$_{61}$	71	**97**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{9}$AGG(CGG)$_{61}$	71	**97**
	(CGG)$_{7}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	27	**157**	(CGG)$_{98}$	98	**98**
	(CGG)$_{7}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	27	**157**	(CGG)$_{98}$	98	**98**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{62}$	72	**98**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{9}$AGG(CGG)$_{62}$	72	**98**
	(CGG)$_{9}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	31	**197**	(CGG)$_{98}$	98	**98**
	(CGG)$_{9}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	31	**197**	(CGG)$_{98}$	98	**98**
	(CGG)$_{9}$AGG(CGG)$_{11}$AGG(CGG)$_{9}$	31	**197**	(CGG)$_{98}$	98	**98**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{100}$	100	**100**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{100}$	100	**100**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{100}$	100	**100**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{9}$AGG(CGG)$_{64}$	74	**100**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{9}$AGG(CGG)$_{64}$	74	**100**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{9}$AGG(CGG)$_{64}$	74	**100**
	(CGG)$_{9}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	29	**189**	(CGG)$_{101}$	101	**101**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{101}$	101	**101**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{101}$	101	**101**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{101}$	101	**101**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{101}$	101	**101**
	(CGG)$_{10}$AGG(CGG)$_{9}$AGG(CGG)$_{9}$	30	**205**	(CGG)$_{101}$	101	**101**

SET 3	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{101}$	101	**101**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{101}$	101	**101**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{102}$	102	**102**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_{102}$	103	**102**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{102}$	102	**102**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{102}$	102	**102**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{67}$	77	**103**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{18}$	38	**198**	$(CGG)_9AGG(CGG)_{67}$	77	**103**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_{103}$	103	**103**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{67}$	77	**103**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{67}$	77	**103**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{67}$	77	**103**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{103}$	105	**103**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{103}$	103	**103**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{104}$	104	**104**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{104}$	104	**104**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{104}$	104	**104**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{69}$	79	**105**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)9AGG(CGG)_{70}$	80	**106**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{106}$	106	**106**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{106}$	106	**106**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{106}$	106	**106**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)9AGG(CGG)_{70}$	80	**106**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{71}$	81	**107**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{71}$	81	**107**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_9AGG(CGG)_{71}$	81	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{71}$	81	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{107}$	107	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{107}$	107	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{67}$	77	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{67}$	77	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{67}$	77	**107**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{107}$	107	**107**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{72}$	82	**108**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{108}$	108	**108**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{72}$	82	**108**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{108}$	108	**108**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{73}$	83	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{73}$	83	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{109}$	109	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{109}$	109	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{73}$	83	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{73}$	83	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_9AGG(CGG)_{73}$	83	**109**
	$(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_9$	31	**209**	$(CGG)_9AGG(CGG)_{73}$	83	**109**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{74}$	84	**110**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{74}$	84	**110**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{74}$	84	**110**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{74}$	84	**110**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{111}$	111	**111**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{111}$	111	**111**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{111}$	111	**111**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{75}$	85	**111**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{111}$	111	**111**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{113}$	113	**113**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{77}$	87	**113**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{78}$	88	**114**
	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	**193**	$(CGG)_9AGG(CGG)_{78}$	88	**114**
	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	**193**	$(CGG)_9AGG(CGG)_{78}$	88	**114**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_{114}$	114	**114**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{74}$	84	**114**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{20}$	41	**216**	$(CGG)_9AGG(CGG)_{78}$	88	**114**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{115}$	115	**115**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{115}$	115	**115**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{79}$	89	**115**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{21}$	41	**201**	$(CGG)_{116}$	116	**116**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{21}$	41	**201**	$(CGG)_{116}$	116	**116**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{116}$	116	**116**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{11}AGG(CGG)_{72}$	84	**116**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{11}AGG(CGG)_{72}$	84	**116**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{81}$	91	**117**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{81}$	91	**117**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{81}$	91	**117**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{81}$	91	**117**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{82}$	92	**118**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{82}$	92	**118**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{10}AGG(CGG)_{78}$	89	**118**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{10}AGG(CGG)_{78}$	89	**118**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{10}AGG(CGG)_{79}$	90	**119**

SET 3	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{10}AGG(CGG)_{79}$	90	**119**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_9AGG(CGG)_{83}$	93	**119**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_{12}$	32	**192**	$(CGG)_{120}$	120	**120**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{84}$	94	**120**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{84}$	94	**120**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{120}$	120	**120**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{120}$	120	**120**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{12}AGG(CGG)_{72}$	85	**120**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{84}$	94	**120**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{10}AGG(CGG)_{81}$	92	**121**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{10}AGG(CGG)_{81}$	92	**121**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{121}$	121	**121**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{121}$	121	**121**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{121}$	121	**121**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	**206**	$(CGG)_{121}$	121	**121**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{86}$	96	**122**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{123}$	123	**123**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{123}$	123	**123**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{123}$	123	**123**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{123}$	123	**123**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG_{)88}$	98	**124**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG_{)88}$	98	**124**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{10}AGG(CGG)_{84}$	95	**124**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{125}$	125	**125**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_9AGG(CGG)_{89}$	99	**125**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{89}$	99	**125**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{89}$	99	**125**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_9AGG(CGG)_{89}$	99	**125**
	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	**193**	$(CGG)_{126}$	126	**126**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{126}$	126	**126**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{127}$	127	**127**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{127}$	127	**127**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{128}$	128	**128**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{10}AGG(CGG)_{88}$	99	**128**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{10}AGG(CGG)_{88}$	99	**128**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{93}$	103	**129**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{93}$	103	**129**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{129}$	129	**129**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{129}$	129	**129**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{129}$	129	**129**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_9AGG(CGG)_{94}$	104	**130**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_9AGG(CGG)_{96}$	106	**132**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{132}$	132	**132**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{132}$	132	**132**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{132}$	132	**132**
	$(CGG)_{10}AGG(CGG)_6AGG(CGG)_9$	27	**193**	$(CGG)_9AGG(CGG)_{97}$	107	**133**
	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	**193**	$(CGG)_{133}$	133	**133**
	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	**193**	$(CGG)_{133}$	133	**133**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{135}$	135	**135**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_9AGG(CGG)_{100}$	110	**136**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{137}$	137	**137**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{138}$	138	**138**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	**205**	$(CGG)_{140}$	140	**140**
	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	**201**	$(CGG)_9AGG(CGG)_{107}$	117	**143**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	**189**	$(CGG)_{145}$	145	**145**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_{111}$	121	**147**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_{119}$	129	**155**
	$(CGG)_{34}$	34	**34**	$(CGG)_{163}$	163	**163**
	$(CGG)_{14}AGG(CGG)_9$	24	**65**	$(CGG)_{165}$	165	**165**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_{175}$	175	**175**
	$(CGG)_{20}$	20	**20**	$(CGG)_9AGG(CGG)_7AGG(CGG)_{40}$	58	**212**
	$(CGG)_{12}AGG(CGG)_{16}$	29	**64**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{35}$	55	**215**
	$(CGG)_9AGG(CGG)_{25}$	35	**61**	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_{34}$	55	**218**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{45}$	65	**225**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{45}$	65	**225**
	$(CGG)_{14}AGG(CGG)_9$	24	**65**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{46}$	66	**226**
	$(CGG)_{14}AGG(CGG)_9$	24	**65**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{46}$	66	**226**
	$(CGG)_{10}AGG(CGG)_{19}$	29	**59**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{47}$	67	**227**
	$(CGG)_{13}AGG(CGG)_9$	23	**61**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{47}$	67	**227**
	$(CGG)_{10}AGG(CGG)_{19}$	29	**59**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{53}$	73	**233**
	$(CGG)_{10}AGG(CGG)_{19}$	29	**59**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{53}$	73	**233**
	$(CGG)_{10}AGG(CGG)_{19}$	29	**59**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{53}$	73	**233**
	$(CGG)_{31}$	31	**31**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{64}$	84	**244**
	$(CGG)_{31}$	31	**31**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{64}$	84	**244**
	$(CGG)_{31}$	31	**31**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{65}$	85	**245**
	$(CGG)_{31}$	31	**31**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{65}$	85	**245**
	$(CGG)_{14}AGG(CGG)_9$	24	**65**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{66}$	86	**246**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{67}$	87	**247**
	$(CGG)_{27}AGG(CGG)_9$	37	**117**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{70}$	90	**250**

SET 3	$(CGG)_{27}AGG(CGG)_9$	37	**117**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{70}$	90	**250**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{73}$	93	**253**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{73}$	93	**253**
	$(CGG)_{12}AGG(CGG)_{10}$	23	**58**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{74}$	94	**254**
	$(CGG)_{29}$	29	**29**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{76}$	96	**256**
	$(CGG)_{29}$	29	**29**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{76}$	96	**256**
	$(CGG)_{14}AGG(CGG)_9$	24	**65**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{76}$	96	**256**
	$(CGG)_{23}AGG(CGG)_{10}$	34	**102**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{76}$	96	**256**
	$(CGG)_{19}AGG(CGG)_9$	29	**85**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{79}$	99	**259**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{80}$	100	**260**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{80}$	100	**260**
	$(CGG)_9AGG(CGG)_{19}$	29	**55**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{80}$	100	**260**
	$(CGG)_{12}AGG(CGG)_{10}$	23	**58**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{80}$	100	**260**
	$(CGG)_{10}AGG(CGG)_{21}$	32	**61**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{81}$	101	**261**
	$(CGG)_{10}AGG(CGG)_9$	20	**49**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{84}$	104	**264**
	$(CGG)_{33}$	33	**33**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{85}$	105	**265**
	$(CGG)_{20}AGG(CGG)_9$	30	**89**	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{70}$	91	**266**
	$(CGG)_{24}$	24	**24**	$(CGG)_9AGG(CGG)_9AGG(CGG)_{105}$	125	**285**
	$(CGG)_{10}AGG(CGG)_{23}$	34	**63**	$(CGG)_{11}AGG(CGG)_9AGG(CGG)_{79}$	101	**291**
	$(CGG)_{10}AGG(CGG)_{23}$	34	**63**	$(CGG)_{11}AGG(CGG)_9AGG(CGG)_{79}$	101	**291**
	$(CGG)_9AGG(CGG)_9AGG(CGG)_7AGG(CGG)_{10}$	38	**758**[§]	$(CGG)_{79}$	79	**79**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9AGG(CGG)_9$	40	**829**[§]	$(CGG)_{101}$	101	**101**
	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9AGG(CGG)_9$	40	**829**[§]	$(CGG)_{101}$	101	**101**
	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_{10}AGG(CGG)_9$	41	**785**[§]	$(CGG)_{133}$	133	**133**

x - number of CGGs; [§]Samples with three AGG interruptions; *Equivalent* (underlined) and *dissimilar* (bold) groups. For detailed information, see Fig. 1.

Supplementary Table 4. Linear regression models of the combination of the *allelic scores*.

	***Equivalent* group**	***Dissimilar* group**
Reference set	Score 2 = 113.21ln(score1) - 361.14 r^2 = 0.3092; r = 0.556; *n* = 74; df = 72; *p* < 0.0001	Score 2 = -95.936ln(score 1) + 605.7 r^2 = 0.1764, r = 0.4199, *n* = 56, *p* = 0.0001
Set 1	Score 2 = 123,13ln(score 1) - 431.47 r^2 = 0.9849; r = 0.992; *n* = 11; df = 9; *p* < 0.0001	Score 2 = -137.95ln(score 1) + 798.92 r^2 = 0.928; r =9634; *n* = 9; df = 7; *p* < 0.0001
Set 2	Score 2 = 77.99ln(score 1) - 176.13 r^2 = 0.7159; r = 0.846; *n* = 29; df = 27; *p* < 0.0001	Score 2 = -121.95ln(score 1) + 753.73 r^2 = 0.8884; r = 0.9495; *n* = 30; df = 28; *p* < 0.0001
Set 3	Score 2 = 79.943ln(score 1) - 208.48 r^2 = 0.7066; r = 0.846; *n* = 194; df = 192; *p* < 0.0001	Score 2 = -82ln(score 1) + 540.81 r^2 = 0.6716; r = 0.8195; *n* = 304; df = 302; *p* < 0.0001

See Supplementary Fig. 2.

Chapter III. Exploring the predictive value of the *FMR1* gene allelic complexity for *in vitro* fertilization success[3]

Abstract

Purpose: We investigated whether *FMR1* allelic complexity, which integrates the CGG repeat length and the number and pattern of AGG interspersions, can be used as a predictor of ovarian reserve and *in vitro* fertilization (IVF) success.

Methods: This was a cohort study that included 124 females with infertility attributed to female factors, who were undergoing intracytoplasmic sperm injection (ICSI). The total CGG repeat lengths and the AGG interspersion patterns of the *FMR1* gene were determined by conventional polymerase chain reaction (PCR) and Triplet-Primed PCR. The allelic complexity (*allelic score*) was determined using a formula previously described.

Results: The combination of the *allelic scores* resulted in a significant stratification of females' samples in the *equivalent* and *dissimilar* groups. No statistically significant differences were observed in ovarian reserve markers and IVF outcomes between two groups. In females from the *dissimilar* group, the *allelic score* of allele 1 significantly negatively correlated with the number of injected metaphase II oocytes and with the number of two pronuclei oocytes.

Conclusions: The correlation found between the allelic complexity of allele 1 and the number of two pronuclei oocytes within females of the *dissimilar* group suggests that *FMR1* allelic complexity can predict IVF success. Females from the *dissimilar* group, thus, appear to be more susceptible to IVF failure than females in the *equivalent* group. Further investigation into the predictive value of the *FMR1* gene and the development of comprehensive genetic markers may offer valuable insights into fertility assessment and assist in optimizing assisted reproductive technologies.

Keywords: *FMR1* allelic complexity, *FMR1* gene, predictor, oocytes with two pronuclei, fertilization success

[3]Rodrigues *et al*., Exploring the predictive value of the *FMR1* gene allelic complexity for *in vitro* fertilization success. Submitted for publication in Journal of Assisted Reproduction and Genetics.

1. Introduction

The fragile X messenger ribonucleoprotein 1 (*FMR1*) gene, located on the long arm of the X chromosome at Xq27.3, plays a significant role in female reproductive health [1]. *FMR1* gene is intricately linked to female infertility through its association with fragile X-associated primary ovarian insufficiency (FXPOI; OMIM #311360) [2,3]. Approximately 20% of female *FMR1* premutation (PM) carriers (55 ≤ CGG repeats ≤ 200) develop FXPOI [4,5]. Typically, these carriers experience hypergonadotropic hypogonadism and have absent or very irregular cycles before the age of 40, therefore increasing the risk of infertility [6–8].

The *FMR1* gene contains a polymorphic region composed of CGG repeats at the 5'UTR. According to the number of the CGG repeats, *FMR1* alleles can be categorized as normal (CGG repeats < 45), intermediate (45–54 CGG repeats), PM (55–200 CGG repeats), and full mutation (> 200 CGG repeats) [9]. Intermediate alleles were associated with idiopathic primary ovarian insufficiency (POI) [10,11]. On the other hand, normal-sized alleles, particularly the low-sized alleles (CGG repeats < 26), have been implicated in reduced *in vitro* fertilization (IVF) pregnancy rates, poor embryonic quality, poor response to ovarian stimulation, diminished ovarian reserve (DOR), and POI in some studies [12–16] while others failed to demonstrate this association [17–20]. Nevertheless, most of the studies that implicated *FMR1* in the increased risk for DOR only considered the total CGG repeat length. Considering that most of non-expanded alleles (CGG repeats < 55) are typically interrupted by AGG triplets, that commonly occur at positions 9 or 10 of the CGG stretches, we previously described a formula that determines the allelic complexity (*allelic score*) of each *FMR1* allele by including the total CGG repeat length and the number and pattern of the AGGs [21]. The length of the CGG repeat and the AGG interspersion pattern are known to play a role in repeat size instability and concomitant risk to offspring [22,23], but their combined effect on infertility remains unknown.

In this study, we aimed to investigate whether *FMR1* allelic complexity can be used as a predictor of ovarian reserve and IVF success. Therefore, samples from females with infertility and who were undergoing intracytoplasmic sperm injection (ICSI) were categorized according to *FMR1* allelic complexity, and correlations with ovarian reserve markers and IVF outcomes were investigated.

2. Material and Methods

2.1. Study design and participants

This was a cohort study that included 124 females with infertility attributed to female factors and who were undergoing ICSI. Females with tubal obstruction were excluded from the study. Females were recruited from the Centre for Medically Assisted Procreation, Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), Unidade Local de Saúde de Santo António (ULSSA) between January 2020 and February 2022.

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the ULSSA (process number 2020.119(097-DEFI/099-CE)). All participants provided written informed consent.

2.2. Demographic and clinical data

Data regarding the age and cause of infertility were obtained for all participants. Data on ovarian reserve markers including follicle-stimulating hormone (FSH) at day 3, anti-Müllerian hormone (AMH), and antral follicle count (AFC]) were obtained. Data regarding IVF outcomes including response to ovarian stimulation (total dose of gonadotrophins, stimulation duration, number of follicles on the trigger day, number of retrieved oocytes, number of immature oocytes and number of aberrant oocytes), oocyte maturation (number of injected metaphase II [MII] oocytes), and fertilization success (number of two pronuclei [2PN] oocytes) were also obtained from all participants. All data were obtained from clinical records.

2.3. *FMR1* CGG repeat region analysis

2.3.1. Total CGG repeat length

The number of CGG repeats was determined using fluorescent polymerase chain reaction (PCR). The forward primer P1- 5'-TTCGGTTTCACTTCCGGTG-3' and the reverse primer P2- FAM labelled - 5'-CCATCTTCTCTTCAGCCCTGC-3' were used. All the PCR components used are detailed in Supplementary Table 1. The amplification protocol involved an initial incubation at 98°C for 5 minutes, followed by 42 cycles of denaturation at 95°C for 1 minute and 10 seconds, annealing at 57°C for 45 seconds, and extension at 68°C for 1 minute and 10 seconds, with a final extension of 10 minutes at 68°C. PCR products were analysed by capillary electrophoresis using ABI PRISM® 3130xl Genetic Analyzer (Applied Biosystems™, Foster City, California, USA) with 500 ROX™ size standard (GeneScan™, Warrington, UK) and were further analysed using

GeneMapper® software version 4.0 (Applied Biosystems™). A previously sequenced sample containing 30 CGGs was used as a control.

The categorization of *FMR1* alleles followed the European Molecular Genetics Quality Network (EMQN) guidelines: normal (CGG repeats < 45), intermediate or "gray-zone" (45 ≤ CGG repeats ≤ 54), PM (55 ≤ CGG repeats ≤ 200), and full mutation (CGG repeats > 200) [9].

2.3.2. AGG interspersion pattern

AGG interruptions were determined by Triplet-Primed PCR (TP-PCR) using the forward primer P1- 5'-GACGGAGGCGCCGCTGCCAGG-3', reserve primer P2- HEX labelled – 5'-TACGCATCCCAGTTTGAGACGGGCCGCCGCCGC-3', and primer "tail" P3- 5'-ACGCATCCCAGTTTGAGACG-3'. All the PCR components are described in Supplementary Table 2. Amplification involved an initial incubation at 98°C for 5 minutes, followed by 15 cycles of denaturation at 98°C for 1 minute, annealing at 55°C for 1 minute, and extension 68°C for 2 minutes, and 30 cycles of denaturation at 98°C for 1 minute, annealing at 55°C for 1 minute, and extension 68°C for 3 minutes. The final extension of 10 minutes was performed at 68°C. PCR products were analysed using capillary electrophoresis as previously described. The AGG interspersion pattern was analysed using the software described above. In samples where AGG interspersion pattern was ambiguous (*n* = 34 samples), the FRAXA PCR kit LabGscan™ (Diagnostica Longwood, Zaragoza, Spain) was used, following the manufacturer's specifications.

2.3.3. *FMR1* allelic complexity

The allelic complexity was calculated separately for each allele using the formula described by Rodrigues et al. (2020) [21]. The formula integrates the number and pattern of AGG interspersions as well as the total CGG repeat length. The resulting value, designated as the *allelic score*, reflects the complexity of the *FMR1* CGG/AGG substructure.

2.3.4. *FMR1* genotypes

The categorization of samples according to *FMR1* genotypes was based on the study by Gleicher et al. (2010a) as follows [24]: normal/normal (N/N), when both alleles were within the new normal range (between 26 and 34 CGG repeats); low/normal (L/N), when one allele was below and the other within the normal range; normal/high (N/H), when one allele was within the normal range and the other was above; low/high (L/H), when

one allele was below and the other was above the normal range; low/low (L/L), when both alleles were below the normal range; and high/high (H/H), when both alleles were above the normal range.

2.4. Statistical analysis

Categorical variables are presented as absolute (*n*) and relative frequencies (%), while continuous variables are presented as the mean ± standard deviation (SD) and range, as indicated in each table. A linearized form of a logarithmic model [i.e., regression of ln(score 1) against score 2] was used to obtain a mathematical model to described the allelic complexity (*allelic score*) relationships between both alleles for each group. Analysis of covariance (ANCOVA) was used to compare the regression models between potentially fertile females (i.e., the reference set which was previously described by Rodrigues et al. (2020) [21] and infertile females, following the methodology outlined by Zar [25]. Principal component analysis (PCA) was performed to test if the ovarian reserve markers and IVF outcomes were able to discriminate between *equivalent* and *dissimilar* groups. For normally distributed data, comparative analysis of the ovarian reserve markers and IVF outcomes between the *equivalent* and *dissimilar* groups were performed using *t*-test. When the data failed normality and homoscedasticity tests, comparisons were performed using the Mann-Whitney test. Pearson correlations were used to assess the relationship between the *allelic score* of each allele, the ovarian reserve markers and IVF outcomes. Chi-square test was used to compare the number of samples (frequency) across the genotype categories between the two groups (*equivalent* and *dissimilar*).

All statistical analyses were performed with SigmaPlot version 14.0 (Systat Software® Inc., Chicago, IL, USA), except for PCA which was performed with Past® version 4.16c (Statistic software) [26]. A statistical significance level of 0.05 was used for all statistical tests.

3. Results

3.1. Demographic and clinical characteristics of the study cohort

This study included 124 infertile females with a mean age of 34.7 ± 3.7 years, ranging from 22 to 40 years. Most females presented multiple causes of infertility, the main causes being ovulatory dysfunction (*n* = 83, 66.8%), endometriosis (*n* = 17, 13.7%), and oocyte factor (*n* = 12, 9.7%). Less frequently (*n* = 12, 9.7%), they presented

hypothyroidism, hyperprolactinemia, POI, DOR, and adenomyosis. Table 1 shows the demographic and clinical characteristics of the study cohort.

Table 1. Demographic and clinical characteristics of the study cohort (*n* = 124 unless otherwise noted).

	Study cohort	
Characteristics	Mean ± SD	Range
Age (years)	34.7 ± 3.7	22 – 40
Markers of ovarian reserve		
Day 3 FSH (mUI/ml)	8.4 ± 11.8 (*n* = 118)	3.3 - 40.0
AMH (ng/ml)	3.1 ± 2.9 (*n* = 123)	0.1 - 18.9
AFC	7.4 ± 3.5 (*n* = 76)	1.0 - 16.0
IVF outcomes		
Response to ovarian stimulation		
Total dose of gonadotrophins (IU/ml)	26582.0 ± 885.3 (*n* = 122)	1025.0 - 7200.0
Stimulation duration (days)	10.2 ± 1.9 (*n* = 123)	5.0 - 16.0
Number of follicles on the trigger day	7.2 ± 4.8 (0 - 21.0) (*n* = 122)	0 - 21.0
Number of retrieved oocytes	11.7 ± 8.5 (*n* = 122)	0 - 46.0
Number of immature oocytes	2.0 ± 2.2 (*n* = 120)	0 - 10.0
Number of aberrant oocytes	0.8 ± 1.6 (*n* = 124)	0 - 12.0
Oocyte maturation		
Number of injected MII oocytes	8.1 ± 5.7 (*n* = 120)	0 - 28.0
Fertilization success		
Number of 2PN oocytes	5.3 ± 4.0 (*n* = 116)	0 - 17.0

2PN, two pronuclei; AFC, antral follicle count; AMH, anti-Müllerian hormone, FSH, follicle-stimulating hormone; MII, metaphase II; n, number of samples; SD, standard deviation.

3.2. *FMR1* gene repeat region characterization

From the 124 samples, 118 presented two normal-sized alleles, ranging from 17 to 44 CGG repeats. Three samples exhibited an intermediate allele with 48, 51, and 52 CGG repeats, while three other samples presented a PM allele with 56, 59, and 75 CGG repeats, respectively (Figure 1).

Considering the 248 alleles, the most frequent CGG repeat length was 30 CGG repeats (*n* = 95, 38.3%). Most alleles presented one or two AGG interruptions (*n* = 232, 93.6%), while 4% of alleles (*n* = 10) presented no AGG interruptions, and 2.4% of alleles (*n* = 6)

presented three AGG interruptions. Overall, 62 distinct AGG interspersion patterns were identified, with the most common being $(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (*n* = 81, 32.7%) and $(CGG)_9GG(CGG)_9AGG(CGG)_9$ (*n* = 25, 10.1%). *FMR1* molecular data can be found in Supplementary Table 3.

The PM alleles were excluded from further analysis because our focus was on normal-sized and intermediate alleles.

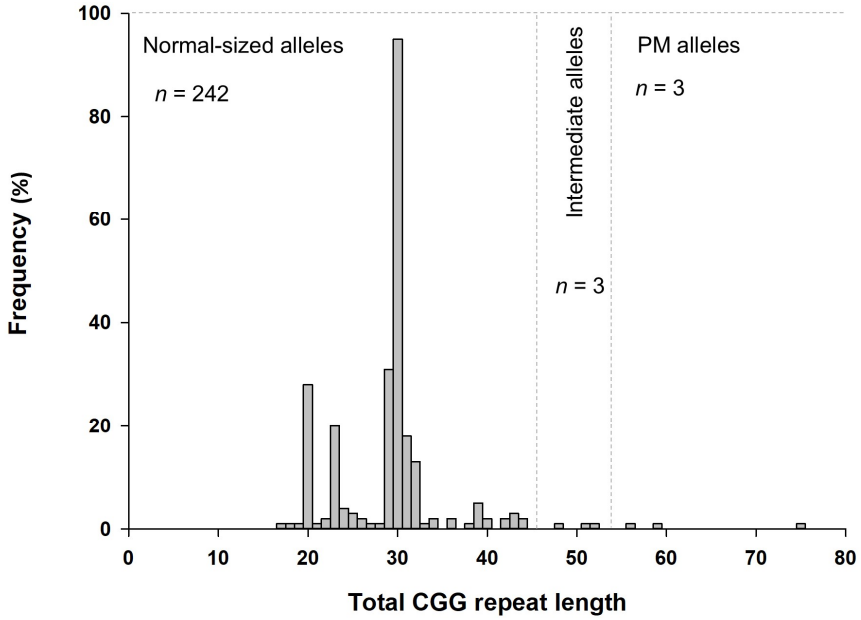


Figure 1. Distribution of total CGG repeat length in the study cohort. The figure depicts the distribution of total CGG repeat lengths among the study participants. The most frequent alleles were characterized by 30 CGG repeats, followed by 29 and 20 CGG repeats. *n* – Number of alleles; PM – Premutation.

3.3. *FMR1 allelic scores* combination and comparison of mathematical models

The summary of the *FMR1* allelic complexity (*allelic score)* results can be found in Table 2. The mean *allelic score* was 125.5 ± 95.6 for allele 1 and 198.9 ± 135.4 for allele 2.

The combination of *allelic scores* resulted in two distinct groups: one containing alleles with similar *allelic scores* (*equivalent* group) and another containing alleles with different *allelic scores* (*dissimilar* group). These were classified as follows: when both alleles presented an *allelic score* > 150 or < 150, the sample was included in the *equivalent* group; when one allele presented an *allelic score* > 150 and the other < 150, the sample was included in the *dissimilar* group. The correlation between the *allelic scores* of each

group was described using a linearized logarithmic model (or mathematical model). Significant correlations were found in the *equivalent* group: r = 0.562; df = 67; $p < 0.0001$ and in the *dissimilar* group: r = -0.417; df = 50; $p = 0.0021$. *Allelic scores* above 700 were obtained in samples with 3 AGG interruptions ($n = 3$ in both the *equivalent* and *dissimilar* groups, Supplementary Table 4, samples 4, 14, 44, 71, 89, and 97, respectively), due to the relevance attributed to the number of AGG interruptions by the formula.

ANCOVA was then used to compare the regression models resulting from the combination of the *allelic scores* of both alleles in each group. Coincident regression lines demonstrated no statistically significant differences in *equivalent* ($F_{(2, 139)} = 0.3023$; $p = 0.7396$) and *dissimilar* ($F_{(2, 99)} = 0.3496$; $p = 0.7058$) groups, when comparing this infertile cohort with the reference set of potentially fertile females (described in Rodrigues et al. (2020) [21]). These results enable the development of a more robust mathematical model that includes all observations: *equivalent* group – Score 2 = -334.6 + 106.6 × ln(score 1) (r = 0.547; df = 141; $p < 0.0001$) and *dissimilar* group – Score 2 = 482.5 - 73.6 × ln(score 1) (r = -0.874; df = 101; $p < 0.0001$).

Table 2. Summary of *FMR1* allelic complexity (*allelic score*) in the study cohort.

	Allele 1 (shorter CGG repeat length)	**Allele 2 (longer CGG repeat length)**
Number of alleles	**121**	**121**
Allelic score		
Mean ± SD	125.5 ± 95.6	198.9 ± 135.4
Median (range)	61.0 (23 - 765)	205.0 (23 - 829)
Most frequent (*n*, %)	205 (34, 28.1)	205 (46, 38.0)
	49 (26, 21.5)	206 (11, 9.1)
	189 (16, 13.2)	201 (10, 8.3)

SD, standard deviation; n, number of alleles; %, percentage.

3.4. Ovarian reserve markers and IVF outcomes according to stratification of *FMR1* allelic complexities

PCA was conducted to test if the ovarian reserve markers and IVF outcomes were able to discriminate between the *equivalent* and *dissimilar* groups. The variables analyzed did not allow a distinct separation between the two groups, as shown in Figure 2. The first principal component (PC1) accounted for 33.4% of the total variance, while the second principal component (PC2) explained an additional 17.3% of the variance. Among the variables analyzed, the number of retrieved oocytes, the number of injected MII oocytes, and the number of 2PN oocytes were identified as the primary contributors to the variance explained by PC1.

In line with these results, no statistically significant differences were found between the *equivalent* and *dissimilar* groups in any of the ovarian reserve markers and IVF outcomes analyzed ($p > 0.05$ for all the variables) (Table 3).

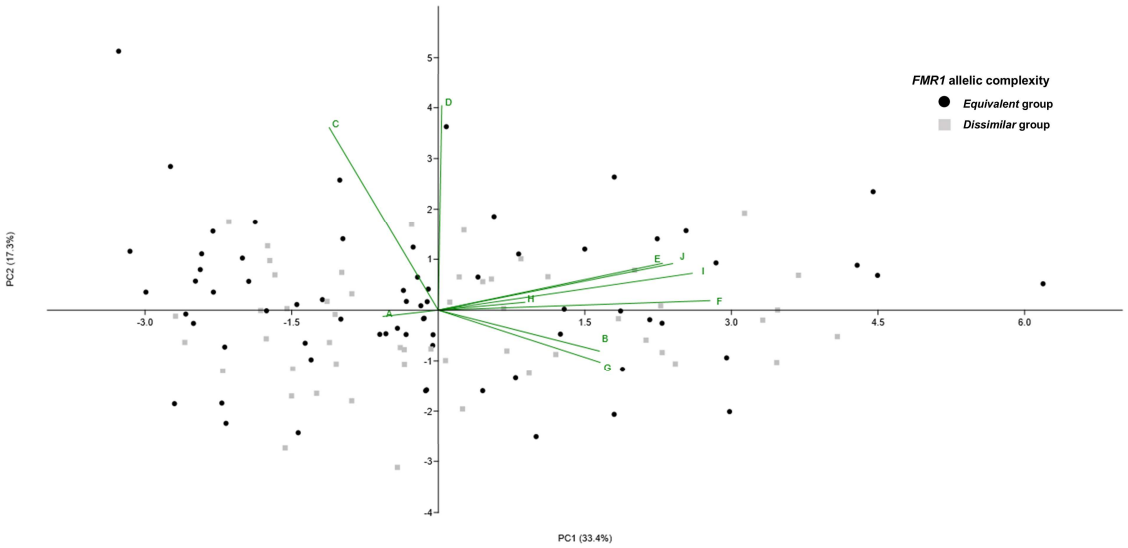


Figure 2. Multivariate statistical analysis results of *FMR1 equivalent* and *dissimilar* groups. The black circles represent samples from the *equivalent* group, while the gray squares represent samples from the *dissimilar* group. The markers of ovarian reserve considered were follicle-stimulating hormone (FSH) (A), anti-Müllerian hormone (AMH) (B), and total dose of gonadotrophins (C). The *in vitro* fertilization (IVF) outcomes evaluated included stimulation days (D), number of follicles on the trigger day (E), number of retrieved oocytes (F), number of immature oocytes (G), number of aberrant oocytes (H), number of injected metaphase II (MII) oocytes (I) and number of oocytes with two pronuclei (2PN) (J).

Table 3. Comparison of the ovarian reserve markers and IVF outcomes between *equivalent* and *dissimilar* groups.

	Equivalent group (*n* = 69[a])	*Dissimilar* group (*n* = 52[a])	*p*-value[b]
Age (years)	34.7 ± 3.9	34.6 ± 3.3	0.550
Markers of ovarian reserve			
Day 3 FSH (mUI/ml)	7.3 ± 2.3 (*n* = 65)	9.8 ± 17.7 (*n* = 51)	0.676
AMH (ng/ml)	2.9 ± 2.4 (*n* = 64)	3.6 ± 3.4 (*n* = 46)	0.238
AFC[c]	6.9 ± 3.3 (*n* = 40)	7.9 ± 3.3 (*n* = 33)	0.205
IVF outcomes			
Response to ovarian stimulation			
Total dose of gonadotrophins (IU/ml)	2811.8 ± 1032.9 (*n* = 68)	2452.5 ± 620.5 (*n* = 51)	0.075
Stimulation duration (days)	10.4 ± 2.0 (*n* = 68)	9.9 ± 1.8	0.08[d]
Number of follicles on the trigger day	6.9 ± 4.7 (*n* = 68)	7.6 ± 5.0 (*n* = 51)	0.451
Number of retrieved oocytes	11.4 ± 9.0 (*n* = 68)	12.0 ± 7.6 (*n* = 51)	0.411
Number of immature oocytes	1.9 ± 2.0 (*n* = 68)	2.4 ± 2.2 (*n* = 49)	0.694
Number of aberrant oocytes	0.9 ± 1.8 (*n* = 68)	0.7 ± 1.5 (*n* = 49)	0.334
Oocyte maturation			
Number of injected MII oocytes	8.0 ± 6.1 (*n* = 68)	8.2 ± 5.0 (*n* = 49)	0.578
Fertilization success			
Number of 2PN oocytes	5.4 ± 4.0 (*n* = 64)	5.1 ± 3.9 (*n* = 48)	0.690

2PN, two pronuclei; AFC, antral follicle count; AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone; MII, metaphase II; *n*, number of samples; SD, standard deviation; [a] Samples size used unless stated otherwise. Values are presented as mean ± SD; [b] *p*-value of statistical test (Mann-Whitney test unless stated otherwise); [c] Variables excluded from PCA; [d] *p*-value of the *t*-test.

3.5. Association of *FMR1* allelic complexity with ovarian reserve markers and with IVF outcomes

We next explored if the *FMR1 allelic score* of each individual allele correlated with the ovarian reserve markers and with the IVF outcomes. In the *equivalent* group, no significant correlations were observed between the *allelic score* of allele 1 or allele 2 and the ovarian reserve markers and IVF outcomes ($p > 0.05$) (Table 4 and Supplementary Table 5). Similarly, no significant correlations were found for allele 2 in the *dissimilar* group (Supplementary Table 5). However, in this group, a significant negative correlation was observed between the *allelic score* of allele 1 and the number of injected MII oocytes (Pearson correlation: r = -0.289, $n = 49$, $p = 0.044$) as well as the number of 2PN oocytes (Pearson correlation: r = -0.311, $n = 48$, $p = 0.031$) (Table 4). Additionally, a significant positive correlation was found between the number of injected MII oocytes and the number of oocytes with 2PN (Pearson correlation: r = 0.859, $n = 48$, $p < 0.001$).

The most frequent CGG repeat length of allele 1 of the *dissimilar* group was 20 CGG repeats (Supplementary Table 4). Considering that alleles with fewer than 26 CGG repeats have been previously associated with a poor fertility prognosis, we analyzed the distribution pattern of the genotypes previously described [27] among the two groups (Figure 3). Most samples from the *dissimilar* group presented alleles with < 26 CGG repeats ($n = 36$, 69.2%), while most samples from the *equivalent* group presented alleles with > 26 CGG repeats ($n = 52$, 75.4%). A statistically significant difference was found in the distribution of the normal/normal and low/normal genotype ($\chi 2 = 35.9$; df = 1; $p < 0.001$).

Table 4. Correlation between *allelic score* of allele 1 with ovarian reserve makers and with IVF outcomes for *equivalent* and *dissimilar* groups.

	***Equivalent* group**			***Dissimilar* group**		
Clinical characteristics	Pearson's correlation coefficient	*p*-value	*n*	Pearson's correlation coefficient	*p*-value	*n*
Markers of ovarian reserve						
Day 3 FSH (mUI/ml)	-0.068	0.588	65	-0.053	0.713	51
AMH (ng/ml)	-0.029	0.819	64	-0.082	0.589	46
AFC	0.015	0.925	40	-0.045	0.806	33
IVF outcomes						
Response to ovarian stimulation						
Total dose of gonadotrophins (IU/ml)	0.096	0.437	68	0.135	0.343	51
Stimulation duration (days)	0.010	0.937	68	-0.078	0.581	52
Number of follicles on the trigger day	0.237	0.0512	68	-0.050	0.726	51
Number of retrieved oocytes	0.060	0.627	68	-0.153	0.283	51
Number of immature oocytes	0.139	0.257	67	-0.168	0.249	49
Number of aberrant oocytes	-0.019	0.876	68	-0.094	0.523	49
Oocyte maturation						
Number of injected MII oocytes	0.0794	0.520	68	-0.289	**0.044**	49
Fertilization success						
Number of 2PN oocytes	0.044	0.731	64	-0.311	**0.031**	48

2PN, two pronuclei; AFC, antral follicle count; AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone; MII, metaphase II; *n*, number of samples; Statistically significant results are marked in bold.

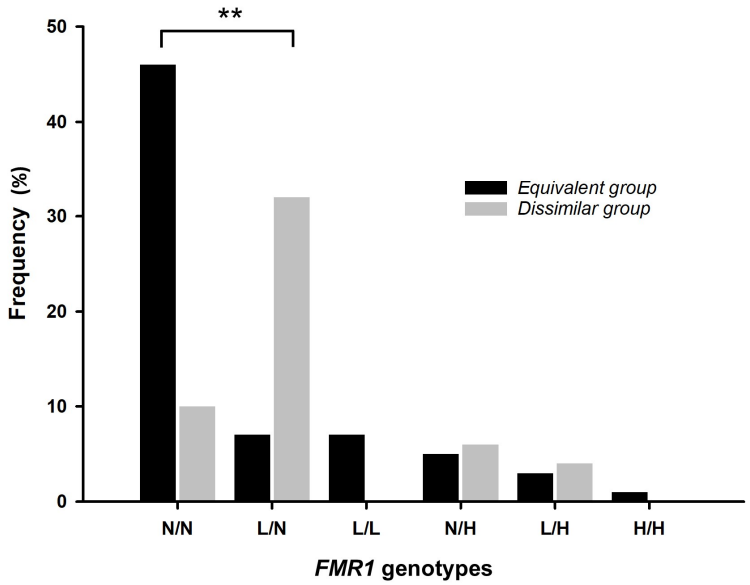


Figure 3. *FMR1* genotype distributions in the *equivalent* and *dissimilar* groups. The black bars represent samples from the *equivalent* group, while the gray bars represent samples from the *dissimilar* group. The genotypes are categorized as follows: N/N – Normal/Normal; L/N – Low/Normal; L/L – Low/Low; N/H – Normal/High; H/H – High/High; L/H – Low/High; $^{**}p < 0.001$.

4. Discussion

In this study, we aimed to investigate whether the *FMR1* gene allelic complexity can be used as a predictor of ovarian reserve and the IVF success. Therefore, we combined the *FMR1 allelic scores* to categorize samples from females with infertility according to allelic complexity, and subsequently we explored its relationship with the ovarian reserve markers and with the IVF outcomes.

The mathematical model derived from the combination of the *allelic scores* of our infertile cohort and was not statistically different from the reference models calculated in the previous study using samples from potentially fertile females [21]. The fact that the combination of the allelic complexity is independent of the clinical condition in each group of samples allows the integration of data from both studies, permitting the development of a more robust models.

We found a significant negative correlation was observed between the *allelic score* of allele 1 (the shorter allele) and the number of 2PN oocytes. The 2PN is the outcome indicative of a successful fertilization [28–30]. This correlation was observed only in the *dissimilar* group that is composed by samples in which the allelic complexity of each

allele is inversely related. Additionally, in cases where the *allelic score* of allele 1 was > 150, very few 2PN oocytes were observed. This observation led us to propose that females falling within this particular subgroup present a higher risk of fertilization failure. One of the females presenting the lowest allelic complexity in allele 1 (*allelic score* = 23) (Supplementary Table 4, sample 114), presented a high rate of fertilization success (12 out of 13 injected MII oocytes presented 2PN) (Supplementary Table 6, sample 114). On the other hand, in the case of the higher allelic complexity (*allelic score* = 205) (Supplementary Table 4, sample 112), only 3 out of 7 injected MII oocytes resulted in 2PN oocytes. Additionally, three oocytes did not fertilize (zero pronuclei) and one degenerated, reflecting poor fertility outcomes (Supplementary Table 6, sample 112). The size range between 35 - 54 CGG repeats has been previously associated with an increased risk of DOR [31,32]. In line with this is the observation that 5 out of 7 samples with high *allelic score* (> 150) in allele 1 have a number of CGG repeats in allele 2 above 34 CGG repeats.

A prior study in polycystic ovary syndrome (PCOS) females revealed disparities in the number of cumulus-oocyte complexes (COCs) and 2PN oocytes when compared to the control group (non-PCOS) [33]. It is thus tempting to speculate that the inherent differences in the study populations from those studies might be explained by the categorization of their samples in distinct groups according to *FMR1* allelic complexity, resorting to our formula. Furthermore, Vale-Fernandes and colleagues (2024) suggested that elevated circulating AMH levels in PCOS patients might not reflect ovarian reserve but rather serve as a marker for disease severity and poorer reproductive outcomes [33]. In fact, in our study, despite the lack of statistical significance, AMH levels tended to be higher in the *dissimilar* group. Other studies have associated alleles with < 26 CGG repeats to DOR, reduced IVF pregnancy rates, poor embryo quality, and poor response to ovarian stimulation [12–14,16,34,35]. Our findings align with previous observations, as the *dissimilar* group was enriched with samples containing fewer than 26 CGG repeats. Nonetheless, allelic complexity enabled us to identify a subgroup at an elevated risk of fertilization failure, underscoring the necessity of considering both the number and pattern of AGG interruptions, as well as the overall CGG repeat size. Furthermore, Quilichini et al. (2024) demonstrated that allelic complexity can serve as an additional predictor of POI risk in females with intermediate and premutated alleles, reinforcing the significance of assessing *FMR1* allelic complexity [36]. Notably, the study by Nunes et al. (2024) reported no association between CGG repeat number and IVF outcomes [37].

This discrepancy might be attributed to their exclusive focus on total CGG repeat length, in contrast to our more comprehensive approach.

Collectively, our findings underscore the critical need to further elucidate the role of *FMR1* gene allelic complexity as a predictor of IVF outcomes. Such knowledge is paramount for developing improved IVF treatments to increase pregnancy and live birth rates.

Curiously, we observed an increased frequency of PM carriers (3 out of 124 cases) compared to the general population [38,39]. PM carriers were referred for genetic counselling due to the risk of FXPOI and other comorbidities, as well as the risk of conceiving children with Fragile X Syndrome (FXS; #300624). Our results justify the need to implement FXS screening protocols in infertile females. The importance of implementing a screening program for females of reproductive age has also been underscored by others [40–42].

Our study is limited by the small number of samples available; future studies employing larger populations are needed to explore of the *FMR1* allelic complexity as a predictor of the fertilization success.

In summary, our findings suggest that *FMR1* allelic complexity can be used as a predictor of fertilization success in females presenting alleles with different *FMR1 allelic scores* (i.e., females from the *dissimilar* group). Additionally, our results underscore the need for FXS screening in infertile females to identify PM alleles, leading to more informed decisions regarding motherhood and the associated risks.

Statements and Declarations

Data availability statement

All data are contained within the article or supplementary material.

Conflict of interest

The authors declare that they have no competing interests.

Funding

This work was supported by the Foundation for Science and Technology /European Social Fund (FCT/ESF) [grant numbers SFRH/BD/136398/2018, COVID/BD/153204/2023 to B. Rodrigues and 10.54499/EXPL/BIA-REP/0423/2021];

UMIB [grant numbers UIDB/00215/2020, UIDP/00215/2020]; ITR [grant number LA/P/0064/2020]; Departamento de Ensino, Formação e Investigação (DEFI) [grant number 2015-DEFI/145/12]; and CESAM [grant numbers UIDP/50017/2020, UIDB/50017/2020, LA/P/0094/2020].

Authors' contributions

P.J., A.J.A.N., and B.R. conceived and designed the study A.J.A.N. and B.R. performed the statistical analysis. B.R. performed the laboratory work and drafted the manuscript. E-V. F. was responsible for recruiting participants and clinical data. V.S., I.M. and R.S. provided critical feedback and contributed to the advancement of the study. All authors have discussed the results and critically reviewed the manuscript.

Acknowledgments

We gratefully acknowledge all the volunteers that enrolled in this research project and all the healthcare professionals involved. Special thanks are due to all the collaborators at the Molecular Genetics Laboratory (Laboratory Genetics Service, Genetics and Pathology Clinic) and the Centre for Medically Assisted Procreation/Public Gamete Bank (Gynaecology Department, Centro Materno-Infantil do Norte Dr. Albino Aroso) both at the ULSSA. Finally, we would like to acknowledge Sofia Nunes, Ph.D. (Scientific ToolBox Consulting, Lisbon, Portugal) for providing technical editing.

References

1. Ma Y, Wei X, Pan H, Wang S, Wang X, Liu X, et al. The prevalence of CGG repeat expansion mutation in FMR1 gene in the northern Chinese women of reproductive age. BMC Med Genet. 2019;20:1–5.

2. 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.

3. Wheeler AC, Bailey DB, Berry-Kravis E, Greenberg J, Losh M, Mailick M, et al. Associated features in females with an FMR1 premutation. J Neurodev Disord. 2014;6:1–14.

4. Allen EG, Glicksman A, Tortora N, Charen K, He W, Amin A, et al. FXPOI: Pattern of AGG interruptions does not show an association with age at amenorrhea among women with a premutation. Front Genet. 2018;9:1–7.

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

6. Allen EG, Charen K, Hipp HS, Shubeck L, Amin A, He W, et al. Refining the risk for fragile X–associated primary ovarian insufficiency (FXPOI) by FMR1 CGG repeat size. Genet Med. 2021;23:1648–55.

7. Fink DA, Nelson LM, Pyeritz R, Johnson J, Sherman SL, Cohen Y, et al. Fragile X Associated Primary Ovarian Insufficiency (FXPOI): Case Report and Literature Review. Front Genet. 2018;9:1–12.

8. Tassone F, Protic D, Allen EG, Archibald AD, Baud A, Brown TW, et al. Insight and Recommendations for Fragile X-Premutation-Associated Conditions from the Fifth International Conference on FMR1 Premutation. Cells. 2023;12:2330.

9. 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;23:417–25.

10. Karimov CB, Moragianni VA, Cronister A, Srouji S, Petrozza J, Racowsky C, et al. Increased frequency of occult fragile X-associated primary ovarian insufficiency in infertile women with evidence of impaired ovarian function. Hum Reprod. 2011;26:2077–83.

11. 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;25:1335–8.

12. Gleicher N, Weghofer A, Oktay K, Barad DH. Can the FMR1 (Fragile X) Gene Serve As Predictor of Response to Ovarian Stimulation? Reprod Sci. 2009;16:462–7.

13. Kushnir VA, Yu Y, Barad DH, Weghofer A, Himaya E, Lee HJ, et al. Utilizing FMR1 gene mutations as predictors of treatment success in human in vitro fertilization. PLoS One. 2014;9:1–9.

14. 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;166:502–7.

15. Lu CL, Li R, Chen XN, Xu YY, Yan LY, Yan J, et al. The 'normal' range of FMR1 triple CGG repeats may be associated with primary ovarian insufficiency in China. Reprod Biomed Online. 2017;34:175–80.

16. Rehnitz J, Alcoba DD, Brum IS, Dietrich JE, Youness B, Hinderhofer K, et al. FMR1 expression in human granulosa cells increases with exon 1 CGG repeat length depending on ovarian reserve. Reprod Biol Endocrinol. 2018;16:1–9.

17. Banks N, Patounakis G, Devine K, DeCherney AH, Widra E, Levens ED, et al. Is FMR1 CGG repeat length a predictor of in vitro fertilization stimulation response or outcome? Fertil Steril. 2016;105:1537-1546.e8.

18. Gleicher N, Kim A, Weghofer A, Shohat-Tal A, Lazzaroni E, Lee HJ, et al. Starting and resulting testosterone levels after androgen supplementation determine at all ages in vitro fertilization (IVF) pregnancy rates in women with diminished ovarian reserve (DOR). J Assist Reprod Genet. 2013;30:49–62.

19. Lledo B, Guerrero J, Ortiz JA, Morales R, Ten J, Llacer J, et al. Intermediate and normal sized CGG repeat on the FMR1 gene does not negatively affect donor ovarian response. Hum Reprod. 2012;27:609–14.

20. Morin SJ, Tiegs AW, Franasiak JM, Juneau CR, Hong KH, Werner MD, et al. FMR1 gene CGG repeat variation within the normal range is not predictive of ovarian response in IVF cycles. Reprod Biomed Online. 2016;32:496–502.

21. Rodrigues B, Vale-Fernandes E, Maia N, Santos F, Marques I, Santos R, et al. Development and Validation of a Mathematical Model to Predict the Complexity of FMR1 Allele Combinations. Front Genet. 2020;11:1–8.

22. Tabolacci E, Nobile V, Pucci C, Chiurazzi P. Mechanisms of the FMR1 Repeat Instability: How Does the CGG Sequence Expand? Int J Mol Sci. 2022;23:1–17.

23. Eichler EE, Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeu SN, et al. Length of uninterrupted CGG repeats determines instability in the FMR1 gene. Nat Genet. 1994;8:88–94.

24. Gleicher N, Weghofer A, Barad DH. Ovarian reserve determinations suggest new function of FMR1 (fragile X gene) in regulating ovarian ageing. Reprod Biomed Online. 2010;20:768–75.

25. Zar JH. Bioestatistical Analysis fifth edition. 2010.

26. Hammer Ø. PAleontological STatistics Version 4.16 - Reference Manual. University of Oslo, Oslo, Norway. URL https://www.nhm.uio.no/english/research/resources/past/. Univ. Oslo. 2024. p. 1–311.

27. Gleicher N, Weghofer A, Lee IH, Barad DH. FMR1 genotype with autoimmunity-associated polycystic ovary-like phenotype and decreased pregnancy chance. PLoS One. 2010;5:1–6.

28. Eppig JJ, O'Brien M, Wigglesworth K. Mammalian oocyte growth and development in vitro. Mol Reprod Dev. 1996;44:260–73.

29. Trebichalská Z, Kyjovská D, Kloudová S, Otevřel P, Hampl A, Holubcová Z. Cytoplasmic maturation in human oocytes: An ultrastructural study. Biol Reprod. 2021;104:106–16.

30. De Vos A, Van De Velde H, Joris H, Van Steirteghem A. In-vitro matured metaphase-I oocytes have a lower fertilization rate but similar embryo quality as mature metaphase-II oocytes after intracytoplasmic sperm injection. Hum Reprod. 1999;14:1859–63.

31. Pastore LM, Young SL, Baker VL, Karns LB, Williams CD, Silverman LM. Elevated Prevalence of 35-44 FMR1 Trinucleotide Repeats in Women With Diminished Ovarian Reserve. Reprod Sci. 2012;19:1226–31.

32. Barasoain M, Barrenetxea G, Huerta I, Télez M, Carrillo A, Pérez C, et al. Study of FMR1 gene association with ovarian dysfunction in a sample from the Basque Country. Gene. 2013;521:145–9.

33. Vale-fernandes E, Moreira M V, Rodrigues B, Pereira S, Leal C. Anti-Müllerian hormone a surrogate of follicular fl uid oxidative stress in polycystic ovary syndrome ? Front Cell Dev Biol. 2024;1–9.

34. 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.

35. 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;88:1024–30.

36. Quilichini J, Perol S, Cuisset L, Grotto S, Fouveaut C, Barbot JC, et al. Stratification of the risk of ovarian dysfunction by studying the complexity of intermediate and premutation alleles of the FMR1 gene. Am J Med Genet Part A. 2024;194:1–10.

37. Nunes ACV, Trevisan CM, Peluso C, Loureiro FA, Dias AT, Rincon D, et al. Low and High-Normal FMR1 Triplet Cytosine, Guanine Guanine Repeats Affect Ovarian Reserve and Fertility in Women Who Underwent In Vitro Fertilization Treatment? Results from a Cross-Sectional Study . DNA Cell Biol. 2024;43:414–24.

38. Tassone F, Iong KP, Tong TH, Lo J, Gane LW, Berry-Kravis E, et al. FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. Genome Med. 2012;4:1–13.

39. Greene R, Pisano MM. Prevalence of CGG Expansions of the FMR1 Gene in a US Population-Based Sample. Birth Defects Res C Embryo Today. 2012;90:133–54.

40. Pesso R, Berkenstadt M, Cuckle H, Gak E, Peleg L, Frydman M, et al. Screening for fragile X syndrome in women of reproductive age. Prenat Diagn. 2000;20:611–4.

41. Bussani C, Papi L, Sestini R, Baldinotti F, Bucciantini S, Bruni V, et al. Premature ovarian failure and fragile X premutation: A study on 45 women. Eur J Obstet Gynecol Reprod Biol. 2004;112:189–91.

42. Rousseau F, Rouillard P, Morel M, Khandjian EW, Morgan K. Prevalence of Carriers of Premutation-Size Alleles of the FMRI Gen and Implications for the Population Genetics of the FragileX Syndrome. Am J Hum Genet. 1995;57:1006–18.

Supporting information

Supplementary Tables

Supplementary Table 1. Fluorescent PCR components and respective final concentration.

Reagents	Final concentration
AccuTaq™ LA 10x Buffer (Sigma-Aldrich®, St. Louis, Missouri, USA)	1 x
Betaine (Sigma-Aldrich®)	1 M
dATP/dCTG/dTTP (Bioline, London, UK)	0.12 mM
dGTP (Bioline)	0.024 mM
7-Deaza-dGTP (Roche®, Basel, Switzerland)	0.44 mM
DMSO (Sigma-Aldrich®)	4.80%
P1- g. *FMR1*_CGG_F (Thermo Fisher Scientific, Waltham, Massachusetts, USA)	0.6 pmol/µl
P2- g. *FMR1*_CGG_R* (Thermo Fisher Scientific)	0.6 pmol/µl
AccuTaq™ LA DNA Polymerase (Sigma-Aldrich®)	0.08 U/µl
gDNA	150 ng
dH_2O	Up to 25 µl

CGG, cytosine-guanine-guanine; dATP, deoxyadenosine triphosphate; dCTG deoxycytidine triphosphate, dGTP, deoxyguanosine triphosphate; dH2O; distilled water; DMSO, dimethyl sulfoxide; dTTP, 2′-deoxythymidine 5′-triphosphate; F, forward; *FMR1*, fragile X messenger ribonucleoprotein 1; gDNA, genomic DNA; LA, long and accurate; P1; primer 1; P2, primer 2; R, reverse; UK, United Kingdom; USA, United Stated of America.

Supplementary Table 2. Triplet-primed-PCR (TP-PCR) components and respective final concentration.

Reagents	Final concentration
PCR Master Mix (Promega®, Madison, Wisconsin, USA)	1 x
Betaine (Sigma-Aldrich®, St. Louis, Missouri, USA)	0.6 M
7-Deaza-dGTP (Roche®, Basel, Switzerland)	0.4 mM
Q-Solution® (Qiagen®, Hilden, Germany)	0.5 x
DMSO (Sigma-Aldrich®)	10%
P1- g. *FMR1*_TP-PCR_F (Thermo Fisher Scientific, Waltham, Massachusetts, USA)	0.4 pmol/µl
P2- g. *FMR1*_TP-PCR_GCC* (Thermo Fisher Scientific)	0.2 pmol/µl
P3- g. *FMR1*_TP-PCR_R (Thermo Fisher Scientific)	0.2 pmol/µl
gDNA	150 ng
dH_2O	Up to 25 µl

dGTP, deoxyguanosine triphosphate; dH2O; distilled water; DMSO, dimethyl sulfoxide; F, forward; *FMR1*, fragile X messenger ribonucleoprotein 1; gDNA, genomic DNA; P1, primer 1; P2, primer 2; PCR, polymerase chain reaction; R, reverse; TP, triplet-primed; USA, United Stated of America.

Supplementary Table 3. Summary of the *FMR1* CGG repetitive region data.

	Allele 1 (shorter CGG repeat length)	Allele 2 (longer CGG repeat length)	Both alleles
Number of alleles	124	124	248
Total CGG repeat length			
Mean ± SD	26.0 ± 4.4	32.6 ± 7.2	29.3 ± 6.8
Median (range)	29.0 (17 - 39)	30.0 (20 - 75)	30 (17 - 75)
Most frequent (*n*, %)	30 (44, 35.5)	30 (51, 41.1)	30 (95, 38.3)
	20 (26, 21.0)	31 (17, 13.7)	29 (31, 12.5)
	29 (20, 16.1)	32 (12, 10.1)	20 (28, 11.3)
Most common $(CGG)_x$ AGG patterns (*n*, %)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (35, 28.2)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (46, 37.1)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$ (81, 32.7)
	$(CGG)_{10}AGG(CGG)_9$ (22, 17.7)	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$ (11, 8.9)	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$ (25, 10.1)
	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$ (17, 13.7)	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$ (9, 7.3)	$(CGG)_{10}AGG(CGG)_9$ (24, 9.7)

AGG, adenine-guanine-guanine; CGG, cytosine-guanine-guanine; *n*, number of alleles; SD, Standard deviation; x, number of CGGs.

Supplementary Table 4. *FMR1* CGG repeat detailed data.

	Allele 1 (shorter CGG repeat length)			Allele 2 (longer CGG repeat length)			***FMR1* sub-genotypes**
Sample	$(CGG)_xAGG$ Pattern	Repeat length	*Allelic score*	$(CGG)_xAGG$ Pattern	Repeat length	*Allelic score*	
1	$(CGG)_{23}$	23	23	$(CGG)_{12}AGG(CGG)_{10}$	23	58	Low/Low
2	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
3	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
4	$(CGG)_9AGG(CGG)9AGG(CGG)_9AGG(CGG)_9$	39	765[§]	$(CGG)_9AGG(CGG)_9AGG(CGG)_{22}$	42	202	High/High
5	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
6	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
7	$(CGG)_{10}AGG(CGG)_{11}$	22	51	$(CGG)_{13}AGG(CGG)_{16}$	30	68	Low/Normal
8	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{12}AGG(CGG)_{10}$	23	58	Low/Low
9	$(CGG)_9AGG(CGG)_9AGG(CGG)_{10}$	30	190	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
10	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9$	20	49	Low/Low
11	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{24}$	44	220	Normal/High
12	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
13	$(CGG)_9AGG(CGG)_9AGG(CGG)_8$	28	188	$(CGG)_9AGG(CGG)_9AGG(CGG)_{10}$	30	190	Normal/Normal
14	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG_9AGG(CGG)_9$	39	829[§]	Normal/High
15	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_{24}$	44	204	Normal/High
16	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
17	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_{34}$	34	34	Low/Normal
18	$(CGG)_9AGG(CGG)_9AGG(CGG)_{10}$	30	190	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{22}$	43	218	Normal/High
19	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
20	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
21	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9$	20	49	Low/Low
22	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
23	$(CGG)_{10}AGG(CGG)_7AGG(CGG)_{10}$	29	198	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
24	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{20}AGG(CGG)_9$	30	89	Low/Normal
25	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
26	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Normal/Normal
27	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal

28	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	Normal/Normal
29	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_{21}$	31	57	Low/Normal
30	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Normal/Normal
31	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Normal/Normal
32	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Normal/Normal
33	$(CGG)_{10}AGG(CGG)_9$	20	49	(CGG)25AGG(CGG)10	36	110	Low/High
34	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
35	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	189	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	189	Normal/Normal
36	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Normal/Normal
37	$(CGG)_9AGG(CGG)_{10}$	20	46	$(CGG)_{10}AGG(CGG)_8AGG(CGG)_9$	29	57	Low/Normal
38	$(CGG)_9AGG(CGG)_{13}$	23	49	$(CGG)_{10}AGG(CGG)_{27}$	38	67	Low/High
39	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
40	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_{12}AGG(CGG)_9$	33	217	Normal/Normal
41	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
42	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Normal/Normal
43	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
44	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)9AGG(CGG)_9AGG(CGG_{)9}AGG(CGG)_9$	39	765[§]	Normal/High
45	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
46	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
47	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
48	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Normal/Normal
49	(CGG)10AGG(CGG)9AGG(CGG)9	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
50	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Normal/Normal
51	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{24}AGG(CGG)_9$	34	105	Low/Normal
52	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
53	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	189	$(CGG)_9AGG(CGG)_{11}AGG(CGG)_9$	32	197	Normal/Normal
54	$(CGG)_9AGG(CGG)_9AGG(CGG)_{10}$	30	190	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	31	193	Normal/Normal
55	$(CGG)_{24}$	24	24	$(CGG)_9AGG(CGG)_{15}$	25	51	Low/Low
56	$(CGG)_9AGG(CGG)_{13}$	23	49	$(CGG)_{10}AGG(CGG)_{20}$	31	60	Low/Normal
57	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
58	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Normal/Normal
59	$(CGG)_{29}$	29	29	$(CGG)_{10}AGG(CGG)_{20}$	31	60	Normal/Normal
60	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{13}AGG(CGG)_9$	23	61	Low/Low
61	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	30	193	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{11}$	32	207	Normal/Normal

62	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{11}$	32	207	Normal/Normal
63	$(CGG)_{13}AGG(CGG)_9$	23	61	(CGG)9AGG(CGG)33	43	69	Low/High
64	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
65	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
66	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Normal/Normal
67	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	29	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	29	205	Normal/Normal
68	$(CGG)_{10}AGG(CGG)_9$	20	49	(CGG)23	23	23	Low/Low
69	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
70	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{26}$	36	62	Low/Normal
71	$(CGG)_9AGG(CGG)_{19}$	29	55	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9AGG(CGG)_9$	39	829[§]	Normal/High
72	$(CGG)_9AGG(CGG)_{10}$	20	46	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{22}$	43	218	Normal/High
73	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_{20}$	31	60	Low/High
74	$(CGG)_9AGG(CGG)_{13}$	23	49	$(CGG)_9AGG(CGG)_9AGG(CGG)_{22}$	42	202	Normal/Normal
75	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_9AGG(CGG)_{32}$	51	212	Low/High
76	$(CGG)_9AGG(CGG)_{20}$	30	56	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Low/High
77	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{30}$	40	66	Normal/Normal
78	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/High
79	$(CGG)_8AGG(CGG)_9$	18	41	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
80	$(CGG)_{10}AGG(CGG)_{19}$	30	59	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
81	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
82	$(CGG)_{25}$	25	25	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	Low/Normal
83	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Low/Normal
84	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{11}$	32	207	Low/Normal
85	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	Low/Normal
86	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Low/Normal
87	$(CGG)_{16}AGG(CGG)_9$	26	73	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Low/Normal
88	$(CGG)_{23}$	23	23	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	Normal/Normal
89	$(CGG)_{25}$	25	25	$(CGG)_9AGG(CGG)_9AGG(CGG)_9AGG(CGG)_9$	39	765[§]	Low/Normal
90	$(CGG)_{10}AGG(CGG)_{19}$	30	59	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/High
91	$(CGG)_{11}AGG(CGG)_9$	21	53	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Normal/Normal
92	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Low/Normal
93	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Low/Normal
94	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
95	$(CGG)_9AGG(CGG)_{10}$	20	46	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal

96	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
97	$(CGG)_{10}AGG(CGG)_6$	17	46	$(CGG)_{10}AGG(CGG)_6AGG(CGG)AGG(CGG)_{10}$	30	750[§]	Low/Normal
98	$(CGG)_{10}AGG(CGG)_{15}$	26	55	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	29	205	Low/Normal
99	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_9AGG(CGG)_{30}$	40	66	Normal/Normal
100	$(CGG)_9AGG(CGG)_{17}$	27	49	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_9$	29	193	Normal/High
101	$(CGG)_{24}$	24	24	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
102	$(CGG)_9AGG(CGG)_9$	19	45	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
103	$(CGG)_9AGG(CGG)_{12}$	22	48	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
104	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
105	$(CGG)_9AGG(CGG)_{10}$	20	46	$(CGG)_{10}AGG(CGG)_8AGG(CGG)_9$	29	201	Low/Normal
106	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	Low/Normal
107	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
108	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
109	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
110	$(CGG)_9AGG(CGG)_{19}$	29	55	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	30	189	Normal/Normal
111	$(CGG)_{10}AGG(CGG)_9$	20	49	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Low/Normal
112	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{20}AGG(CGG)_9$	30	89	Normal/Normal
113	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	$(CGG)_{10}AGG(CGG)_{42}$	52	82	Normal/High
114	$(CGG)_{23}$	23	23	$(CGG)_9AGG(CGG)_{12}AGG(CGG)_9$	32	201	Low/Normal
115	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_9AGG(CGG)_9AGG(CGG)_9$	29	189	Low/Normal
116	$(CGG)_9AGG(CGG)_{14}$	24	50	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
117	$(CGG)_9AGG(CGG)_{10}AGG(CGG)_{10}$	31	194	$(CGG)_{10}AGG(CGG)_{37}$	48	77	Normal/High
118	(CGG)20AGG(CGG)9	30	89	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Normal/Normal
119	$(CGG)_{10}AGG(CGG)_{13}$	24	53	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_{10}$	31	206	Low/Normal
120	$(CGG)_{13}AGG(CGG)_9$	23	61	$(CGG)_{10}AGG(CGG)_9AGG(CGG)_9$	30	205	Low/Normal
121	$(CGG)_{12}AGG(CGG)_{10}$	23	58	$(CGG)9AGG(CGG)_9AGG(CGG)_{11}$	31	191	Low/Normal

AGG, adenine-guanine-guanine; CGG, cytosine-guanine-guanine; *FMR1*, fragile X messenger ribonucleoprotein 1; x, number of CGGs; [§]Samples with three AGG interruptions; *Equivalent* (white background) and *dissimilar* (gray background) groups.

Supplementary Table 5. Correlation between the *allelic score* of allele 2, ovarian reserve markers, and IVF outcomes for *equivalent* and *dissimilar* groups.

	***Equivalent* group**			***Dissimilar* group**		
Clinical characteristics	Pearson´s correlation coefficient	*p*-value	*n*	Pearson´s correlation coefficient	*p*-value	*n*
Markers of ovarian reserve						
Day 3 FSH (mUI/ml)	-0.0280	0.825	65	0.00533	0.970	51
AMH (ng/ml)	-0.119	0.349	64	-0.0402	0.791	46
AFC	0.122	0.455	40	-0.0934	0.605	33
IVF outcomes						
Response to ovarian stimulation						
Total dose of gonadotrophins (IU/ml)	0.0417	0.736	68	-0.0822	0.567	51
Stimulation duration (days)	0.0199	0.872	68	0.0434	0.760	52
Number of follicles on the trigger day	0.0399	0.747	68	-0.0162	0.910	51
Number of retrieved oocytes	-0.101	0.414	68	0.0231	0.872	51
Number of immature oocytes	0.0522	0.672	67	0.152	0.299	49
Number of aberrant oocytes	0.0315	0.799	68	0.210	0.148	49
Oocyte maturation						
Number of injected MII oocytes	-0.0913	0.459	68	0.000245	0.999	49
Fertilization success						
Number of 2PN oocytes	-0.149	0.240	64	-0.0959	0.517	48

2PN, two pronuclei; AFC, antral follicle count; AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone; MII, metaphase II; *n*, number of samples.

Supplementary Table 6. Detailed description of ovarian reserve markers and IVF outcomes for each sample.

Sample	Day 3 FSH (mIU/ml)	AMH (ng/ml)	AFC	Total dose of gonadotrophins (IU/ml)	Stimulation duration (days)	Number of follicles on the trigger day	Number of retrieved oocytes	Number of immature oocytes	Number of aberrant oocytes	Number of injected MII oocytes	Number of oocytes with 2PN
1	5.3	4.2	4	2400	11	7	33	5	4	24	15
2	6.2	1.5	6	3300	11	11	27	7	2	18	8
3	5.7	4.4		1025	7	11	22	2	0	10	8
4	6.1	2.1	5	4500	11	11	14	3	0	11	8
5	4.2	2.7		2400	11	7	9	0	0	9	8
6	10.6	0.7		5550	13	12	13	1	1	12	8
7	7.5	1.9	12	2175	10	11	22	5	0	17	7
8	12.9	0.6	4	4800	10	1	2	0	0	2	1
9	4.9	2.8	7	1575	7	5	15	8	0	7	6
10	6.8	1.4	5	3600	12	2	3	0	0	3	2
11	5.1	1.4	12	2700	9	3	3	1	0	2	1
12	6.3		5	4725	11	3	1	0	1	0	NA
13	4	4.9		2925	13	16	21	3	12	6	3
14		0.9		4500	15	4	2	0	1	1	1
15	10.7	0.8	4	3300	11	5	3	2	0	1	0
16	8	1.4	5	3450	12	6	3	1	0	2	2
17	6.57	11.2		2475	11	7	4	1	0	3	2
18	8.6	1.5	12								
19	5			2450	10	13	5	0	0	5	0
20	7		8	2900	9	11	24	6	2	16	5
21	5.2	4.5		1650	11	5	10	1	0	9	5
22	13.9	0.1	1	3150	11	1	1	0	0	1	1
23	6	2.6		1800	8	8	10	3	2	5	3
24	9.8	3.4	8	3500	12	21	26	4	1	21	17
25	7.5	4.7		2625	12	8	11	2	0	9	7
26		1.6	3	2700	12	7	10	3	0	7	6

27	5.4	2.6		1800	8	5	14	4	1	9	5
28	5.2	2.4	12	3150	14	11	17	0	0	17	14
29	9.4	0.7	6	3000	10	5	7	1	1	5	5
30	7.4	1.3	8	3750	13	3	5	0	1	4	4
31	11	0.6	8	2100	7	1	1	0	0	1	1
32	4.1	3.9		1575	7	12	25	7	4	14	5
33	7.5	0.7	4	7200	16	2	3	0	0	3	2
34	7.3	4.9		2250	10	6	8	0	0	8	5
35	9.6	3		3000	10	5	11	2	1	8	7
36	5.7	3.1	8	3000	10	11	8	2	0	6	4
37	7.2	3.1	16	2250	10	3	10	1	2	7	4
38	7.1	1.3		2250	10	5	12	1	0	11	6
39	3.3	5.2		2700	9	8	17	0	0	17	8
40	5	4.8		2025	9	5	13	5	6	8	2
41	5.9	6.2	5	2475	11	16	24	4	2	19	16
42	11.6	0.7	6	3000	10	6	7	1	1	5	2
43	10.3	2.2	6	3600	12	4	2	2	0	0	NA
44	7.4		8	1650	6	4	4	0	1	4	2
45	9.1	0.9	5	4200	14	5	10	1	0	9	5
46	5.1	5.5		1525	12	12	16	3	2	11	7
47	5.5	8.2		2250	10	6	14	2	2	10	8
48		2.8	4	2400	8	0	13	6	0	7	5
49	8.8	2.7	13	2250	10	0	7	1	1	6	3
50	7.3	3.5		3162.5	12	12	22	3	1	18	11
51	4.8	1	8	3600	12	5	9	1	0	8	5
52	4.9	1.5	8	2400	8	4	7	0	1	6	3
53	6.7	2.6		2700	12	10	14	1	1	12	8
54	9	2.9	8	2887.5	11	3	5	2	0	3	1
55	7.7	0.8	4	3300	11	3	2	0	0	2	2
56		0.3	2	1800	6	1	1	0	1	0	NA
57	7.6	2.3	10	2850	10	4	1	0	0	1	0
58	6.7	1.1	7	3300	11	5	4	1	0	3	2
59	10.5	1	6	2100	7	0	1	1	0	0	NA
60	8.5	0.5	3	3000	10	1	3	0	1	2	1

61	6.1			1750	10	19	46	4	0	28	16
62	11.6	1.7		2250	10	6	10	3	0	7	5
63	4.4	10.9		2025	9	13	25	3	2	10	8
64	8	1.9		3300	12	10	9	2	0	7	6
65	7.4	3.7		2700	12	12	16	1	2	13	9
66	7.3	6.7		2500	10	11	6	0	0	6	5
67	7	7		1350	9	11	23	1	0	11	7
68	6.4	5.6		1650	11	0	16	1	1	6	6
69	11.7	1.4	10	3000	10	8	10	1	1	8	6
70	8.1	2.7	10	1750	12	8	11	1	1	9	5
71	12.5	0.8		2475	11	5	8	3	0	5	3
72	8.4	1.5	6	2100	9	2	9	2	0	7	3
73	7			1800	6	3	4	3	0	1	1
74	4.7	10.5		2025	9						
75	5.3	2.3	14	3300	12	8	12	0	0	12	9
76	6.6		5	1800	10	14	18	1	0	17	11
77	6.7			1300	8	6	6	0	0	6	3
78	7.4	3.3	14		8	8	21	2	0	4	0
79	8.8	1.5	4	2400	8	3	3	0	0	3	1
80	9.3	1.1	10	2250	10	2	6	1	0	0	NA
81	10.9	2.8	14	1800	8	4	14	6	1	7	3
82	10.9	0.6	4	3300	11	5	5	1	0	4	4
83	5.5	2.6	8	3300	12	2	5	0	0	5	4
84		4.3	10	2025	9	5	10	2	1	8	6
85	5	2.8		2700	12	6	6	0	0	6	5
86	6.7	1.7	10	1950	7	4	6	0	0	6	4
87	8.4	5.3		1875	9	11	15	2	0	13	5
88	5.9	3.3		1800	8	10	20	4	1	15	11
89	4.1	11.8		1800	12	14	20	4	2	14	9
90	9	2.7		2025	9	12	11	1	0	10	1
91	7.5	0.9	4	2700	9	6	7	3	0	4	3
92	9.4	3.7	12	3000	10	8	15	4	1	10	5
93	6.4	1.6		1975	10	12	24	7	0	17	13
94	6.2	0.9	6	3900	13	0	7	0	0	7	2
95	6.7	4	14	2700	9	10	13	6	2	5	4

96	11.6	2	7	3300	11	6	13	3	8	10	1
97	6.3	2.7	5	1500	10	10	18	4	1	14	10
98	8.4	4.8		2925	13	8	14	6	6	2	2
99	8.1	2.4	9	3000	11	8	16	4	1	11	9
100	7.5	4.8		3050	14	16	6	1	2	3	2
101	6.1			2000	10	18	28	2	1	6	5
102	4.9	10.2		1700	12	21	25	0	0	13	12
103	4.7	6		3175	12	7	24	0	0	24	17
104	4.6	1.7		3300	11	11	15	3	1	11	7
105	6.24	2.5	10	2200	10	7	9	4	1	4	3
106	4.5	7.2	5	1500	5	2	0	NA	NA	NA	NA
107	9.1	5.9	9	2200	9	11	25	9	2	14	9
108	133	2.3	8	2500	10	6	5	1	0	4	1
109	12.2	0.8	2	3000	10	1	0	NA	NA	NA	NA
110	9.5	1.7	6	3600	12	5	6	0	0	6	2
111	7.4	18.9		2250	10	14	32	1	0	13	9
112	7.7	1.1	7	3000	10	6	8	1	0	7	3
113	8.3	1.8	7	3000	10	9	10	0	1	4	4
114	8.1	4.7		2350	12	11	16	3	0	13	12
115	6.2	2.2	5	3000	10	3	6	0	0	6	1
116	6.7	2.9	7	2700	9	0	11	4	0	7	7
117	6	1.3	6	2100	7	4	5	1	0	4	3
118	10.2		4	2400	8	2	6	0	0	6	3
119	5.24			2475	11	18	15	0	0	8	2
120	4.9	1.2	6	3000	10	2	1	0	0	1	1
121	6.1	4.6	12	1800	8	14	20	1	2	17	7

2PN, two pronuclei; AFC, antral follicle count; AMH, anti-Müllerian hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone; MII, metaphase II; NA, not applicable. *Equivalent* (white background) and *dissimilar* (gray background) groups.

Chapter IV. Cumulus cell DNA damage linked to fertilization success in females with an ovulatory dysfunction phenotype[4]

Abstract

Intracytoplasmic sperm injection (ICSI) is a widely used technique in fertility centers. ICSI success depends on both nuclear and cytoplasmic oocyte maturation. Cumulus cells, which surround the oocytes, play a pivotal role in oocyte competence. However, the significance of DNA damage in cumulus cells as a marker of fertilization success remains largely unexplored. This study aims to investigate the relationship between DNA damage in cumulus cells of females undergoing ICSI, and oocyte competence, with a focus on *in vitro* fertilization (IVF) outcomes. We employed the alkaline comet assay to assess DNA damage levels (%TDNA) in cumulus cells and whole blood from 22 potentially fertile females and 35 infertile females, including 20 with an ovulatory disfunction phenotype. Our results revealed significant differences between the levels of %TDNA in cumulus cells and blood. Females with an ovulatory dysfunction phenotype exhibited higher levels of %TDNA in cumulus cells compared to potentially fertile females. Additionally, within the group of females with ovulatory dysfunction, a significant correlation was observed between %TDNA levels and the number of oocytes with two pronuclei. Our findings suggest that blood does not accurately reflect DNA damage in cumulus cells, which was correlated with the fertilization success in females with ovulatory dysfunction. High levels of %TDNA in cumulus cells were associated with a higher likelihood of successful fertilization. Moreover, our results imply that low levels of %TDNA may be linked to oocytes that fail to complete maturation and, consequently, do not fertilize (oocytes with zero pronuclei). Further research with larger cohorts is necessary to validate these findings and to explore potential applications in female fertility. However, our study provides evidence that DNA damage in cumulus cells may serve as a valuable biomarker for predicting fertilization success and oocyte competence.

Keywords: cumulus cells, DNA damage, comet assay, ovulatory dysfunction, fertilization success, two pronuclei

[4]Rodrigues *et al*., Cumulus cell DNA damage linked to fertilization success in females with an ovulatory dysfunction phenotype. Accepted for publication in Frontiers in Cell and Development Biology, October 24th, 2024.

1. Introduction

Infertility is a complex disease characterized by the inability to achieve a clinical pregnancy after 12 months or more of regular unprotected sexual activity [1–3]. This disease affects a significant number of reproductive-aged couples globally, leading to increased reliance on assisted reproductive technologies (ART) [4,5]. Intracytoplasmic sperm injection (ICSI) is a widely utilized technique in fertility centers, with a fertilization rate between 50% and 80% [6,7]. Nevertheless, the success rate of ICSI can vary depending on the underlying cause of infertility [8]. In clinical practice, the oocytes for ICSI are selected based on their nuclear maturation status, cytoplasmic structure and extracytoplasmic morphology [9,10]. The presence of the first polar body serves as a biomarker of nuclear maturity, indicating that the oocyte is at metaphase II (MII) [6,9,11–15]. The success of ICSI depends on both the nuclear and cytoplasmic maturity of the oocyte, which are essential for the development of pronuclei and the subsequent completion of fertilization [16,17]. It is crucial that the process of nuclear and cytoplasmic maturation are coordinated to establish optimal conditions for successful fertilization [18,19].

The cumulus-oocyte complex (COC), composed of an oocyte surrounded by cumulus cells, is essential for oocyte development and maturation. Cumulus cells provide a supportive microenvironment, facilitating the transfer of vital signals and nutrients via gap junctions. Additionally, these cells contribute significantly to glycolytic activity, serving as the primary energy source for oocyte maturation [20,21]. This bi-directional communication is essential for maintaining oocyte quality [20,22–24]. Cumulus cells play a pivotal role in COC expansion, driven by the secretion of hyaluronic acid, a process essential for meiotic maturation, ovulation, and fertilization. Furthermore, these cells support fertilization by releasing factors like prostaglandins, which enhance sperm motility and the acrosome reaction. Cumulus cells also contribute to early embryonic development, aiding in cleavage and blastocyst formation [20,21]. At the conclusion of oocyte maturation and COC expansion, gap junctions between the oocyte and cumulus cells gradually close, leading to a loss of cell-to-cell connections [20,25]. As a result, the cumulus cells, undergo apoptosis, characterized by DNA fragmentation [26,27]. The close proximity between cumulus cells and oocyte suggests their potential as a biomarker for predicting oocyte competence. While studies in male infertility have linked sperm DNA damage to reproductive failure and lower fertilization rates [28–30], research in female infertility regarding the use of cumulus cells as a biomarker for fertilization

success has yield inconsistent findings [24,31,32]. For instance, Raman et al. (2001) found a link between the comet tail moment and the percentage of fertilization following ICSI [33]. In contrast, Tola et al. (2018) demonstrated that DNA damage in cumulus cells is not associated with oocyte and embryo quality or ICSI success [34]. These contradictory results further underscore the importance of elucidating the relationship between cumulus cell DNA damage and oocyte competence.

The primary objective of this study was to investigate the correlation between DNA damage in cumulus cells and oocyte competence in the *in vitro* fertilization (IVF) outcomes. To complement this analysis, we investigated blood samples as a surrogate marker for DNA damage in cumulus cells, aiming to establish a reliable and accessible method for assessing DNA damage in a broader population. Cumulus cells and whole blood were collected from infertile females and those with male factor-related infertility undergoing ICSI. DNA damage was quantified using the alkaline comet assay, and comparisons were made between sets with different fertility status, to explore correlations between DNA damage levels, ovarian reserve markers and with IVF outcomes.

2. Material and Methods

2.1. Study design and population

This study investigated the outcomes of ICSI in non-smoking females (*n* = 57). Two sets were defined based on the cause of infertility: set 1 (potentially fertile), consisted of 22 females experiencing male factor-related infertility, obtained by convenience sampling; set 2, comprised 35 infertile females, arbitrarily selected using Random Sequence Generator (https://www.random.org/, accessed on 05 November 2022) from a larger cohort. This cohort included females diagnosed with various female factor infertility etiologies: ovulatory dysfunction (*n* = 83, 66.9%), endometriosis (*n* = 17, 13.7%), and oocyte factor (*n* = 12, 9.7%). Less frequent diagnoses (*n* = 12, 9.7%) included hypothyroidism, hyperprolactinemia, diminished ovarian reserve, adenomyosis, obesity, among others. Age-range were matched between the two groups to control for potential confounding effects of age on hormonal secretory patterns [35].

Participants were recruited between mid-January/2020, and February/2022 from the Centre for Medically Assisted Procreation of the Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN-ULSSA). Recruitment was temporarily paused between March and November/2020 due to COVID-19 pandemic. The study was approved by the Ethics

Committee of the Unidade Local de Saúde de Santo António (ULSSA) (process number 2020.119 /097-DEFI/099-CE). Prior to enrollment, all participants provided written informed consent after receiving detailed information about the study's objectives, procedures, potential risks, and benefits. Participants were given the opportunity to ask questions and ensure their understanding before signing the consent form. It was emphasized that participation would not affect their medical treatment, and that collected data would be used exclusively for research purposes. Consent was obtained during the initial phase of ovarian stimulation to align with the timing of sample collection.

2.2. Fertility-related outcomes and demographic variables

Baseline demographic data, including age and infertility etiology were collected from all participants. Ovarian reserve markers were assessed, encompassing day 3 follicle-stimulating hormone [FSH], anti-Müllerian hormone [AMH] within 6 months prior to ovarian stimulation, and antral follicle count [AFC] (Table 1). Additionally, IVF outcomes were captured, including ovarian stimulation response (total gonadotrophin dose, stimulation duration, follicle count on the trigger day, and oocyte retrieval), oocyte quality (mature MII oocytes), and fertilization success (pronuclear configuration: 0PN, 1PN, 2PN, 3PN) (Table 2).

2.3. Alkaline comet assay

2.3.1. Collection of whole blood and cumulus cells

EDTA K3 tubes (VACUETTE®, Greiner AG, Kremsmünster, Austria) were used to collect whole blood samples from each female on the day of follicular puncture. The samples were immediately stored at -70°C until further use. Cumulus cells were also obtained on the same day, following oocyte denudation using enzymatic (ICSI Cumulase® enzyme [Origio®, Måløv, Denmark]) and mechanical methods. Cells from all oocytes retrieved per female were pooled and cryopreserved in liquid nitrogen. All samples were anonymized to ensure confidentiality

2.3.2. Positive and negative controls

Negative and positive controls were established using a pool of blood samples obtained from five non-smoking female volunteers within a comparable age range as the study participants. The positive control was exposed to a concentration of 5 mM methyl methanesulfonate (Merck KGaA, Darmstadt, Germany) for a duration of one hour at 37°C. Following this treatment, the cells were cryopreserved in liquid nitrogen.

2.3.3. Assessment of DNA damage

The alkaline comet assay was performed, following minor modifications in the protocol based on the original methods described by Singh et al. (1988) [36], subsequently adapted by Abreu et al. (2017) and Collins et al. (2023) [37,38]. Briefly, thawed cumulus cell pellets at 37 °C, were washed with PBS (Sigma-Aldrich®), and resuspended in 0.8% low-melting point agarose (LMA) (Sigma-Aldrich®). Thawed whole blood samples on ice at room temperature were directly resuspended in 0.8% LMA. Two-gel slides were prepared for each sample by dropping 70 µl of cell suspension into each gel, on a frosted slide precoated with 1% normal melting point agarose (GRS Agarose LE, Grisp, Porto, Portugal) and covered with 20×20 coverslips. Slides were placed at 4ºC for 5 minutes to solidify. Coverslips were removed, and slides were immersed in a freshly prepared lysis solution (2.5 M NaCl, 0.1 M EDTA, 0.01 M Tris, 10 M NaOH, and 1 % Triton X-100, pH = 10) at 4°C for 90 minutes in the dark. Slides were placed in an electrophoresis tank with cold buffer (0.3 M NaOH and 0.01 M EDTA, pH = 13) at 4°C for 30 minutes, to allow DNA unwinding. Electrophoresis was carried out at 4 °C for 20 minutes at 25 V (1v/cm). Slides were washed with neutralizing buffer (PBS, Sigma-Aldrich®), at 4°C for 10 minutes, dehydrated/fixed with ethanol 96% (10 minutes) (Merck KGaA, Frankfurter, Darmstadt, Germany), and let to dry overnight. Dried slides were stained with SYBR™ Gold (Invitrogen™, Waltham, MA, USA) at room temperature for 30 minutes. Analyses were performed using Motic BA410E epi-fluorescence microscope (Motic®, Barcelona, Spain), with 400x magnification, and Comet Assay IV (Instem®, Staffordshire, UK) software. A total of 150 cells (75 per gel) were scored per sample type (cumulus and blood cells) for each female.

According to the Organization for Economic Co-operation and Development (OECD) guidelines for in vivo mammalian alkaline comet assay (test no. 489), the percentage of DNA in the comet tail (%TDNA) was used to evaluate DNA damage at cell level on a scale from 0 to 100% [39].

2.4. Statistical analysis

The minimum sample size (n = 33) was estimated based on Pearson's correlation coefficient, assuming an expected correlation of 0.5, a 95% confidence level, and a statistical power of 80% [40]. To compare fertility-related outcomes (total dose of gonadotrophins and stimulation duration), *t*-test were conducted. When the data failed normality and homoscedasticity, Mann-Whitney test were employed (FSH, AMH, AFC,

number of follicles on the trigger day, number oocytes retrieved, number of injected MII oocytes, number of 0PN, 1PN, 2PN, 3PN oocytes and %TDNA levels). Descriptive statistics for variables that did not follow a normal distribution were reported as median with interquartile range (quartile 1, quartile 3), while mean ± standard deviation (s.d.) was used for normally distributed data. A two-way analysis of variance (ANOVA) was employed to evaluate differences in DNA damage levels (log-%TDNA), considering fertility status and type of cells as independent variables. The study adhered to a fully randomized factorial design, ensuring independent and randomly selected datasets. To account for potential interactions between explanatory variables, a factorial ANOVA model was utilized. Following two-way ANOVA, Dunnett's multiple comparison test was conducted. Pearson correlation coefficient was used to assess the linear correlation (linear association) between log-%TDNA and fertility-related outcomes (FSH, log-AMH, AFC, total dose of gonadotrophins, stimulation duration number of follicles on the trigger day, number of oocytes retrieved, number of injected MII oocytes, square-root of the number of oocytes with 0PN and square-root of the number of oocytes with 2PN). All statistical analyses were performed using SigmaPlot version 14.0 (Systat Software® Inc., Chicago, IL, USA). A statistical significance level of 0.05 was used for all tests.

3. Results

3.1. Demographic and clinical characteristics

This study included 22 females with male-related infertility with a mean age of 33.6 ± 3.2 years, ranging from 28 to 39 years and 35 infertile females a mean age of 34.8 ± 3.5 years, ranging from 25 to 39 years. The primary causes of infertility were ovulatory dysfunction (*n* = 20, 57.1%), oocyte factor (*n* = 6, 17.1%), and endometriosis (*n* = 5, 14.3%). Additional causes, including hypothyroidism, hyperprolactinemia, diminished ovarian reserve, and adenomyosis were less prevalent (*n* = 4, 11.5%). Tables 1 and 2 summarize the demographic and clinical characteristics of the two sets. Consistent with previous reports, most participants exhibited an ovulatory dysfunction phenotype, characterized by irregular or absent menstruation and anovulation [41,42].

3.2. Infertile and potentially fertile females showed similar conventional fertility-related outcomes

Ovarian reserve markers, including FSH, AMH, and AFC, were assessed in both sets (Table 1). Despite matching for age, no significant differences were found in ovarian reserve markers between infertile and potentially fertile females (Table 1). Furthermore, IVF outcomes, including total dose of gonadotrophins, stimulation duration, number of follicles on the trigger day, number of oocytes retrieved and MII oocytes injected were comparable between the two sets. Subsequently, we evaluated the fertilization rates by categorizing the oocytes into non-fertilized (0PN), abnormally fertilized (1PN and 3PN), and normally fertilized (2PN). No statistically significant differences were found in the number of oocytes with 0PN, 1PN, 3PN as well as 2PN (Table 2). These findings suggest that infertile and potentially fertile females may exhibit similar ovarian reserve characteristics and IVF outcomes, highlighting the complexity of infertility and the potential for successful outcomes in both sets.

Table 1. Descriptive analysis of the ovarian reserve markers according to fertility status.

	Set 1		**Set 2**		
	n	**Median (Q1, Q3)**	***n***	**Median (Q1, Q3)**	***p*-value[a]**
Age (years)	22	33.5 (31.0, 35.5)	35	35.0 (32.0, 38.0)	0.103[b]
Ovarian reserve markers					
Day 3 FSH (mUI/ml)	22	7.8 (6.2, 9.6)	33	6.7 (5.1, 8.6)	0.087
AMH (ng/ml)	17	8.5 (5.5, 3.1)	33	7.8 (4.6, 15.0)	0.822
AFC	16	8.0 (6.0, 11.5)	23	7.0 (6.0, 9.0)	0.687

Values are presented as median (Quartile 1, Quartile 3); Set 1: Females with male factor-related infertility; Set 2: Infertile females; FSH: Follicle-stimulating hormone; AMH: Anti-Müllerian hormone; AFC: Antral follicle count; [a]*p*-value of Mann-Whitney test; [b]*p*-value of *t*-test.

Table 2. Descriptive analysis of the IVF outcomes according to fertility status.

	Set 1		Set 2		
	n	**Median (Q1, Q3)**	***n***	**Median (Q1, Q3)**	***p*-value[a]**
Response to ovarian stimulation					
Number of follicles on the trigger day	22	7.5 (4.8, 11.5)	35	6.0 (2.0, 11.0)	0.289
Number of oocytes retrieved	22	11.5 (8.8, 16.5)	35	10.0 (6.0, 16.0)	0.367
Oocyte maturation					
Number of injected MII oocytes	22	9.0 (4.0, 12.5)	35	7.0 (4.0, 11.0)	0.267
Fertilization success					
Number of oocytes with 0PN	22	2.0 (0, 4.3)	34	1.0 (1.0, 3.0)	0.427
Number of oocytes with 1PN	22	0 (0, 1.0)	34	0 (0, 0)	0.257
Number of oocytes with 2PN	22	5.0 (3.0, 9.0)	34	5.0 (2.0, 7.0)	0.506
Number of oocytes with 3PN	22	0 (0, 0.3)	34	0 (0, 0)	0.059
	n	**Mean ± s.d.**	***n***	**Mean ± s.d.**	***p*-value[b]**
Therapeutic regime					
Total dose of gonadotrophins (IU/ml)	22	2519.3 ± 766.9	35	2684.3 ± 706.9	0.205
Stimulation duration (days)	22	10.7 ± 1.6	35	10.5 ± 1.9	0.288

Values are presented as median (Quartile 1, Quartile 3) for non-parametric data and mean ± standard deviation (s.d.) for normally distributed data; Set 1: Females with male factor-related infertility; Set 2: Infertile females; 0PN: Zero pronuclei; 1PN: One pronuclei; 2PN: Two pronuclei; 3PN: Three pronuclei; [a]*p*-value of Mann-Whitney test; [b]*p*-value of *t*-test.

3.3. Distinct levels of DNA damage in blood and cumulus cells

To evaluate the extent of DNA damage in blood and cumulus cells, the comet assay was simultaneously conducted on control samples. Cells with no visible DNA in the tail (negative control, Figure 1A) served as a baseline, while those with DNA in the tail (positive control, Figure 1B) indicated DNA damage. A variation coefficient of 15% and 20%, was observed in the positive and negative controls, respectively, ensuring the assay's reproducibility [43].

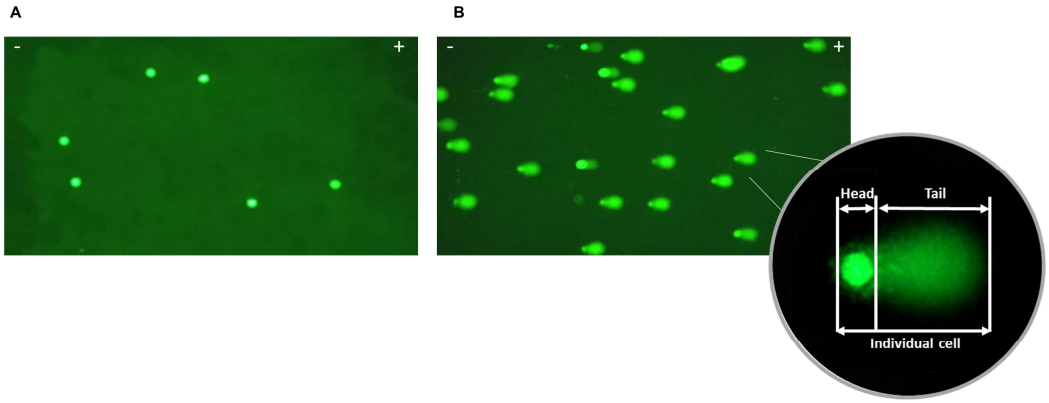


Figure 1. Visualization of DNA damage in cumulus cells after alkaline comet assay. (A) Negative control: intact nuclei without DNA damage (%TDNA = 1.16); (B) Positive control: exposure to MMS resulted in significant high DNA damage and insignificant head (%TDNA = 49.66). Tail intensity represents the percentage of DNA fragments the comet's tail, indicating the extent of DNA damage; - cathode; + anode.

An association analysis of DNA damage levels (%TDNA) in blood and cumulus cells from the same individuals revealed no significant correlation (Pearson correlation: r = -0.095, n = 57, p = 0.481). To investigate the influence of tissue type and fertility status on DNA damage a two-way ANOVA was performed. Log-%TDNA values exhibited normality (Shapiro-Wilk test: W= 0.984; p = 0.189) and equal variance (Brown-Forsythe test: p = 0.182). The ANOVA results demonstrated a significant effect of tissue type on log-%TDNA ($F_{(1, 110)}$ = 93.286, p < 0.001) (Figure 2A), but no interaction between tissue type and fertility status ($F_{(1, 110)}$ = 3.832, p = 0.053), nor a significant effect of fertility status ($F_{(1, 55)}$ = 0.002, p = 0.964) (Figure 2B). These findings suggest that while tissue type significantly influences DNA damage, fertility status does not. Similar results were observed when analyzing %TDNA levels according to the fertility status (Figure 3). Supplementary Table 1 provides detailed statistical analyses for each sample.

A Mann-Whitney test revealed no statistically significant differences in %TDNA levels between cumulus cells from infertile and potentially fertile females (U = 274.5; n_1 = 22; n_2 = 35; p = 0.07). Given this lack of statistical significance, we hypothesized that heterogeneity of infertility causes might be influencing these outcomes. When analyzing only infertile females with ovulatory dysfunction (20 out of 35 cases) we observed significant higher %TDNA levels in this group compared to potentially fertile females (Mann-Whitney test: U = 135.5; n_1 = 20; n_2 = 22; p = 0.034), as depicted in Figure 4. Importantly, no significant differences were observed in fertility-related outcomes between these two sets (p > 0.05 all variables).

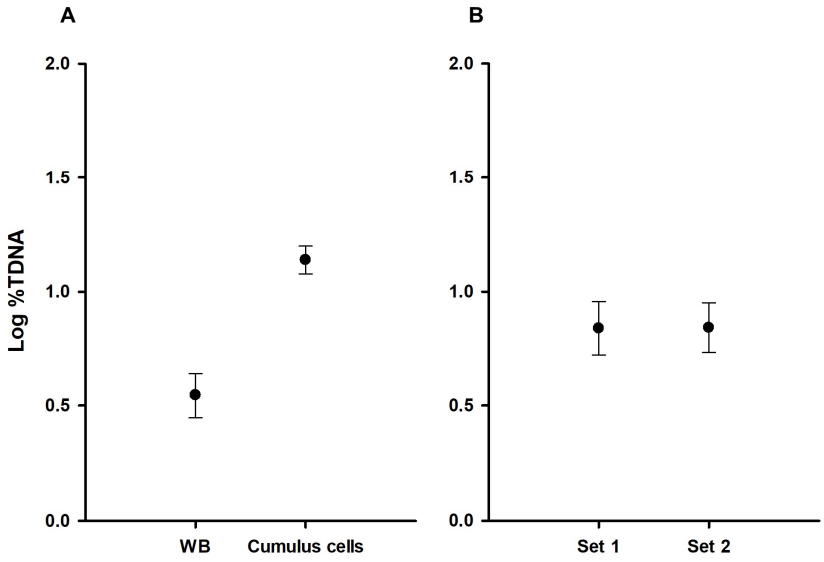


Figure 2. Comparative analysis of DNA damage levels according to sets and tissues. (A) Comparison of DNA damage levels in whole blood (WB) and cumulus cells; (B) Comparison between set 1 (females with male factor-related infertility) and set 2 (infertile females). %TDNA – levels of DNA damage.

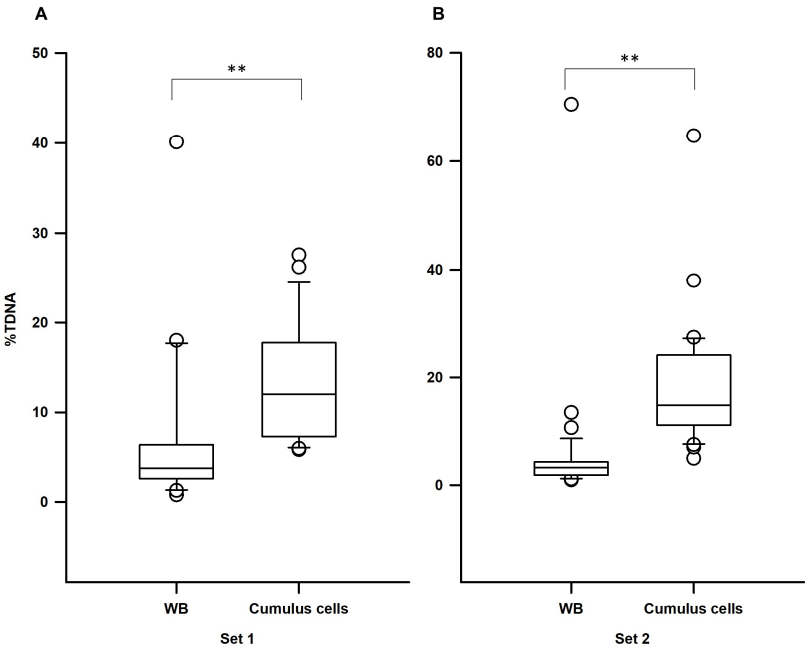


Figure 3. Tissue-specific differences in DNA damage levels in each set. (A) Set 1 (females with male factor-related infertility): cumulus cells exhibited significantly higher DNA damage levels compared to whole blood (WB); (B) Set 2 (infertile females): similarly, to set 1, cumulus cells displayed significantly elevated DNA damage levels compared to WB. %TDNA – DNA damage levels; $^{**}p < 0.001$, calculated using the Mann-Whitney test.

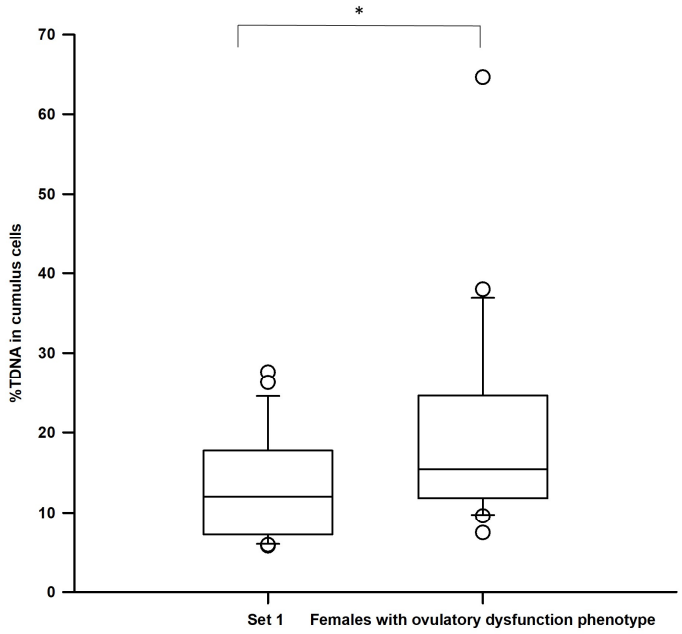


Figure 4. **Comparison of DNA damage levels between set 1 (females with male factor-related infertility) and infertile females with ovulatory dysfunction.** %TDNA – DNA damage levels; *p < 0.05, calculated using the Mann-Whitney test.

3.4. Impact of cumulus cell DNA damage in females with an ovulatory dysfunction phenotype

To investigate the association between DNA damage in cumulus cells and ovarian function and competence, we examined the correlation between log-%TDNA levels and ovarian reserve markers (FSH, log-AMH and AFC) and IVF outcomes. Pearson correlation analysis revealed no significant correlations between log-%TDNA and the ovarian reserve markers (Table 3). Additionally, no correlations were observed between log-%DNA and IVF outcomes, including total dose of gonadotrophins, stimulation duration, number of follicles on the trigger day, number of oocytes retrieved, number of injected MII oocytes and square-root of number of 0PN oocytes) (Table 4). A significant positive correlation was found between log-%TDNA and the square-root of the number of 2PN oocytes (Pearson correlation: r = 0.508, n = 19, p = 0.026), indicating an association with fertilization success (Table 4, Figure 5A). In addition, infertile females with ovulatory dysfunction and a high number of 0PN oocytes (≥ 8) exhibited lower levels of %TDNA levels in cumulus cells compared to those with fewer 0PN oocytes (Figure 5B, Supplementary Table 2). Although this result did not reach statistical significance

(Mann-Whitney test: U = 14.0; n_1 = 3; n_2 = 16; p = 0.288), it suggests a potential association between reduced %TDNA levels and oocyte competence, particularly in cases with a high number of 0PN oocytes.

Table 3. Summary of correlation analysis with ovarian reserve markers in females with ovulatory dysfunction.

	Correlation coefficients with levels of DNA damage in cumulus cells**	*p*-value	*n*
Markers of ovarian reserve			
Day 3 FSH (mUI/ml)	-0.247	0.294	20
AMH (ng/ml)**	-0.112	0.639	20
AFC	0.153	0.674	10

**Log-transformed data; FSH: Follicle-stimulating hormone; AMH: Anti-Müllerian hormone; AFC: Antral follicle count.

Table 4. Summary of correlation analysis with IVF outcomes in females with ovulatory dysfunction.

	Correlation coefficients with levels of DNA damage in cumulus cells**	*p*-value	*n*
Therapeutic regime			
Total dose of gonadotrophins (IU/ml)	0.225	0.339	20
Stimulation days	0.204	0.388	20
Response to ovarian stimulation			
Number of follicles on the trigger day	-0.053	0.825	20
Number of oocytes retrieved	0.054	0.821	20
Oocyte maturation			
Number of injected MII oocytes	0.323	0.165	20
Fertilization success			
Number of oocytes with 0PN++	-0.080	0.746	19
Number of oocytes with 2PN++	0.508	0.026S	19

**Log-transformed data; ++Square root-transformed data; SSignificant correlation; MII: Metaphase II; 0PN: Zero pronuclei; 2PN: Two pronuclei.

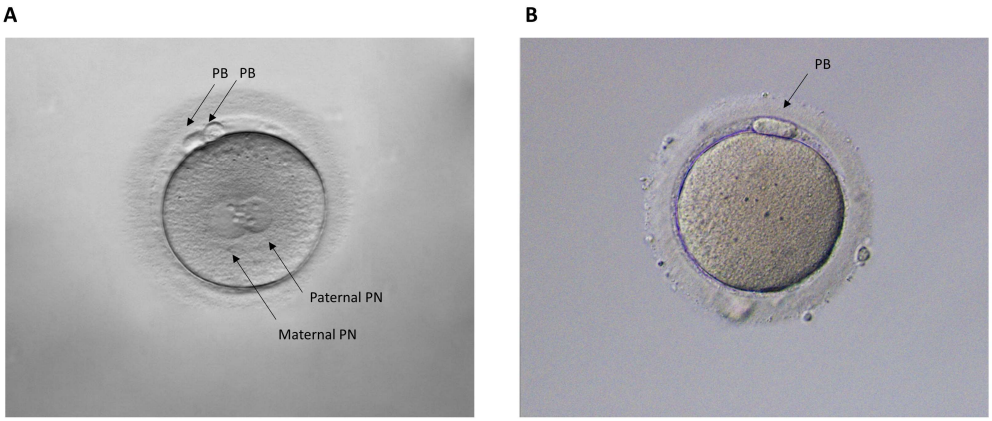


Figure 5. Representative human MII oocytes. (A) Normal fertilization: oocyte with two pronuclei (2PN, maternal and paternal) and two polar bodies; (B) Unfertilized oocyte: zero pronuclei (0PN) and a single polar body (PB).

4. Discussion

This study aimed to investigate the relationship between DNA damage in cumulus cells of females undergoing ICSI and oocyte competence, with a focus on IVF outcomes.

Blood samples have been considered a valuable tool for evaluating global DNA damage, for this we evaluated whether this tissue could mirror the effects of DNA damage in cumulus cells. Our findings suggest that it may not be a suitable sentinel for cumulus cell DNA damage, posing a limitation for future studies due to the inherent challenges associated in obtaining cumulus cells.

This study comprised a set of infertile females with diverse etiologies of infertility, with a predominance of ovulatory dysfunction phenotypes, consistent with previous findings [41]. Our results revealed a significant increase in the levels of %TDNA in cumulus cells of females with ovulatory dysfunction compared to potentially fertile females. Although the sample size was relatively small, the robustness of our conclusions was supported by strong statistical power. Notably, we observed a significant positive correlation between the levels of %TDNA in cumulus cells and the number of 2PN oocytes, suggesting that DNA damage level in cumulus cells may indeed be relevant predictor of fertilization success and a potential indicator of oocyte competence. In alignment with our hypothesis, a sample with higher levels of %TDNA (27%) exhibited a high fertilization success rate, yielding twelve 2PN oocytes from thirteen injected MII oocytes (sample 24, Supplementary Tables 1 and 2). Conversely, a sample with lower %TDNA levels (7%)

demonstrated a low rate of fertilization success, resulting in only one 2PN oocyte from ten injected MII oocytes (sample 37, Supplementary Tables 1 and 2). The development of pronuclei and the subsequent completion of fertilization depend on the nuclear and cytoplasmatic maturation, i.e. oocyte competence [13,18,44,45]. Interestingly, in sample 37, despite the injection of ten MII oocytes nine did not fertilize (nine 0PN [zero pronuclei] oocytes), Supplementary Table 2). While other factors may have contributed, it is plausible that the 0PN oocytes were immature, likely due to insufficient cytoplasmic maturation, leading to no fertilization. The preserved DNA integrity of cumulus cells, as indicated by lower levels of % TDNA, could potentially serve as a reliable biomarker of oocyte competence, addressing a significant knowledge gap in reproductive biology. Our findings align with those of Raman et al. (2001) who demonstrated a correlation between the comet tail moment and the percentage of fertilization following ICSI [33]. Furthermore, Lourenço et al. (2014) observed elevated caspase activity, one of the apoptotic markers they analyzed, in cumulus cells from fertilized oocytes compared to unfertilized ones, suggesting a potential link to oocyte quality and fertilization ability [46]. However, the literature presents conflicting evidence regarding this association [47,48]. Høst et al. (2000) and Lee et al. (2001) reported results suggesting that high levels of cumulus cell apoptosis may compromise fertilization [31,49]. Abu-Hassan et al. (2006) found no clear association between cumulus cell apoptosis levels and oocyte fertilization outcomes [32]. Similarly, Tola et al. (2018) demonstrated a lack of correlation between DNA integrity of cumulus cells and oocyte and embryo quality or ICSI success [32].

The discrepancies in these findings may be attributed to variations in study design, such as evaluation of DNA damage in pooled or individually isolated cumulus cells, sample size, and the methodologies employed to assess apoptosis, including the incidence of apoptosis, apoptotic markers, percentage of viable cells, or DNA fragmentation. Additionally, the specific fertility parameters evaluated in each study may contribute to the observed differences. The detachment and apoptosis of cumulus cells during COC expansion, as documented by Szołtys et al. (2000) and Luciano et al. (2004), underscore the critical nature of this cellular interaction [25,50]. While the precise mechanisms governing the relationship between DNA damage and apoptosis during oocyte maturation remain enigmatic, further investigations with larger sample sizes and additional biological endpoints are essential to elucidate these pathways and explore potential therapeutic applications for addressing female infertility.

5. Conclusion

Overall, these findings suggest that DNA damage in cumulus cells may serve as a predictive marker of fertilization success. Furthermore, the study provides insights into the association between low levels of total DNA damage and the absence of cytoplasmic maturation. These results highlight the potential significance of cumulus cell DNA damage as an indicator of oocyte development and maturation. A more comprehensive understanding of the intricate relationship between cumulus cells and oocyte development could lead to the identification of valuable biomarkers for assessing oocyte quality, maturation, fertilization, embryonic development, and cumulus cell function, ultimately contributing to the optimization of routine IVF procedures.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contributions

Paula Jorge, António J A Nogueira, and Bárbara Rodrigues conceived and designed the study. António J A Nogueira and Bárbara Rodrigues performed the statistical analysis. Paula Jorge critically reviewed the manuscript and is responsible for raising funds for this study and its publication. Bárbara Rodrigues and Vanessa Sousa performed the laboratory work and drafted the manuscript. Solange Costa, Filipa Esteves, Joana Pires and Isabel Gaivão helped with the optimization of the comet assay and comet analysis. Emídio Vale-Fernades was responsible for recruiting participants and clinical data. Daniela Sousa, Raquel Brandão, and Carla Leal collected biological material. Solange Costa and João Paulo Teixeira provided critical feedback and contributed to the advancement of the study. All authors have discussed the results and reviewed the manuscript. All authors have approved the submitted version and any substantially modified version that involves the author's contribution to the study.

Funding

This work was supported by national funds: FCT/ESF (Foundation for Science and Technology /European Social Fund) – Project Reference SFRH/BD/136398/2018 and COVID/BD/153204/2023 to B. Rodrigues and FCT/FSE – Project Reference EXPL/BIA-REP/0423/2021—X-EPIFERTILITY. UMIB - Unit for Multidisciplinary Research in Biomedicine is funded by the FCT Portugal (grant numbers UIDB/00215/2020, and

UIDP/00215/2020), and ITR - Laboratory for Integrative and Translational Research in Population Health (LA/P/0064/2020). DEFI (Departamento de Ensino, Formação e Investigação) – Reference 2015-DEFI/145/12. Thanks are due to FCT/Ministry of Science, Technology, and Higher Education (MCTES) for the financial support to CESAM (UIDP/50017/2020+UIDB/50017/2020), through national funds. F. Esteves work is supported by FCT and ESF through the PhD individual grant UI/BD/150783/2020.

Acknowledgements

We gratefully acknowledged all the volunteers for participating in this study and all the healthcare professionals involved. A special thanks to Rosário Santos, Ph.D., and all the collaborators at the Molecular Genetics Laboratory, Laboratory Genetics Service, Genetics and Pathology Clinic, Santo António Local Health Unit (ULSSA), for providing an environment conducive to the progress of my project, and for the continuous support, availability, and insightful discussions about the research. We would also like to acknowledge the Portuguese Foundation for Science and Technology (FCT) and Ministry of Science, Technology and Higher Education (MCTES) for the financial support to Ph.D. project (SFRH/BD/136398/2018 & COVID/BD/153204/2023) and CESAM (UIDP/50017/2020, UIDB/50017/2020 & LA/P/0094/2020). Finally, we would like to acknowledge Sofia Nunes, Ph.D. (Scientific ToolBox Consulting, Lisbon, Portugal) for providing technical editing

References

1. Vander Borght M, Wyns C. Fertility and infertility: Definition and epidemiology. Clin Biochem. 2018;62:2–10.

2. Gurunath S, Pandian Z, Anderson RA, Bhattacharya S. Defining infertility-a systematic review of prevalence studies. Hum Reprod Update. 2011;17:575–88.

3. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Fertil Steril. 2017;108:393–406.

4. Ombelet W. WHO fact sheet on infertility gives hope to millions of infertile couples worldwide. Facts Views Vis Obgyn. 2020;12:249–51.

5. Nelson SM, Telfer EE, Anderson RA. The ageing ovary and uterus: New biological insights. Hum Reprod Update. 2013;19:67–83.

6. Rubino P, Viganò P, Luddi A, Piomboni P. The ICSI procedure from past to future: A systematic review of the more controversial aspects. Hum Reprod Update. 2016;22:194–227.

7. Nerl. Q V., Lee B, Rosenwalks Z, Machaca K, Palermo GD. Understanding fertilization through intracytoplasmic sperm injection (ICSI). Cell Calcium. 2013;55:24–37.

8. Song J, Liao T, Fu K, Xu J. ICSI Does Not Improve Live Birth Rates but Yields Higher Cancellation Rates Than Conventional IVF in Unexplained Infertility. Front Med. 2021;7:1–7.

9. Rienzi L, Ubaldi F, Iacobelli M, Minasi MG, Romano S, Greco E. Meiotic spindle visualization in living human oocytes. Reprod Biomed Online. 2005;10:192–8.

10. Lemseffer Y, Terret ME, Campillo C, Labrune E. Methods for Assessing Oocyte Quality: A Review of Literature. Biomedicines. 2022;10:4–11.

11. Sun B, Yeh J. Identifying fertilization-ready metaphase II stage oocytes beyond the microscope: a proposed molecular path forward. F S Rev. 2021;2:302–16.

12. Holubcová Z, Kyjovská D, Martonová M, Páralová D, Klenková T, Otevřel P, et al. Egg maturity assessment prior to ICSI prevents premature fertilization of late-maturing oocytes. J Assist Reprod Genet. 2019;36:445–52.

13. Trebichalská Z, Kyjovská D, Kloudová S, Otevřel P, Hampl A, Holubcová Z. Cytoplasmic maturation in human oocytes: An ultrastructural study. Biol Reprod. 2021;104:106–16.

14. VEECK LL. Extracorporeal Maturation: Norfolk, 1984. Ann N Y Acad Sci. 1985;442:357–67.

15. VEECK LL. Oocyte Assessment and Biological Performance. Ann N Y Acad Sci. 1988;541:259–74.

16. Mobarak H, Heidarpour M, Tsai PSJ, Rezabakhsh A, Rahbarghazi R, Nouri M, et al. Autologous mitochondrial microinjection; A strategy to improve the oocyte quality and subsequent reproductive outcome during aging. Cell Biosci. 2019;9:1–15.

17. Eppig JJ, O'Brien M, Wigglesworth K. Mammalian oocyte growth and development in vitro. Mol Reprod Dev. 1996;44:260–73.

18. Rienzi L, Balaban B, Ebner T, Mandelbaum J. The oocyte. Hum Reprod. 2012;27:2–22.

19. Swain JE, Pool TB. ART failure: Oocyte contributions to unsuccessful fertilization. Hum Reprod Update. 2008;14:431–46.

20. Turathum B, Gao EM, Chian RC. The function of cumulus cells in oocyte growth and maturation and in subsequent ovulation and fertilization. Cells. 2021;10:1–18.

21. Martinez CA, Rizos D, Rodriguez-Martinez H, Funahashi H. Oocyte-cumulus cells crosstalk: New comparative insights. Theriogenology. 2023;205:87–93.

22. Roberts R, Franks S, Hardy K. Culture environment modulates maturation and metabolism of human oocytes. Hum Reprod. 2002;17:2950–6.

23. Barcena P, López-Fernández C, García-Ochoa C, Obradors A, Vernaeve V, Gosálvez J, et al. Detection of DNA damage in cumulus cells using a chromatin dispersion assay. Syst Biol Reprod Med. 2015;61:277–85.

24. Baratas A, Gosálvez J, de la Casa M, Camacho S, Dorado-Silva M, Johnston SD, et al. Cumulus Cell DNA Damage as an Index of Human Oocyte Competence. Reprod Sci. 2022;29:3194–200.

25. Luciano AM, Modina S, Vassena R, Milanesi E, Lauria A, Gandolfi F. Role of Intracellular Cyclic Adenosine 3′,5′-Monophosphate Concentration and Oocyte-Cumulus Cells Communications on the Acquisition of the Developmental Competence during In Vitro Maturation of Bovine Oocyte. Biol Reprod. 2004;70:465–72.

26. Di Giacomo M, Camaioni A, Klinger FG, Bonfiglio R, Salustri A. Cyclic AMP-elevating agents promote cumulus cell survival and hyaluronan matrix stability, thereby prolonging the time of mouse oocyte fertilizability. J Biol Chem. 2016;291:3821–36.

27. Majtnerová P, Roušar T. An overview of apoptosis assays detecting DNA fragmentation. Mol Biol Rep. 2018;45:1469–78.

28. Agarwal A, Majzoub A, Baskaran S, Selvam MKP, Cho CL, Henkel R, et al. Sperm DNA fragmentation: A new guideline for clinicians. World J Mens Health. 2020;38:412–71.

29. Haddock L, Gordon S, Lewis SEM, Larsen P, Shehata A, Shehata H. Sperm DNA fragmentation is a novel biomarker for early pregnancy loss. Reprod Biomed Online. 2021;42:175–84.

30. Simon L, Emery B, Carrell DT. Sperm DNA fragmentation: Consequences for reproduction. Adv Exp Med Biol. 2019;1166:87–105.

31. Høst E, Mikkelsen AL, Lindenberg S, Smidt-Jensen S. Apoptosis in human cumulus cells in relation to maturation stage and cleavage of the corresponding oocyte. Acta Obstet Gynecol Scand. 2000;79:936–40.

32. Abu-Hassan D, Koester F, Shoepper B, Schultze-Mosgau A, Asimakopoulos B, Diedrich K, et al. Comet assay of cumulus cells and spermatozoa DNA status, and the relationship to oocyte fertilization and embryo quality following ICSI. Reprod Biomed Online. 2006;12:447–52.

33. Raman RS, Chan PJ, Corselli JU, Patton WC, Jacobson JD, Chan SR, et al. Comet assay of cumulus cell DNA status and the relationship to oocyte fertilization via intracytoplasmic sperm injection. Hum Reprod. 2001;16:831–5.

34. Tola EN, Koşar PA, Karatopuk DU, Sancer O, Oral B. Effect of DNA damage of cumulus oophorus cells and lymphocytes analyzed by alkaline comet assay on oocyte quality and intracytoplasmic sperm injection success among patients with polycystic ovary syndrome. J Obstet Gynaecol Res. 2018;45:609–18.

35. van den Beld AW, Kaufman JM, Zillikens MC, Lamberts SWJ, Egan JM, van der Lely AJ. The physiology of endocrine systems with ageing. Lancet Diabetes Endocrinol. 2018;6:647–58.

36. Singh NP, McCoy MT, Tice RR, Schneider EL. A simple technique for quantitation of low levels of DNA damage in individual cells. Exp Cell Res. 1988;175:184–91.

37. Abreu A, Costa C, Pinho e Silva S, Morais S, do Carmo Pereira M, Fernandes A, et al. Wood smoke exposure of Portuguese wildland firefighters: DNA and oxidative damage evaluation. J Toxicol Environ Heal - Part A Curr Issues. 2017;80:596–604.

38. Collins A, Gajski G, Unit M, Vodenkova S, Republic C, Republic C, et al. Measuring DNA damage with the comet assay : a compendium of protocols. Nat Protoc. 2023;18:929–89.

39. OECD/OCDE. Test No. 489.OECD GUIDELINE FOR THE TESTING OF CHEMICALS: In Vivo Mammalian Alkaline Comet Assay. Adopt. 2016. 2016.

40. Bonett DG, Wright TA. Sample size requirements for estimating Pearson, Kendall and Spearman correlations. Psychometrika. 2000;65:23–8.

41. Carson SA, Kallen AN. Diagnosis and Management of Infertility: A Review. JAMA - J Am Med Assoc. 2021;326:65–76.

42. Munro MG, Balen AH, Cho S, Critchley HOD, Díaz I, Ferriani R, et al. The FIGO Ovulatory Disorders Classification System. Russ J Hum Reprod. 2023;29:116–36.

43. Collins AR, El Yamani N, Lorenzo Y, Shaposhnikov S, Brunborg G, Azqueta A. Controlling variation in the comet assay. Front Genet. 2014;5:1–6.

44. Coticchio G, Dal Canto M, Renzini MM, Guglielmo MC, Brambillasca F, Turchi D, et al. Oocyte maturation: Gamete-somatic cells interactions, meiotic resumption, cytoskeletal dynamics and cytoplasmic reorganization. Hum Reprod Update. 2014;21:427–54.

45. Liao Z, Li Y, Li C, Bian X, Sun Q. Nuclear transfer improves the developmental potential of embryos derived from cytoplasmic deficient oocytes. iScience. 2023;26:107299.

46. Lourenço B, Sousa AP, Almeida-Santos T, Ramalho-Santos J. Relation of cumulus cell status with single oocyte maturity, fertilization capability and patient age. J Reprod Infertil. 2014;15:15–21.

47. Corn CM, Hauser-Kronberger C, Moser M, Tews G, Ebner T. Predictive value of cumulus cell apoptosis with regard to blastocyst development of corresponding gametes. Fertil Steril. 2005;84:627–33.

48. Høst E, Gabrielsen A, Lindenberg S, Smidt-Jensen S. Apoptosis in human cumulus cells in relation to zona pellucida thickness variation, maturation stage, and cleavage of the corresponding oocyte after intracytoplasmic sperm injection. Fertil Steril. 2002;77:511–5.

49. Lee KS, Joo BS, Na YJ, Yoon MS, Choi OH, Kim WW. Cumulus cells apoptosis as an indicator to predict the quality of oocytes and the outcome of IVF-ET. J Assist Reprod Genet. 2001;18:490–8.

50. Szołtys M, Tabarowski Z, Pawlik A. Apoptosis of postovulatory cumulus granulosa cells of the rat. Anat Embryol (Berl). 2000;202:523–9.

Supporting information

Supplementary material

1. Supplementary Tables

Supplementary Table 1. Levels of DNA damage (%TDNA) in whole blood and cumulus cells of females from sets 1 (females with male factor-related infertility; **bold samples**) and 2 (infertile females; underlined samples).

Sample	Levels of %TDNA in whole blood	Levels of %TDNA in cumulus cells
1	8.2	7.4
2	2.1	18.3
3	5.3	6.6
4	1.5	9.6
5	3.4	10.8
6	7.2	5.8
7	2.6	6.5
8	18	26.2
9	4.3	12
10	6	10.3
11	4	6.3
12	16.9	13.3
13	4.5	27.6
14	40.1	5.9
15	3	17.6
16	0.8	12.7
17	2.5	12
18	2.6	19.6
19	2.7	11.1
20	1.3	12
21	4	20.5
22	3.2	13.5
23	3.3	11.9
24	3.3	27.1
25	1.1	16
26	3.5	12.3
27	1.4	11.3
28	3.5	38
29	6.4	22.1
30	1.2	11.8
31	5.9	7.6
32	1.9	8.8
33	70.6	14.8
34	2.6	11
35	3	15
36	2.4	18.9
37	1.6	7.4
38	4.9	27.4
39	1.1	9.6
40	1.9	17.3

41	3.2	7
42	1.6	64.7
43	13.5	14.8
44	10.5	22.1
45	3.3	12.7
46	1	24.9
47	3.5	4.9
48	3.9	9.5
49	2.7	11.2
50	1.8	24.1
51	4.4	19
52	7.3	21.9
53	4.2	13.6
54	1.8	25.4
55	2.8	25
56	3.6	7.7
57	2.5	24.3

Supplementary Table 2. Fertility-related outcomes of females from sets 1 (females with male factor-related infertility; **bold samples**) and 2 (infertile females; underlined samples).

Sample	Age	Day 3 FSH	LH	AMH	AFC	Total dose of gonadotrophins (IU/ml)	Stimulation duration (days)	Nº of follicles on the trigger day	Nº of oocytes retrieved	Nº of injected MII oocytes	Nº of oocytes with 0PN	Nº of oocytes with 1PN	Nº of oocytes with 2PN	Nº of oocytes with 3PN
1	34	6.3	3.6	2.2	14	1800	8	11	14	11	4	0	6	0
2	38	10.5	9.1	1.3	-	3900	13	13	14	14	4	1	8	0
3	29	4.9	3.6	10	9	2400	12	14	33	3	15	1	0	13
4	39	6.9	7.3	1.4	-	2250	9	4	5	4	0	0	4	0
5	32	8.8	9.2	2	5	1950	9	3	7	4	3	0	1	0
6	35	12.1	8.5	2.7	5	3300	11	7	11	11	2	0	9	0
7	35	8.8	4	1.3	10	2175	10	6	9	5	2	0	3	0
8	35	10.4	6.9	4.4	10	2200	11	9	16	15	5	1	7	0
9	33	5.7	1.6	3.8	25	1350	9	15	11	10	5	1	4	0
10	33	7.3	4.1	2.3	-	3600	12	1	9	4	1	0	3	0
11	35	8	10.5	2.7	-	1500	10	2	8	7	2	0	3	1
12	39	10	5.1	1.1	4	3900	13	5	4	3	0	0	1	2
13	30	6.6	8.9	5.8	-	2500	13	10	37	18	0	0	14	0
14	38	9.4	27.5	2.3	10	2800	13	18	8	4	0	0	4	0
15	31	13.4	6	-	2	3750	13	6	9	7	0	1	6	0
16	30	8.7	6	2.7	12	2250	10	1	13	12	5	0	6	0
17	31	7.6	9.8	2.8	6	2000	10	8	11	6	2	0	3	1
18	34	8.1	5	-	7	2025	9	6	12	11	7	0	4	0
19	37	5.2	9.8	-	5	2250	10	11	24	20	3	1	13	1
20	32	6.8	5.3	-	6	3300	11	7	12	8	0	0	6	0
21	28	7.4	8.3	-	16	2025	9	9	21	18	3	0	15	0
22	31	5.3	4.8	6.8	-	2200	11	17	18	12	0	0	12	0
23	28	4.4	13.1	10.9	-	2025	9	13	25	10	1	0	8	0
24	34	4.9	14.1	10.2	-	1700	12	21	25	13	1	0	12	0
25	39	4.1	10.4	3.9	-	1575	7	12	25	14	8	0	5	0
26	38	7.4	17.5	18.9	-	2250	10	14	32	13	3	0	9	0
27	32	7	14.6	7	-	1350	9	11	23	11	2	0	7	0
28	39	4.8	8.7	1	8	3600	12	5	9	8	1	0	5	0
29	34	4.9	8.3	1.2	6	3000	10	2	1	1	0	0	1	0
30	36	9.5	15.8	1.7	6	3600	12	5	6	6	3	0	2	0

31	32	6.1	8.4	6.9	-	2475	11	0	6	4	1	1	2	0
32	37	5.2	6.6	2.4	12	3150	14	11	17	17	1	0	14	0
33	31	5.5	5.7	2.6	8	3300	12	2	5	5	1	0	4	0
34	38	6.4	6.7	5.6	-	1650	11	0	16	6	0	0	6	0
35	33	8.8	8.7	2.7	13	2250	10	0	7	6	3	0	3	0
36	35	6.2	6.1	2.5	10	2200	10	7	9	4	1	0	3	0
37	25	9	8.8	2.7	-	2025	9	12	11	10	9	0	1	0
38	30	8.1	7.7	2.4	9	3000	11	8	16	11	2	0	9	0
39	36	9.1	8.6	0.9	5	4200	14	5	10	9	1	0	5	0
40	32	8.4	7.6	4.8	-	2925	13	8	14	2	0	0	2	0
41	39	5.7	5.1	3.1	8	3000	10	11	8	6	1	0	4	0
42	34	4.7	4.2	6	-	3175	12	7	24	24	2	1	17	0
43	35	4.6	4	1.7	-	3300	11	11	15	11	2	0	7	0
44	31	7.1	6.1	1.3	-	2250	10	5	12	11	4	1	6	0
45	30	6.1	5	4.6	12	1800	8	14	20	17	9	0	7	0
46	38	2.7	2.2	-	5	3000	10	1	13	11	3	0	7	0
47	36	6.8	5.4	2.5	7	3000	10	3	6	5	1	1	3	0
48	39	8.3	5.9	1.8	7	3000	10	9	10	4	0	0	4	0
49	39	6.2	4.2	-	6	4200	14	23	10	7	0	0	5	1
50	37	11	7.2	0.6	8	2100	7	1	1	1	0	0	1	0
51	36	10.5	6.8	1	6	2100	7	0	1	0	NA	NA	NA	NA
52	38	7.6	4.9	2.3	10	2850	10	4	1	1	1	0	0	0
53	38	6.7	3.7	1.1	7	3300	11	5	4	3	1	0	2	0
54	34	11.6	3.9	0.7	6	3000	10	6	7	5	3	0	2	0
55	37	133	6.3	2.3	8	2500	10	6	5	4	2	0	1	1
56	34	-	-	1.6	3	2700	12	7	10	7	0	1	6	0
57	33	-	-	2.8	4	2400	8	0	13	7	2	0	5	0

FSH: Follicle-Stimulating Hormone (mIU/ml); LH: Luteinizing Hormone (mIU/ml); AMH: Anti-Müllerian Hormone (ng/ml); AFC: Antral Follicle Count; MII: Metaphase II; Nº: Number; NA: Not Applicable; 0PN: Zero pronuclei; 1PN: One pronuclei; 2PN: Two Pronuclei; 3PN: Three pronuclei; -: missing.

Chapter V. Use of the *FMR1* gene methylation status to assess the X-Chromosome Inactivation pattern: a stepwise analysis[5]

Abstract

X-chromosome inactivation (XCI) is a developmental process to compensate the imbalance in the dosage of X-chromosomal genes in females. A skewing of the XCI pattern may suggest a carrier status for an X-linked disease or explain the presence of a severe phenotype. In these cases, it is important to determine the XCI pattern, conventionally using the gold standard Human Andro-gen-Receptor Assay (HUMARA), based on the analysis of the methylation status at a polymorphic CAG region in the first exon of the human androgen receptor gene (*AR*). The aim of this study was to evaluate whether the methylation status of the fragile mental retardation protein translational regulator gene (*FMR1*) can provide an XCI pattern similar to that obtained by HUMARA. A set of 48 female carriers of *FMR1* gene normal-sized alleles was examined using two assays: HUMARA and a *FMR1* methylation PCR (mPCR). Ranges were defined to establish the XCI pattern using the methylation pattern of the *FMR1* gene by mPCR. Overall, a 77% concordance of the XCI patterns was obtained between the two assays, which led us to propose a set of key points and a stepwise analysis towards obtaining an accurate result for the XCI pattern and to minimize the underlying pitfalls.

Keywords: X-chromosome inactivation pattern; HUMARA; *FMR1* methylation status; methylation PCR

[5]Rodrigues *et al*., Use of the *FMR1* Gene Methylation Status to Assess the X-Chromosome Inactivation Pattern: A Stepwise Analysis. Genes (Basel). 2022 Feb 25;13(3):419.

1. Introduction

Females are usually less susceptible to recessive pathogenic variants in X-chromosome genes, being asymptomatic with a complete absence of clinical features or showing a phenotype less severe than those males presenting the same pathogenic variants [1]. This is assumed to be due to the X-chromosome inactivation (XCI) phenomena, where preferential inactivation of the X-chromosome with the normal allele results in an incomplete penetrance or in variable clinical presentations [2–8]. Such an event can complicate risk assessment and genetic counselling in females, such as in the case of *FMR1* full mutation carriers [9]. A skewed XCI pattern may also be observed in the general population, known to increase with the aging process by mechanisms not fully understood, and it has been associated with females with idiopathic premature ovarian insufficiency or recurrent pregnancy loss [10–18]. The methylation pattern of the CpG islands adjacent to the polymorphic CAG repeat in the androgen receptor (*AR*) gene located near the X-inactivation center correlates well with the methylation of the X-chromosome [19]. The method described by Allen et al. (1992) to analyze the *AR* gene methylation status, designated Human Androgen-Receptor Assay (HUMARA), is considered the gold standard assay to determine the XCI pattern [20]. The frequency of non-informativity, as well as the technical challenges (e.g., allele dropout, presence of stutter bands, among others) inherent to the properties of GC-rich polymorphic regions, justifies the testing of other loci with CpG sites susceptible to inactivation [21]. The CGG repetitive region in the promotor of the fragile mental retardation protein translational regulator (*FMR1*) gene, when expanded to over 200 triplets (full mutation), leads to methylation and gene silencing and, consequently, Fragile X syndrome (FXS; OMIM #3000624) [22,23]. Past studies have assessed the utility of using the *FMR1* locus (Xq27.3) to determine the XCI pattern. Analysis was done by visual inspection of PCR fragments, after digestion with a methylation-sensitive endonuclease [24,25]. Nowadays, there are more appropriate methods to determine the methylation status of expanded *FMR1* alleles (in the context of FXS diagnosis), resorting to capillary electrophoresis coupled with specific software [26,27]. In this study, we used the AmplideX mPCR *FMR1* Kit (mPCR) and aimed to assess the XCI pattern using the *FMR1* gene methylation status in samples with normal-sized alleles, compared to the corresponding pattern obtained using HUMARA. We used samples from females showing random and skewed XCI patterns, which enabled the establishment of intervals for XCI pattern categorization when using the *FMR1* gene. The absence of a complete overlap between the results of both assays led us to propose a set of key points and a stepwise analytical procedure

for accurate XCI pattern result assessment, while minimizing the effect of the pitfalls underlying the use of this locus.

2. Material and Methods

2.1. Retrospective study

The results of the X-chromosome inactivation (XCI) pattern category of 74 female samples, with a mean age of 25.2 ± 15.1 years (range 1–68), tested between January 2014 and February 2021 at our laboratory in a diagnostic context, were re-analyzed. This data set includes carriers of X-autosome translocation, Fragile-X full mutation, and presumably pathogenic recessive X-linked variants.

2.2. Study cohort

Forty-one females with fertility issues and seven oocyte donors (*n* = 48), with a mean age of 33.8 ± 5.0 years (range 20–40), were selected from a set of females previously recruited at the Centre for Medically Assisted Procreation and Public Gamete Bank, Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), Centro Hospitalar Universitário do Porto (CHUPorto), as part of B. Rodrigues's Ph.D. research project (2020.119/097-DEFI/099-CE). Genomic DNA extracted from peripheral blood samples were used to assess the methylation status of the *FMR1* and *AR* genes, as well as the number of CGG and CAG repeats, respectively.

2.3. Molecular Studies

2.3.1. *AR* gene methylation status

The *AR* methylation status was determined using the Human Androgen-Receptor *(AR*) Assay, HUMARA (hereafter Assay A), adapted from Allen et al. (1992) [20]. Briefly, 150 ng of DNA was digested with the methylation-sensitive enzyme *Hha*I (New England BioLabs®, Ipswich, Massachusetts, USA) and incubated overnight at 37°C, followed by 20 min of inactivation at 70°C. The amplification of the digested and undigested fragments was performed using PCR Master Mix (Promega®, Madison, WI, USA) and NED-labelled forward primer, with the following thermal cycling conditions: denaturation of 5 min at 95°C, 29 cycles of 45 s at 95°C, 30 s at 62°C, and 30 s at 72°C, with a final extension of 15 min at 72°C. PCR products were resolved on ABI PRISM® 3130xl Genetic Analyzer (Applied Biosystems™, Foster City, CA, USA) using 500 ROX™ size standard (Gene Scan™, Warrington, UK) and were analysed using GeneMapper® software version 4.0 (Applied Biosystems™). The percentage of *AR* allele methylation

was determined using the area values, normalized to the corresponding undigested allele area. Assay A cut-offs used were previously defined (Supplementary Table S1) [21].

2.3.2. *FMR1* gene methylation status

The *FMR1* promoter methylation status was determined using the AmplideX® mPCR *FMR1* kit (AmplideX® mPCR *FMR1* Kit, Asuragen®, Austin, TX, USA) (hereafter Assay B), according to the manufacturer's instructions [26]. The optimized protocol and components of the kit used a total of 160 ng of DNA and the methylation-sensitive endonuclease *Hpa*II. Undigested fragments were amplified with FAM-labelled primers, whereas digested fragments (after methylation-sensitive *Hpa*II restriction) were amplified using HEX-labelled primers. PCR products were resolved as described above, using a capillary electrophoresis (CE) injection at 2 kV for 5 s, as recommended for normal alleles with a signal intensity beyond the CE instrument saturation limit. This assay determines both the number of CGG repeats and the methylation percentage of each *FMR1* allele. Fragment size and height were determined using GeneMapper® software version 4.0 (Applied Biosystems™). The percentage of the allele methylation status was determined using the ratio of allele height (digested/undigested) and was normalized to the sum of the ratios digested/undigested of each allele.

2.4. Statistical analysis: reproducibility and performance comparison of assays

Statistical analysis was performed with SigmaPlot version 14.0 (Systat Software® Inc., Chicago, IL, USA) and the Real Statistics Resource Pack for Excel (available at https://www.real-statistics.com/, accessed on 22-12-2021). The Shapiro−Wilk test was used to assess the normality in the retrospective study.

Four samples were arbitrarily selected (http://vassarstats.net/, accessed on 12-09-2021) for each random and skewed category, as established by Assay A. These samples were used to assess the reproducibility of each assay, addressed in the context of the analytical replication of the results (internal reproducibility). The assays were performed in triplicate in three independent experiments.

The Assay B methylation ranges and categories, i.e., random, moderate, and high, were selected to maximize the similarity between the results of both assays.

The internal reproducibility between replicates for each assay was performed with the Brown−Forsythe test, considering the following four XCI patterns: random, moderately

skewed, highly skewed, and completely skewed. The Kolmogorov−Smirnov test was used to compare the two assays in terms of the XCI pattern categories. The Chi-square test was used to compare the number of samples (frequency) by XCI pat-tern categories between the two assays. A significance level of 0.05 was considered for all of the statistical tests.

3. Results

3.1. Retrospective analysis of samples tested by HUMARA

HUMARA determines the methylation percentage of the *AR* gene alleles and is commonly applied to establish the pattern of XCI. Among the retrospective cases (*n* = 74), a high number of uninformative samples with respect to the XCI pattern and the polymorphic locus (CAG homozygosity) was observed in 10.8% (*n* = 8/74). The informative samples were categorized according to ranges previously defined (Supplementary Table S1) [21]. The majority of the samples (71.2%, *n* = 47/66) showed a random XCI pattern, 15.2% (*n* = 10/66) showed moderately skewed, 1.5% (*n* = 1/66) highly skewed, and 12.1% (*n* = 8/66) completely skewed patterns. In this retrospective analysis, the full spectrum of normally distributed XCI patterns was observed (Shapiro−Wilk test: W = 0.9770, *p* = 0.870; mean [54:46] ± [23:75] and median [55:45]). Borderline results such as [78:22] (*n* = 2/66), [80:20] (*n* = 1/66), [81:19] (*n* = 1/66), [82:18] (*n* = 1/66), [88:12] (*n* = 1/66), and [90:10] (*n* = 3/66), were observed among 13.6% of the informative samples (*n* = 66). When combining the results that were either borderline or uninformative (homozygous), the XCI pattern was not (accurately) defined in around 23% (*n* = 17/74) of the samples. This observation prompted us to analyse if the methylation percentage of another locus, namely the *FMR1* gene, could be used to determine the XCI pattern, thereby reducing the number of inconclusive results.

3.2. Establishment of the intervals for XCI pattern determination using *FMR1* mPCR

We determined the methylation status of the *AR* and *FMR1* genes using Assay A and B, respectively, in 48 samples. The XCI patterns were categorized according to the ranges previously defined for Assay A (Supplementary Table S1) [21]. Table 1 shows the ranges used after the Assay B values were adjusted, to minimize the differences between the two assays. Assay A results showed a random XCI pattern in 75% of samples (*n* = 36/48), high skewing in 12.5% (*n* = 6/48), and moderate skewing in 12.5% (*n* = 6/48) (Table 2). Two samples showed borderline results: [19:81] (*n* = 1/48) and [91:9] (*n* = 1/48). Assay

B revealed a random XCI pattern in 77% of the samples (*n* = 37/48), 16.7% (*n* = 8/48) had moderate skewing, and only 6.3% (*n* = 3/48) with high skewing (Table 2). Borderline results were obtained in four samples: [73:27] (*n* = 2/48), [90:10] (*n* = 1/48) and [91:9] (*n* = 1/48). For both assays, a random XCI pattern was observed in the majority of the samples, and no case was seen to have complete skewing (>99:1 or >1:99).

Table 1. Categories established for XCI pattern determination using *FMR1* mPCR (Assay B).

	Assay B			
XCI Pattern categories	**Random**	**Moderately skewed**	**Highly skewed**	**Completely skewed**
Ranges	[25:75-75:25] [75:25-25:75]	[11:89-25:75[ ]75:25-89:11]	[1:99-11:89[ ]89:11-99:1]	[0:100-1:99[ ]99:1-100:0]

Table 2. Summary of the results obtained in the forty-eight samples.

	XCI pattern categories		
Assay	**Random** ***n*** **(%)**	**Moderately skewed** ***n*** **(%)**	**Highly skewed** ***n*** **(%)**
A	36 (75 %)	6 (12.5 %)	6 (12.5 %)
B	37 (77 %)	8 (16.7 %)	3 (6.3 %)
Partial χ^2 (df = 1)	0.014	0.286	1.00
p **value***	0.907	0.593	0.317

* Chi-square value calculation was based on absolute frequencies in each category ($\chi 2 = 1.299$; df = 2; $p = 0.522$); *n* = Number of samples in the category; % = Percentage of samples in the category.

3.3. Performance comparison showed no differences between the assays

Table 2 summarizes the results obtained after testing 48 samples using the two assays and categorized according to the newly established XCI pattern based on the *FMR1* methylation status. In Figure 1, representative electropherograms of samples number 29 (Figure 1A), 37 (Figure 1B), 39 (Figure 1C), and 45 (Figure 1D) are shown. When the XCI pattern categories were considered, i.e., shape of the distributions per category, no significant differences were found between the assays (Kolmogorov-Smirnov Test: D = 0.333, $p = 0.976$; Table 2). Up to 77% of the samples showed concordant results in both assays: 66% of samples (*n* = 32/48) with a random pattern, 6.4% (*n* = 3/48) were moderately skewed, and 4.2% (*n* = 2/48) were highly skewed (Supplementary Table S2).

Reproducibility within each assay was evaluated by testing eight samples in triplicate. A significant Brown−Forsythe test result (Assay A: F* = 2089.739, df = 7, $p < 0.005$; Assay B: F* = 1704.674, df = 7, $p < 0.005$) suggested our data did not meet the assumption of homogeneity of variances and is indicative of a low internal reproducibility.

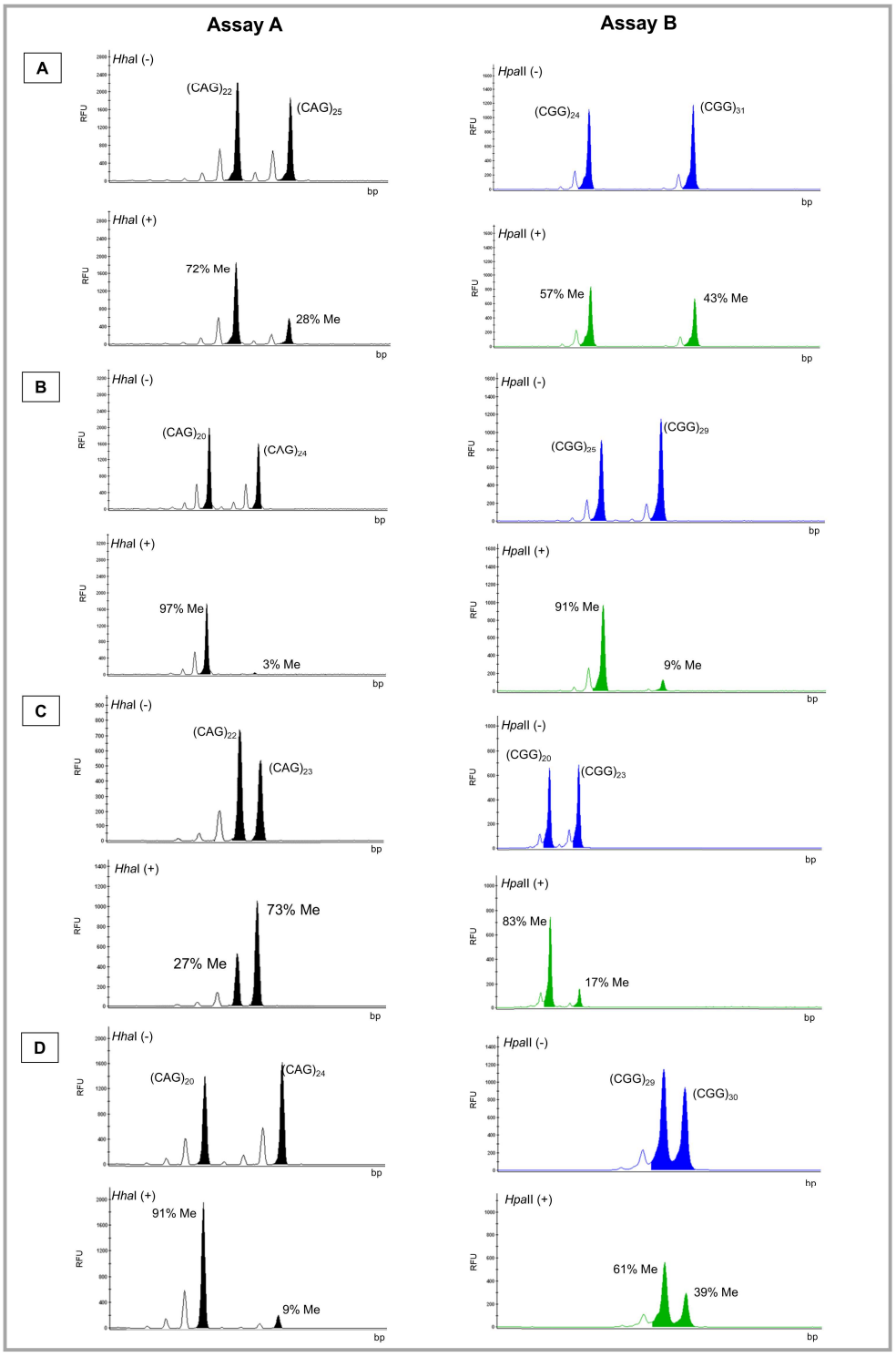


Figure 1. Electropherogram of the A and B Assay results for samples number 29 (A), 37 (B), 39 (C), and 45 (D). Digested allele (+) and undigested allele (−). Number of CAG (AR) and CGG (FMR1) repeats and respective percentage of methylation (Me). Blue (FAM-labelled) – undigested allele; green (HEX-labelled) – digested allele; RFU – Relative fluorescence units; bp – Base pairs.

3.4. Small and large repeat number differences among *FMR1* alleles explain distinct XCI pattern categorizations

We investigated whether samples showing small differences (one or two repeats) in *FMR1* allele sizes could fall into distinct categories when comparing the results of both assays. In fact, in four samples where *FMR1* alleles differed by less than two re-peats, the XCI pattern category was distinct among the assays (samples number 40, 42, 45, and 47; Figure 1D and Supplementary Table S2). We further speculated that a large difference in the number of repeats among *FMR1* alleles (≥7 CGGs) might also contribute to distinct XCI pattern categorizations, due to the preferential amplification of the smaller allele. This was observed in five cases (Supplementary Table S2; samples number 38, 41, 44, 46, and 48). Borderline results, such as that of sample 39 (Figure 1C and Supplementary Table S2), also led to the distinct categorization of the XCI pattern among the two assays. We speculate that the cumulative effect of preferential amplification of the smaller allele and inaccuracy introduced due to the presence of stutter bands, particularly in samples with small differences in allele repeat numbers (≤2), may influence the quantification of the methylation percentage and consequently the XCI pattern.

4. Discussion

In the current study, we compared the results of 48 females using *FMR1* methylation status and HUMARA in order to establish the *FMR1* XCI pattern range limits. Our results revealed 77% agreement regarding the XCI pattern categorization between the two assays. Furthermore, analysis of the discordant samples led us to suggest a set of key points that should be taken into consideration when evaluating the XCI pattern using the percentage of *FMR1* methylation. In samples where alleles differed by up to two trinucleotide repeats, determination of the XCI pattern might be challenging. We also observed that when the stutter band from the larger allele, with a similar amplification intensity, overlapped with the primary band of the smaller allele, it interfered with the calculation of the area of the true allele, hampering an accurate XCI categorization. Previous reports suggest that inaccurate XCI pattern categorization might occur in samples with AR alleles differing by ≤2 repeats, due to difficulty in discriminating the true allele area value from the overlapping stutter bands [21,28]. In samples where this is observed, the XCI pattern should not be based on the obtained methylation pattern. In alleles that differ by ≥7 repeats, the use of *FMR1* allele methylation percentage might also become problematic due to amplification bias. In order to minimize the putative preferential amplification of the smaller allele, we suggest testing these samples in

triplicate, as our analysis showed that the use of mean results of replicates was more accurate even in cases of low internal reproducibility. We suggest that these key-points should also be considered when using methylation of other X-linked loci by any methodology. Nevertheless, further studies, such as an increased number of test samples, are warranted. In fact, one of the major limitations of our study is the small number of samples, as well as the trial-and-error process used to establish the *FMR1*-based XCI pattern categories, to maximize the number of samples that matched the HUMARA XCI categories.

Furthermore, we cannot exclude that the mechanism of epigenetic control of the *FMR1* gene is different from that of the *AR* gene. This could be the reason underlying the distinct results observed in cases with high skewing obtained by Assay A, which is not reproduced in Assay B. Of note, the methylation status of the *FMR1* normal alleles was determined using a commercial mPCR kit designed to determine the percentage of methylation of the expanded alleles with the purpose of diagnosing Fragile-X syndrome. Other methodologies or methylation-sensitive enzymes should also be tested. In addition, other candidate genes for further additional inactivation tests could be of interest, particularly those previously described [29,30]. Nevertheless, in might be difficult to find a gene that completely mirrors the methylation status of the *AR* gene, allowing for the extrapolation of the X-chromosome methylation pattern. Although the large majority of the X-chromosome genes are subjected to methylation, about 15% escaped XCI [31,32]. Furthermore, an incomplete correlation between XCI pattern and the expression of the *AR* gene was reported. The discrepancy between *AR* locus methylation and *AR* expression was observed in healthy females with a skewed XCI pattern, suggesting that the CpG methylation did not extend to the promoter or at least was unable to inhibit the *AR* promotor [33]. This could also explain why in some samples, XCI skewing was observed using Assay A and not Assay B (Supplementary Table S2; Sample number 47). Future expression studies should be undertaken to support this hypothesis.

We suggest the use of the *FMR1* methylation status to infer the XCI pattern, after triplicate testing, when (i) alleles differ by more than two and less than seven CGG repeats, (ii) in the absence of amplification bias (maximum difference between undigested allele area/height of: 1:1.3 or 1.3:1), and (iii) the result is established outside borderline (± 2) limits. More studies are needed to further increase the power of the statistical analysis and to validate the ranges established for each pattern category, as well as the expression analysis of both *AR* and *FMR1* genes. Nevertheless, this study shows that

the *FMR1* gene methylation mirrors the methylation of the *AR* gene and can be used to determine the XCI pattern, with a particular impact on samples where *AR* locus homozygosity is observed. Our results warrant a new line of investigation to further increase the knowledge of the epigenetic mechanisms that regulate XCI and the methylation of *FMR1* and *AR* genes.

5. Conclusion

This study shows that *FMR1* gene methylation mirrors the methylation of the *AR* gene and can be used to determine the XCI pattern, with a particular impact in samples where *AR* locus homozygosity is observed. Our results warrant a new line of investigation to further increase the knowledge of the epigenetic mechanisms that regulate XCI and the methylation of *FMR1* and *AR* genes.

Supplementary Materials

The following are available online at https://www.mdpi.com/article/10.3390/genes13030419/s1. Table S1: References ranges used for X-chromosome inactivation pattern determination by HUMARA (Assay A). Table S2: Summary of the results of *AR* (Assay A) and *FMR1* (Assay B) genes methylation pattern.

Author Contributions

B.R., P.J. and A.J.A.N. conceived and designed the study, A.J.A.N. performed the statistical analysis with B.R., who also carried out laboratory work, analysed the data, and drafted the manuscript. E.V.-F., A.G., V.S., N.M., I.M., and R.S. provided critical feedback, helped conduct the research, and contributed toward the manuscript. A.G. and R.S. were responsible for the retrospective studies of XCI (HUMARA) in our laboratory. All of the authors discussed the final results and critically reviewed the manuscript.

Funding

This work was supported by national funds: FCT/FSE (Fundação para a Ciência e a Tecnologia) – Project Reference EXPL/BIA-REP/0423/2021—X-EPIFERTILITY and DEFI (Departamento de Ensino, Formação e Investigação) – Reference 2015-DEFI/145/12. Thanks are due to the financial support to CESAM (UID/AMB/50017/2019), to UMIB (UIDB/00215/2020, and UIDP/00215/2020) and ITR (LA/P/0064/2020) to FCT/MCTES through national funds (SFRH/BD/136398/2018 to

B.R.), and the co-funding by the FEDER, within the PT2020 Partnership Agreement and Compete 2020.

Institutional Review Board Statement

This research was approved by the Hospital's Ethics Committee (2020.119 /097-DEFI/099-CE) as part of B. Rodrigues's Ph.D. studies.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

Data are contained within the article or supplementary material.

Acknowledgments

We gratefully acknowledge the Center for Medically Assisted Procreation/Public Gamete Bank, Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), Centro Hospitalar Universitário do Porto (CHUPorto). Without the invaluable help of all collaborators, our work would not have been possible. We also would like to acknowledge Carmina Serrano (Palex Medical) for the availability, critical input, and discussion about the results of the XCI pattern using mPCR *FMR1* Kit.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Migeon BR. X-linked diseases: susceptible females. Genet Med. 2020;22:1156–74.

2. Patrat C, Ouimette J, Rougeulle C. X chromosome inactivation in human development. Development. 2020;147.

3. Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB, Dupre Y, et al. DNA methylation profiles of human active and inactive X chromosomes. Genome Res. 2011;21:1592–600.

4. Galupa R, Heard E. X-chromosome inactivation: A crossroads between chromosome architecture and gene regulation. Annu Rev Genet. 2018;52:535–66.

5. Vacca M, Della Ragione F, Scalabrì F, D'Esposito M. X inactivation and reactivation in X-linked diseases. Semin Cell Dev Biol. 2016;56:78–87.

6. Viggiano E, Politano L. X chromosome inactivation in carriers of fabry disease: Review and meta-analysis. Int J Mol Sci. 2021;22.

7. Soltanzadeh P, Friez MJ, Dunn D, von Niederhausern A, Gurvich OL, Swoboda KJ, et al. Clinical and genetic characterization of manifesting carriers of DMD mutations. Neuromuscul Disord. 2010;20:499–504.

8. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nat Lett. 2005;434:400–4.

9. Taylor AK, Safanda JF, Fall MZ, Quince C, Lang KA, Hull CE, et al. Molecular Predictors of Cognitive Involvement in Female Carriers of Fragile X Syndrome. JAMA J Am Med Assoc. 1994;271:507–14.

10. Van den Veyver IB. Skewed X inactivation in X-linked disorders. Semin Reprod Med. 2001;19:183–91.

11. Sharp A, Robinson D, Jacobs P. Age- and tissue-specific variation of X chromosome inactivation ratios in normal women. Hum Genet. 2000;107:343–9.

12. Sandovici I, Naumova AK, Leppert M, Linares Y, Sapienza C. A longitudinal study of X-inactivation ratio in human females. Hum Genet. 2004;115:387–92.

13. Shvetsova E, Sofronova A, Monajemi R, Gagalova K, Draisma HHM, White SJ, et al. Skewed X-inactivation is common in the general female population. Eur J Hum Genet. 2018;27:455–65.

14. Zito A, Davies MN, Tsai PC, Roberts S, Andres-Ejarque R, Nardone S, et al. Heritability of skewed X-inactivation in female twins is tissue-specific and associated with age. Nat Commun. 2019;10:1–11.

15. Wong CCY, Caspi A, Williams B, Houts R, Craig IW, Mill J. A longitudinal twin study of skewed X chromosome-inactivation. PLoS One. 2011;6:1–5.

16. Hatakeyama C, Anderson CL, Beever CL, Peñaherrera MS, Brown CJ, Robinson WP. The dynamics of X-inactivation skewing as women age. Clin Genet. 2004;66:327–32.

17. Miranda-Furtado CL, Luchiari HR, Chielli Pedroso DC, Kogure GS, Caetano LC, Santana BA, et al. Skewed X-chromosome inactivation and shorter telomeres associate with idiopathic premature ovarian insufficiency. Fertil Steril. 2018;110:476-485.e1.

18. Uehara S, Hashiyada M, Sato K, Sato Y, Fujimori K, Okamura K. Preferential X-chromosome inactivation in women with idiopathic recurrent pregnancy loss. Fertil Steril. 2001;76:908–14.

19. Liu X, Walsh J, Mburu P, Kendrick-jones J, Cope JT, Steel KP, et al. A promotor mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat Genet. 1997;15:57–61.

20. Allen RC, Y.Zoghbi H, It Annemarie B. Moseley HMR, Belmont JW. Methylation of Hpall and Hhal Sites Near the Polymorphic CAG Repeat in the Human Androgen-Receptor Gene Correlates with X Chromosome Inactivation. Am J Hum Genet. 1992;51:1229-12:1229–39.

21. Amos-Landgraf JM, Cottle A, Plenge RM, Friez M, Schwartz CE, Longshore J, et al. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. Am J Hum Genet. 2006;79:493–9.

22. Nobile V, Pucci C, Chiurazzi P, Neri G, Tabolacci E. Dna methylation, mechanisms of FMR1 inactivation and therapeutic perspectives for fragile X syndrome. Biomolecules. 2021;11:1–17.

23. Kraan CM, Godler DE, Amor DJ. Epigenetics of fragile X syndrome and fragile X-related disorders. Dev Med Child Neurol. 2019;61:121–7.

24. Carrel L, Willard HF. An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. Am J Med Genet. 1996;64:27–30.

25. Wolff DJ, Schwartz S, Carrel L. Molecular determination of X inactivation pattern correlates with phenotype in women with a structurally abnormal X chromosome. Genet Med. 2000;2:136–41.

26. Asuragen I. Amplidex mPCR FMR1 kit - Protocol Guide. Asuragen, Inc. 2019;1–15.

27. Grasso M, Boon EMJ, Filipovic-Sadic S, Van Bunderen PA, Gennaro E, Cao R, et al. A novel methylation PCR that offers standardized determination of FMR1 methylation and CGG repeat length without southern blot analysis. J Mol Diagnostics. 2014;16:23–31.

28. Karasawa M, Tsukamoto N, Yamane A, Okamoto K, Maehara T, Yokohama A, et al. Analysis of the distribution of CAG repeats and X-chromosome inactivation status of HUMARA gene in healthy female subjects using improved fluorescence-based assay. Int J Hematol. 2001;74:281–6.

29. Bertelsen B, Tümer Z, Ravn K. Three new loci for determining X chromosome inactivation patterns. J Mol Diagnostics. 2011;13:537–40.

30. Musalkova D, Minks J, Storkanova G, Dvorakova L, Hrebicek M. Identification of novel informative loci for DNA-based X-inactivation analysis. Blood Cells, Mol Dis. 2015;54:210–6.

31. Berletch JB, Yang F, Xu J, Carrel L, Disteche CM. Genes that escape from X inactivation. Hum Genet. 2011;130:237–45.

32. Disteche CM, Berletch JB. X-chromosome inactivation and escape. J Genet. 2015;94:591–9.

33. Swierczek SI, Piterkova L, Jelinek J, Agarwal N, Hammoud S, Wilson A, et al. Methylation of AR locus does not always reflect X chromosome inactivation state. Blood. 2016;119:100–10.

Supporting Information

1. Supplementary Figures

Table S1. References ranges used for X-chromosome inactivation pattern determination by HUMARA (Assay A).

	Assay A			
XCI pattern categories	**Random**	**Moderately skewed**	**Highly skewed**	**Completely skewed**
Ranges*	[20:80-80:20] [80:20-20:80]	[10:90-20:80[ ]80:20-90:10]	[1:99-10:90[ ]90:10-99:1]	[0:100-1:99[ ]99:1-100:0]

*Ranges were previously defined by Amos-Landgraf et al., [21].

Table S2. Summary of the results of *AR* (Assay A) and *FMR1* (Assay B) genes methylation pattern. In the material and method section, the calculations on the X-chromosome inactivation (XCI) pattern and respective category on the basis of the percentage (%) of methylation of each gene are explained.

	Assay A			Assay B			
Sample number	% methylation		XCI pattern	% methylation		XCI pattern	$(CGG_n)_{allele\ 2}$ - $(CGG_n)_{allele\ 1}$
	Allele 1	Allele 2		Allele 1	Allele 2		
1	48	52	Random	44	56	Random	7
2	42	58	Random	30	70	Random	28
3	56	44	Random	34	66	Random	2
4	71	29	Random	36	64	Random	1
5	70	30	Random	57	43	Random	3
6	65	35	Random	32	68	Random	11
7	50	50	Random	49	51	Random	12
8	27	73	Random	37	63	Random	1
9	63	37	Random	41	59	Random	11
10	60	40	Random	48	52	Random	2
11	56	44	Random	49	51	Random	13
12	37	63	Random	69	31	Random	5
13	48	52	Random	56	44	Random	7
14	55	45	Random	40	60	Random	13
15	28	72	Random	35	65	Random	10
16	41	59	Random	58	42	Random	10
17	63	37	Random	45	55	Random	6
18	55	45	Random	38	62	Random	6
19	66	34	Random	69	31	Random	3
20	40	60	Random	48	52	Random	1
21	58	42	Random	43	57	Random	14
22	72	28	Random	40	60	Random	11
23	32	68	Random	44	56	Random	1
24	47	53	Random	59	41	Random	2
25	43	57	Random	56	44	Random	9
26	59	41	Random	59	41	Random	10
27	44	56	Random	45	55	Random	9
28	64	36	Random	44	56	Random	11
29	72	28	Random	57	43	Random	7
30	35	65	Random	73	27	Random	23
31	29	71	Random	28	72	Random	2
32	23	77	Random	73	27	Random	5
38	71	29	Random	17	83	Moderately skewed	11
39	27	73	Random	83	17	Moderately skewed	3
40	41	59	Random	82	18	Moderately skewed	1
41	52	48	Random	79	21	Moderately skewed	7
33	82	18	Moderately skewed	79	21	Moderately skewed	13
34	83	17	Moderately skewed	22	78	Moderately skewed	16
35	83	17	Moderately skewed	15	85	Moderately skewed	10
42	83	17	Moderately skewed	32	68	Random	1
43	19	81	Moderately skewed	35	65	Random	3
44	83	17	Moderately skewed	97	3	Highly skewed	7
45	91	9	Highly skewed	61	39	Random	1
46	93	7	Highly skewed	81	19	Moderately skewed	9
47	93	7	Highly skewed	28	72	Random	1
48	94	6	Highly skewed	72	28	Random	7
36	93	7	Highly skewed	91	9	Highly skewed	1
37	97	3	Highly skewed	91	9	Highly skewed	4

n = number of CGGs

Chapter VI. Experimental and dataset resources

This chapter outlines the acquisition of data and biological samples obtained from oocyte donors and females experiencing fertility challenges undergoing intracytoplasmic sperm injection (ICSI) following informed consent. Participant recruitment and sample management procedures are detailed. Furthermore, the chapter elaborates on processing and storage protocols tailored to the specific characteristics of each sample type. A total of 210 participants were enrolled in the study and various types of samples were collected from each participant. The results derived from these samples are present in detail throughout the thesis with particular emphasis on chapters III, IV and V.

1. The importance of a biorepository for female fertility and *FMR1* research

The establishment of a comprehensive database, coupled with a biorepository, represents a significant advancement in biomedical research, particularly in the field of female fertility. This integrated resource serves as a valuable asset for conducting studies and scientific research related to female fertility.

The primary objective of creating this database and biorepository was not only to support the current doctoral research project but also to lay the foundation for future scientific research. By safeguarding these valuable biological samples and ensuring their long-term accessibility, we can explore novel research questions and hypotheses related to female fertility without the need for additional sample collection efforts. The ability to repurpose samples and reanalyze existing data is a key advantage, offering significant savings in terms of time and resources. Moreover, the preservation of high-value and difficult-to-obtain samples, such as CCs, is of paramount importance. These samples are often scarce and challenging to collect, making their availability crucial for advancing research in this field. By maximizing the utility of these precious resources, the biorepository not only supports current studies but also ensures that future research can build upon this foundation, driving progress in our understanding of the complex mechanisms underlying female fertility.

In the subsequent section, we will delve into the detailed methodology employed for the harvesting and storage of various biological samples, ensuring their integrity and suitability for research purposes.

2. Biological sample collection, processing and storage

2.1. Sample collection process

2.1.1. Population recruited as part of the Ph.D. project

Participant recruitment occurred between January 16th, 2020, and February 18th, 2022, at the Centre for Medically Assisted Procreation/ Public Gamete Bank, Centro Materno-Infantil do Norte Dr. Albino Aroso (CMIN), Unidade Local de Saúde de Santo António (ULSSA). A total of thirty oocyte donors (*n* = 30) with a mean age of 27.7 ± 3.8 (range 20 - 33) and one hundred and eighty females (*n* = 180) with fertility issues undergoing intracytoplasmic sperm injection (ICSI). No exclusion criteria were applied to oocyte donors, all females who accepted to volunteer were included in the study. In case of females with fertility issues, inclusion or exclusion was determined based on the fertilization technique performed. Females undergoing techniques other than ICSI were excluded.

Informed consent was obtained from all participants involved in the study. I collaborated in the preparation of the informed consent provided by the clinician to the participants (Annex 1).

2.1.2. Biological material

Peripheral blood, follicular fluid and CCs were collected from each participant.

2.1.2.1. Peripheral blood

Peripheral blood samples were collected using ethylenediamine tetra acetic acid (EDTA)-containing tubes and stored at room temperature in an upright position in the laboratory until further use.

2.1.2.2. Follicular fluid and granulosa cumulus cells

During the follicular aspiration procedure (follicular puncture), the cumulus-oocyte complexes (COCs) and follicular fluid (FF) were retrieved from the follicles. After being recovered from FF, the COCs were exposed to Cumulase® (Origio®, Måløv, Denmark) enzyme for maximum 30 seconds, which weakens the link between the oocyte and the granulosa CCs. Subsequently, the oocyte was mechanically denuded by gentle pipetting in flushing medium (Origio®). Isolated CCs were diluted in DMEM/F-12 medium supplemented with GlutaMAX™ (Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12, Gibco™, Thermo Fisher Scientific, Waltham, Massachusetts, USA), and FF was incubated in a 5% humidified atmosphere with CO_2 at 37ºC, and stored until further use.

Figure 1 illustrates the process of COCs retrieval and the subsequent steps involved in the CCs isolation.

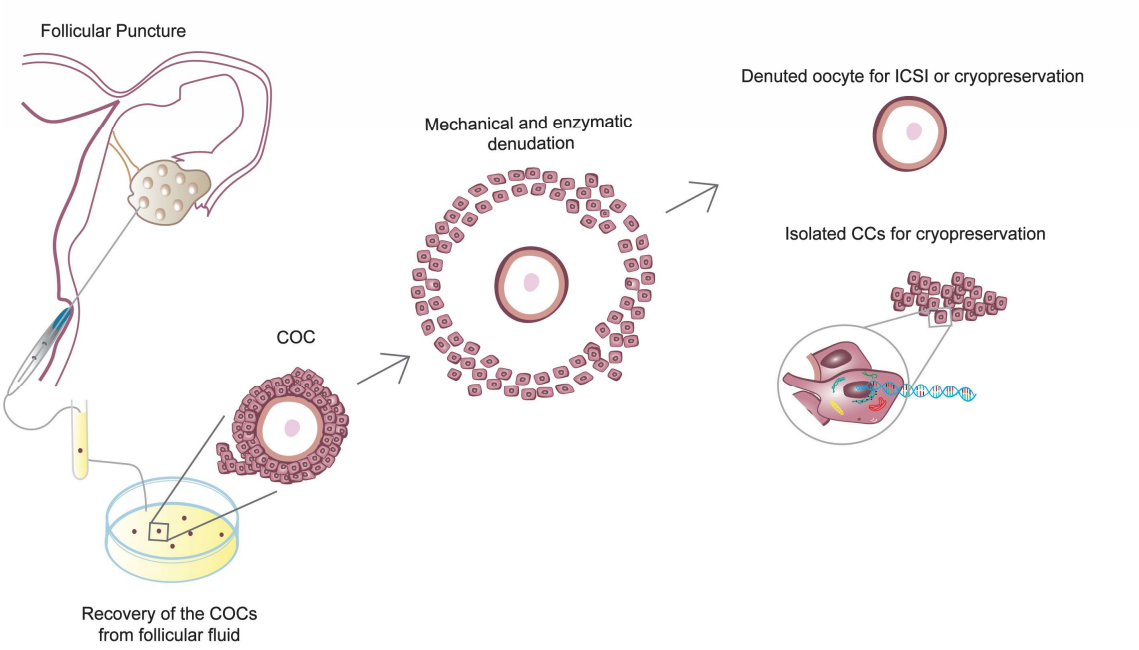


Figure 1. Schematic representation of cumulus cell isolation. Following follicular puncture, the follicular aspirate, containing cumulus-oocyte complexes (COCs), is transferred to a petri dish for microscopic recovery of COCs. Oocytes are subsequently denuded, and the cumulus cells (CCs) are isolated. Isolated CCs are diluted in DMEM/F12 medium supplemented with GlutaMAX™ (Thermo Fisher Scientific), and the follicular fluid (FF) is incubated in a humidified atmosphere with 5% CO_2 at 37°C. FF – Follicular fluid; COCs – Cumulus-oocyte complexes; CCs – Cumulus cells; ICSI – Intracytoplasmic sperm injection.

2.2. Processing and long-term storage of biological samples

Before processing, the samples were pseudonymized using a key-coding system: INF (females with fertility issues undergoing ICSI) and DOV (oocyte donors). Samples were processed and stored on the day of collection to minimize degradation, following specific protocols designed by our team to ensure their integrity during long-term storage. Additionally, for peripheral blood samples, prior to storage, blood plasma separation and DNA extraction was performed (Annex 2).

3. Biorepository and associated data

3.1. Clinical information

Clinical information of each participant was obtained by the team of health professionals involved in the process, and coded as INF or DOV. Subsequently, all acquired clinical

data were systematically integrated and organized into an Excel spreadsheet (Figure 2A).

Age, weight, height, body mass index (BMI), smoking habits, anti-Müllerian hormone levels (AMH), number of antral follicles in each ovary, stimulation protocol, drug, total dose of follicle-stimulating hormone (FSH), stimulation days, trigger, number of COCs retrieved, endometrium thickness and glucose levels were obtained from oocyte donors and females with fertility issues undergoing ICSI. In the case of oocyte donors, information such as donation number, number of children, and ethnicity was obtained, along with counts of immature, lysed, abnormal, metaphase II, and vitrified oocytes. For females with fertility issues undergoing ICSI it was still obtained type of infertility, attempt number, duration of infertility, cause of infertility, follicle-stimulating hormone (FSH), luteinizing hormone (LH) and estradiol (E2). E2 and progesterone values on the trigger day were also obtained. Additionally, other parameters, such as oocyte quality and fertilization success are detailed in Figure 2A.

3.1.1. Challenges

The lack of standardized clinical record-keeping practices limited the scope and depth of the clinical data available for analysis. To address these data consistency challenges, we advocate for the adoption of standardized data extraction protocols. Specifically, we propose implementing the protocol outlined in Annex 3, which was meticulously designed for data collection from females with ovulatory dysfunction. Given the prevalence of ovulatory dysfunction as a primary cause of infertility and the growing interest in its genetic underpinnings, this protocol offers a robust framework for ensuring data quality and comparability. Moreover, this protocol can be readily adapted for future participant recruitment initiatives, thereby facilitating ongoing research and analysis.

3.2. Laboratory studies

The database was designed for ongoing updates, with a key objective being the incorporation of study results as biological samples are utilized. A significant contribution from our team was the integration of laboratory data into the database, particularly focusing on characterizing the CGG repetitive tract of the *FMR1* gene. This data is crucial for the current study's primary objective of investigating the *FMR1* gene's influence on female reproductive health. By combining comprehensive clinical data with genetic information, we can identify correlations and patterns that may be overlooked in studies examining only one aspect. This integrated approach allows us to explore not only the

association between the *FMR1* gene and fertility but also its potential impact on this critical process.

Summary of studies and results for all biorepository samples: CGG repeat length of the *FMR1* gene, *FMR1* allele categories, sub-genotype classification, AGG number, AGG interspersion pattern, *allelic score*, *FMR1* allelic complexity classification, *FMR1* methylation status, CAG repeat length of the *AR* gene, *AR* methylation status, and X-chromosome inactivation pattern (Figure 2B).

3.2.1. Ongoing studies

Our research group is currently investigating the potential influence of cumulus cell DNA integrity on oocyte quality. To date, we have assessed the DNA integrity of blood and CCs in a cohort of fifty-seven samples (n = 57), employing parameters such as tail length, tail intensity (% of DNA in tail), and tail moment. These preliminary findings are being further analyzed, with ongoing evaluation of DNA integrity in the remaining samples. Consequently, the results from the fifty-seven samples have been incorporated into our established database.

Consistent updates to both the database and biorepository are essential to maintain their utility for future research.

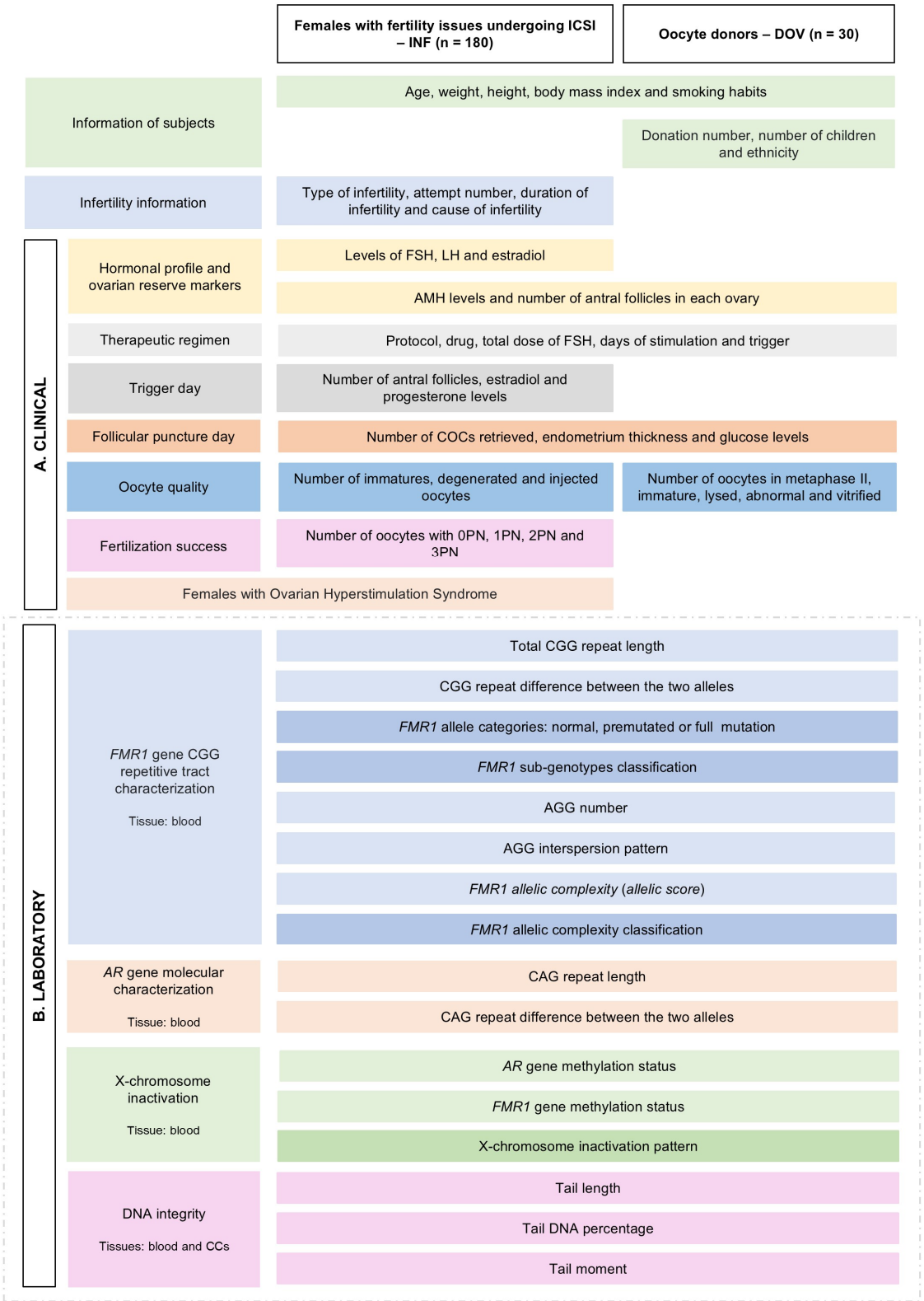


Figure 2. Summary of data categories included in the database. FSH – Follicle-stimulating hormone; LH – Luteinizing hormone; AMH – Anti-Müllerian hormone; AFC – Antral follicle count; COCs – Cumulus-oocyte complexes; PN – Pronuclei; CCs – Cumulus cells.

4. Leveraging a comprehensive and adaptable database to advance understanding of female reproductive health

The meticulously organized database facilitates rapid access to desired information, streamlining data extraction, analysis, and inter-institutional information exchange. Its adaptability allows for continuous updates, ensuring its relevance and currency with evolving research. Despite challenges encountered, access to a diverse, expansive, and adaptable database that integrates clinical information with laboratory data and has the potential to incorporate various data types in the future enables comprehensive and extensive analyses. This may contribute to a deeper understanding of the complex biological processes underlying female reproductive health, which have remained elusive.

Consequently, this database may facilitate the identification of associations between genetic factors and various parameters of fertility, including ovarian stimulation, oocyte quality, and fertilization success. The insights gained from this database have the potential to inform the discovery of novel biomarkers and early interventions aimed at improving reproductive health.

4

Discussion and future perspectives

This doctoral thesis aimed to investigate the influence of the *FMR1* gene's repetitive tract complexity on female fertility, with a particular emphasis on its impact on various fertility outcomes. The exploration of the *FMR1* repetitive tract was essential to i) introduce a new approach to evaluate the CGG/AGG substructure; ii) establish links with fertility outcomes following the stratification of samples based on allelic complexity; iii) identify potential risk factors associated with poor fertility prognosis in females undergoing intracytoplasmic sperm injection (ICSI) procedures. The results of this thesis yielded significant findings that contribute to a deeper understanding of the *FMR1* gene's role in female fertility. These findings have the potential to stimulate new research questions and directions, bridging the gap between basic science and clinical applications. Ultimately, these findings may lead to the development of practical solutions for improving fertility outcomes.

Algorithm development

A significant limitation in previous research linking the *FMR1* gene to fertility has been the exclusive focus on the total length of the CGG repeat [1–8], overlooking other relevant characteristics of this repetitive region. While recent studies have explored the influence of AGG interruptions on the development of fragile X-associated primary ovarian insufficiency (FXPOI, OMIM #311360) in female premutation carriers, a comprehensive analysis considering the cumulative effect of total CGG repeat length, AGG presence, and AGG interspersion pattern has been lacking [9–11].

To address this gap, this investigation developed a novel algorithm that offers a more nuanced perspective on the analysis of this polymorphic repetitive region (Chapter I). The formula devised by our team, incorporating the number and pattern of AGG interspersions in conjunction with the total CGG repeat length, generates an *allelic score* that reflects the complexity of the CGG/AGG substructure. The identification of distinct behaviors arising from allelic complexity combinations has led to the stratification of the population into two well-defined groups.

The intricate allelic landscape of the *FMR1* gene enable robust stratification of study populations across various clinical phenotypes, including potentially fertile females (Chapter I), *FMR1* premutation (PM) carriers (Chapter II), and infertile females (Chapter III). These findings underscore the robustness of the employed methodology. Moreover, the stratification employed in this study demonstrates its relevance by providing valuable insights into the impact of the *FMR1* gene on infertility phenotypes. This approach holds the potential to elucidate the association between allelic complexity within the *FMR1* gene and distinct fertility outcomes observed in infertile females.

Identification of *FMR1* premutation carriers

During this research project three *FMR1* PM carriers were identified within a cohort 124 infertile females (Chapter III), recruited at the Centro Materno Infantil do Norte Dr. Albino Aroso, Centro Hospitar Universitário de Santo António, Unidade Local de Saúde de Santo António. This institution serves as the primary referral center for two major regions: Douro Litoral, constituted by the Metropolitan area of Porto and Tâmega e Sousa, and Trás-os-Montes e Alto Douro, constituted by the Douro, Alto Tâmega and Terras de Trás-os-Montes [12]. The estimated prevalence of *FMR1* PM in the general population ranges from 1 in 150 to 1 in 300 females [13–17]. In our cohort of infertile females, a notably high prevalence was observed (1 in 41), consistent with previous studies in populations with specific infertility phenotypes [18,19]. These findings underscore the importance of implementing carrier screening protocols in fertility centers and emphasize the need for comprehensive genetic counseling to inform females about the risks associated with CGG expansion, including the risk of having a child with fragile X syndrome (FXS, OMIM #300624). Beyond the risk of developing fragile X-associated primary ovarian insufficiency (FXPOI, OMIM # 311360), *FMR1* PM carriers may also experience other comorbidities such as autoimmune disorders, chronic pain conditions, endocrine disorders, and mental health issues [20,21]. Therefore, family screening is essential for understanding the genetic landscape, identifying additional mutated or expanded cases, and providing appropriate guidance and follow-up. Expanding *FMR1* PM screening to females of reproductive age can increase the likelihood of successful pregnancies by broadening reproductive options, such as fertility preservation through oocyte retrieval, *in vitro* fertilization (IVF), and preimplantation genetic diagnosis [22,23]. Females/ couples with knowledge about of their reproductive options can make more empowered decisions regarding procreation.

A previous study reported the occurrence of *FMR1* PM in females with primary ovarian insufficiency (POI) within the Portuguese population [24]. However, a definitive assessment of its prevalence, especially among infertile females, remains necessary. This information is crucial for the implementation of effective genetic counseling strategies in families with a history of POI, enabling informed reproductive decision-making. Moreover, our preliminary findings indicate a potentially significant prevalence of *FMR1* PM in this specific population, warranting further investigation with a larger cohort to establish precise prevalence estimates.

Deciphering X-Chromosome Inactivation with *FMR1* Methylation: New Categories and Considerations

Skewed X-chromosome inactivation (XCI), characterized by the preferential inactivation of one X chromosome, can result in imbalanced expression of *FMR1* alleles, potentially influencing the severity of the fragile X PM phenotype [25,26]. For example, females with fragile X PM may be more susceptible to FXPOI if the normal *FMR1* allele is preferentially silenced (methylated) on the inactive X-chromosome [27]. Our study (Chapter V) demonstrated a correlation between *FMR1* methylation and Androgen Receptor (*AR*) gene methylation (HUMARA), a widely recognized gold standard for XCI studies. Our results showed that evaluation of *FMR1* allele methylation using methylation-specific PCR (mPCR) allows assessment of the XCI pattern. This suggests a potential association between these processes, despite their generally considered independence. This study is the first to establish a set of categories for determining XCI patterns using *FMR1* mPCR. We identified key factors essential for accurate XCI assessment using this method: i) *FMR1* alleles should differ by no more than two and less than seven CGG repeats; ii) absence of amplification bias; and iii) results should fall outside the limit of error (±2%). A primary advantage of our approach is its potential to reduce inconclusive results compared to HUMARA, especially in cases with borderline or uninformative data (homozygous *AR* gene).

Refining Risk Prediction for FXPOI: A Combined Analysis of CGG Repeats, AGG number and pattern

Understanding the risk associated with developing FXPOI has significant implications for family planning and the overall health of PM carriers. The risk of FXPOI in PM carries is nonlinearly correlated with the total CGG repeat length [28,29]. Allen et al. (2021) reported that PM carriers with CGG repeat length between 85 and 89 have a risk of developing FXPOI [30]. Our study (Chapter III) identified three PM carriers with CGG repeat lengths of 56, 59, and 75. While these carriers exhibit a decreased risk of developing FXPOI, they experience fertility problems, including ovulatory dysfunction. The PM carrier with a CGG repeat length of 75 exhibited characteristics typical of FXPOI, including diminished ovarian reserve, poor response to ovarian stimulation, and impaired fertilization. However, the other two demonstrated an unexpected excessive response to controlled ovarian stimulation, contradicting prior findings [31,32]. This hyper-responsiveness necessitated the postponement of embryo transfer due to the elevated risk of ovarian hyperstimulation syndrome.

Previous studies have yielded conflicting results regarding the influence of AGG interruptions on FXPOI development [9,10]. A comprehensive evaluation of the combined effects of the total CGG repeat length and the number and pattern of AGG interruptions has not been undertaken. In our study (Chapter II), we employed a novel approach that considers the combined allelic complexity of both the normal and the permutated alleles. Despite exhibiting similar allelic complexity, the three premutated alleles were associated with distinct clinical phenotypes. Notably, normal-sized alleles did not display comparable complexity. Based on these observations, we hypothesize that a combined assessment of allelic complexity may serve as a valuable predictor of FXPOI risk. To address the limitations of sample size and the inherent uncertainty regarding individual FXPOI risk development, we conducted a comprehensive literature review to identify collaborators for a larger-scale study (willing to share results on both normal and PM alleles). By pooling data from multiple studies, we were able to analyze a more robust dataset and investigate the relationship between allelic complexity and the development of FXPOI. Our results demonstrated evidence of a link between the allelic complexity of the normal allele and the age at amenorrhea, a clinical manifestation associated with FXPOI development. These findings suggest that a combined analysis of allelic complexity for both the normal and PM alleles of the *FMR1* gene may be a valuable tool for early identification of females at increased risk of FXPOI. In a recent study, that employed our allelic complexity formula to assess PM and intermediate alleles. The authors demonstrated the utility of allelic complexity in estimating the risk of POI/DOR and emphasized the importance of considering the normal allele [33]. By refining risk prediction for FXPOI in female PM carriers, we aim to improve the likelihood of achieving successful future reproductive outcomes. However, further validation studies are necessary, including the exploration of additional clinical outcomes.

FMR1 Allelic Complexity: A Potential Predictor of Fertilization Success

Chapter III explores the potential predictive value of *FMR1* allelic complexity in determining ovarian function and IVF success. The study focuses on the impact of normal-sized alleles (CGG repeats less than 45) and intermediate alleles (CGG repeats between 45 and 54), extending beyond the established risk factors associated with PM alleles (CGG repeats between 55 and 200). Our results revealed a significant negative correlation between the allelic complexity of allele 1 (shorter in size) and the number of successfully fertilized oocytes (two pronuclei [2PN] oocytes). Females carrying allele 1 with an allelic complexity exceeding 150 may be at an elevated risk of fertilization failure. These findings suggest that *FMR1* allelic complexity could potentially serve as a predictive marker for IVF success in females undergoing ICSI. It is hypothesized that

oocyte cytoplasmic immaturity may contribute to fertilization failure, and adjustments to *in vitro* procedures aimed at enhancing oocyte maturation could potentially improve success rates for females with different *FMR1* allelic complexities. The association between allelic complexity and cytoplasmic maturity warrants further investigation and robust validation in larger patient cohorts.

Cumulus Cell DNA Integrity and Oocyte Competence

Given the intimate relationship between cumulus cells (CCs) and the oocyte, the DNA integrity of CCs has been investigated as a potential marker for oocyte quality and its impact on fertilization [34–41]. However, previous studies have produced conflicting results, and research on this tissue is often hindered by its limited accessibility. In Chapter IV, we evaluated the potential of cumulus cell DNA damage as a predictor of oocyte competence and subsequent fertilization success in females with diverse infertility phenotypes. Additionally, we explored the use of DNA damage in whole blood as a potential marker for DNA damage in CCs. Our findings in Chapter IV demonstrated that the extent of DNA damage in whole blood does not accurately reflect DNA damage in CCs. This limitation poses challenges for studies involving larger populations due to the difficulties associated with collecting CCs. Furthermore, our results revealed a paradoxical positive correlation between DNA damage levels in CCs and successful fertilization (evidenced by 2PN oocytes) in samples from females with ovulatory dysfunction. In addition, females with high incidence of unfertilized oocytes (zero pronuclei) exhibited lower levels of DNA damage. This intriguing observation suggests that fertilization failure could be attributed to incomplete oocyte maturation, particularly cytoplasmic immaturity. Further studies are necessary to elucidate the precise relationship between low DNA damage and oocyte maturation. Nonetheless, our findings tentatively indicate that CCs with higher DNA damage levels may be associated with oocytes in an optimal state (good oocyte competence) for sperm reception, leading to a greater likelihood of fertilization success.

Is *FMR1* a Potential Predictor of Oocyte Competence?

Given the observed association between *FMR1* allelic complexity and the number of fertilized oocytes (2PN), as well as the increased risk of fertilization failure in females carrying allele 1 with allelic complexity greater than 150, we hypothesize that these females may exhibit a higher propensity for oocyte maturation failure. Oocyte maturity is well-established as crucial for successful fertilization [42,43]. The formation of pronuclei and subsequent completion of fertilization are contingent upon both nuclear and cytoplasmic maturity of the oocyte. Although nuclear maturity can be assessed by the

presence of the first polar body, a specific marker for cytoplasmic maturity remains elusive. Nonetheless, both aspects contribute significantly to fertilization success [44–50]. Future investigations are warranted to explore the potential of the *FMR1* gene as a predictive biomarker for oocyte competence, particularly with respect to cytoplasmic maturity.

Unravelling the Link Between Allelic Complexity and DNA Damage

The *FMR1* allelic complexity and DNA damage levels in CCs appear to be correlated with IVF success, as indicated by the number of 2PN oocytes. These findings necessitate further investigation into the underlying mechanisms. Specifically, it is crucial to elucidate the relationship between allelic complexity and DNA damage, and how these factors collectively impact IVF outcomes. To address this (Chapter IV), an initial stratification of the ovulatory dysfunction population based on allelic complexity was conducted. However, due to the limited sample size, definitive conclusions were challenging to draw. To overcome this bottleneck, the established biological sample repository, coupled with the comprehensive clinical and laboratory database (Chapter VI), will be fundamental in identifying suitable samples for future investigations. These investigations will focus on exploring the potential association between DNA damage in CCs and the complexity of the *FMR1* repetitive tract.

References

1. Banks N, Patounakis G, Devine K, DeCherney AH, Widra E, Levens ED, et al. Is FMR1 CGG repeat length a predictor of in vitro fertilization stimulation response or outcome? Fertil Steril. 2016;105:1537-1546.e8.

2. 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.e1.

3. Gustin SLF, Ding VY, Desai M, Leader B, Baker VL. Evidence of an age-related correlation of ovarian reserve and FMR1 repeat number among women with “normal” CGG repeat status. J Assist Reprod Genet. 2015;32:1669–76.

4. 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;166:502–7.

5. Morin SJ, Tiegs AW, Franasiak JM, Juneau CR, Hong KH, Werner MD, et al. FMR1 gene CGG repeat variation within the normal range is not predictive of ovarian response in IVF cycles. Reprod Biomed Online. 2016;32:496–502.

6. Batiha O, Shaaban ST, Al-Smadi M, Jarun Y, Maswadeh A, Alahmad NA, et al. A study on the role of FMR1 CGG trinucleotide repeats in Jordanian poor ovarian responders. Gene. 2021;767:145174.

7. Cogendez E, Ozkaya E, Eser AÇ, Eken M, Karaman A. Can FMR1 CGG repeat lengths predict the outcome in ICSI cycles? Ginekol Pol. 2022;93:735–41.

8. Jin X, Zeng W, Xu Y, Jin P, Dong M. Cytosine–guanine–guanine repeats of FMR1 gene negatively affect ovarian reserve and response in Chinese women. Reprod Biomed Online. 2023;49:103779.

9. Allen EG, Glicksman A, Tortora N, Charen K, He W, Amin A, et al. FXPOI: Pattern of AGG interruptions does not show an association with age at amenorrhea among women with a premutation. Front Genet. 2018;9:1–7.

10. Friedman-Gohas M, Kirshenbaum M, Michaeli A, Domniz N, Elizur S, Raanani H, et al. Does the presence of AGG interruptions within the CGG repeat tract have a protective effect on the fertility phenotype of female FMR1 premutation carriers? J Assist Reprod Genet. 2020;37:849–54.

11. Lekovich J, Man L, Xu K, Canon C, Lilienthal D, Stewart JD, et al. CGG repeat length and AGG interruptions as indicators of fragile X–associated diminished ovarian reserve. Genet Med. 2018;20:957–64.

12. Rodrigues-Martins D, Vale-Fernandes E, Leal C, Barreiro M. Influence of Women´s Residence Region on Assisted Reproduction Treatments-Experience of a Tertiary Center in Northern Portugal. J Bras Reprod Assist. 2022;26:73–7.

13. Tassone F, Iong KP, Tong TH, Lo J, Gane LW, Berry-Kravis E, et al. FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. Genome Med. 2012;4:1–13.

14. Hantash FM, Goos DM, Crossley B, Anderson B, Zhang K, Sun W, et al. 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. Genet Med. 2011;13:39–45.

15. Ma Y, Wei X, Pan H, Wang S, Wang X, Liu X, et al. The prevalence of CGG repeat expansion mutation in FMR1 gene in the northern Chinese women of reproductive age. BMC Med Genet. 2019;20:1–5.

16. Greene R, Pisano MM. Prevalence of CGG Expansions of the FMR1 Gene in a US Population-Based Sample. Birth Defects Res C Embryo Today. 2012;90:133–54.

17. Seltzer MM, Baker MW, Hong J, Maenner M, Greenberg J, Mandel D. Prevalence of CGG expansions of the FMR1 gene in a US population-based sample. Am J Med Genet Part B Neuropsychiatr Genet. 2012;159 B:589–97.

18. Dean DD, Agarwal S, Kapoor D, Singh K, Vati C. Molecular Characterization of FMR1 Gene by TP-PCR in Women of Reproductive Age and Women with Premature Ovarian Insufficiency. Mol Diagnosis Ther. 2018;22:91–100.

19. Agustí I, Méndez M, Borrás A, Goday A, Guimerà M, Peralta S, et al. Prevalence of the FMR1 Gene Premutation in Young Women with a Diminished Ovarian Reserve Included in an IVF Program: Implications for Clinical Practice. Genes (Basel). 2024;15:8:1008.

20. 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.

21. Allen EG, Charen K, Hipp HS, Shubeck L, Amin A, He W, et al. Clustering of comorbid conditions among women who carry an FMR1 premutation. Genet Med. 2020;22:758–66.

22. Tassone F, Protic D, Allen EG, Archibald AD, Baud A, Brown TW, et al. Insight and Recommendations for Fragile X-Premutation-Associated Conditions from the Fifth International Conference on FMR1 Premutation. Cells. 2023;12:2330.

23. Ain Q, Hwang YH, Yeung D, Panpaprai P, Iamurairat W, Chutimongkonkul W, et al. Population-based FMR1 carrier screening among reproductive women. J Assist Reprod Genet. 2024.

24. Neves AR, Pais AS, Ferreira SI, Ramos V, Carvalho MJ, Estevinho A, et al. Prevalence of Cytogenetic Abnormalities and FMR1 Gene Premutation in a Portuguese Population with Premature Ovarian Insufficiency. Acta Med Port. 2021;34:580–5.

25. Migeon BR. X-linked diseases: susceptible females. Genet Med. 2020;22:1156–74.

26. Sun Z, Fan J, Wang Y. X-Chromosome Inactivation and Related Diseases. Genet Res (Camb). 2022;2022:1391807.

27. Hadd AG, Filipovic-Sadic S, Zhou L, Williams A, Latham GJ, Berry-Kravis E, et al. A methylation PCR method determines FMR1 activation ratios and differentiates premutation allele mosaicism in carrier siblings. Clin Epigenetics. 2016;8:1–9.

28. Mailick MR, Hong J, Greenberg J, Smith L, Sherman S. Curvilinear Association of CGG Repeats and Age at Menopause in Women with FMR1 Premutation Expansions. Am J Med Genet B Neuropsychiatr Genet. 2014;0:705–11.

29. Ennis S, Ward D, Murray A. Nonlinear association between CGG repeat number and age of menopause in FMR1 premutation carriers. Eur J Hum Genet. 2006;14:253–5.

30. Allen EG, Charen K, Hipp HS, Shubeck L, Amin A, He W, et al. Refining the risk for fragile X–associated primary ovarian insufficiency (FXPOI) by FMR1 CGG repeat size. Genet Med. 2021;23:1648–55.

31. Bibi G, Malcov M, Yuval Y, Reches A, Ben-Yosef D, Almog B, et al. The effect of CGG repeat number on ovarian response among fragile X premutation carriers undergoing preimplantation genetic diagnosis. Fertil Steril. 2010;94:869–74.

32. Persico T, Tranquillo ML, Seracchioli R, Zuccarello D, Sorrentino U. PGT-M for Premature Ovarian Failure Related to CGG Repeat Expansion of the FMR1 Gene. Genes (Basel). 2024;15:6.

33. Quilichini J, Perol S, Cuisset L, Grotto S, Fouveaut C, Barbot JC, et al. Stratification of the risk of ovarian dysfunction by studying the complexity of intermediate and premutation alleles of the FMR1 gene. Am J Med Genet Part A. 2024;194:1–10.

34. Abu-Hassan D, Koester F, Shoepper B, Schultze-Mosgau A, Asimakopoulos B, Diedrich K, et al. Comet assay of cumulus cells and spermatozoa DNA status, and the relationship to oocyte fertilization and embryo quality following ICSI. Reprod Biomed Online. 2006;12:447–52.

35. McKenzie LJ, Pangas SA, Carson SA, Kovanci E, Cisneros P, Buster JE, et al. Human cumulus granulosa cell gene expression: A predictor of fertilization and embryo selection in women undergoing IVF. Hum Reprod. 2004;19:2869–74.

36. Raman RS, Chan PJ, Corselli JU, Patton WC, Jacobson JD, Chan SR, et al. Comet assay of cumulus cell DNA status and the relationship to oocyte fertilization via intracytoplasmic sperm injection. Hum Reprod. 2001;16:831–5.

37. Tola EN, Koşar PA, Karatopuk DU, Sancer O, Oral B. Effect of DNA damage of cumulus oophorus cells and lymphocytes analyzed by alkaline comet assay on oocyte quality and intracytoplasmic sperm injection success among patients with polycystic ovary syndrome. J Obstet Gynaecol Res. 2018;45:609–18.

38. Roberts R, Franks S, Hardy K. Culture environment modulates maturation and metabolism of human oocytes. Hum Reprod. 2002;17:2950–6.

39. Barcena P, López-Fernández C, García-Ochoa C, Obradors A, Vernaeve V, Gosálvez J, et al. Detection of DNA damage in cumulus cells using a chromatin dispersion assay. Syst Biol Reprod Med. 2015;61:277–85.

40. Baratas A, Gosálvez J, de la Casa M, Camacho S, Dorado-Silva M, Johnston SD, et al. Cumulus Cell DNA Damage as an Index of Human Oocyte Competence. Reprod Sci. 2022;29:3194–200.

41. Turathum B, Gao EM, Chian RC. The function of cumulus cells in oocyte growth and maturation and in subsequent ovulation and fertilization. Cells. 2021;10:1–18.

42. Marteil G, Richard-Parpaillon L, Kubiak JZ. Role of oocyte quality in meiotic maturation and embryonic development. Reprod Biol. 2009;9:203–24.

43. Swain JE, Pool TB. ART failure: Oocyte contributions to unsuccessful fertilization. Hum Reprod Update. 2008;14:431–46.

44. Trebichalská Z, Kyjovská D, Kloudová S, Otevřel P, Hampl A, Holubcová Z. Cytoplasmic maturation in human oocytes: An ultrastructural study. Biol Reprod. 2021;104:106–16.

45. Coticchio G, Dal Canto M, Renzini MM, Guglielmo MC, Brambillasca F, Turchi D, et al. Oocyte maturation: Gamete-somatic cells interactions, meiotic resumption,

cytoskeletal dynamics and cytoplasmic reorganization. Hum Reprod Update. 2014;21:427–54.

46. Rienzi L, Balaban B, Ebner T, Mandelbaum J. The oocyte. Hum Reprod. 2012;27:2–22.

47. Eppig JJ, O'Brien M, Wigglesworth K. Mammalian oocyte growth and development in vitro. Mol Reprod Dev. 1996;44:260–73.

48. Shu Y, Gebhardt J, Watt J, Lyon J, Dasig D, Behr B. Fertilization, embryo development, and clinical outcome of immature oocytes from stimulated intracytoplasmic sperm injection cycles. Fertil Steril. 2007;87:1022–7.

49. Chian RC, Buckett WM, Tan SL. In-vitro maturation of human oocytes. Reprod Biomed Online. 2004;8:148–66.

50. Younis JS, Radin O, Izhaki I, Ben-Ami M. Does first polar body morphology predict oocyte performance during ICSI treatment? J Assist Reprod Genet. 2009;26:561–7.

Rights and permissions

Rights and permissions

Bárbara Rodrigues, Emídio Vale-Fernandes, Vanessa Sousa, Isabel Marques, Rosário Santos, António J A Nogueira, Paula Jorge. Exploring the predictive value of the *FMR1* gene allelic complexity for *in vitro* fertilization success. Submitted for publication in Journal of Assisted Reproduction and Genetics.

Bárbara Rodrigues*, Vanessa Sousa*, Filipa Esteves, Emídio Vale-Fernandes, Solange Costa, Daniela Sousa, Raquel Brandão, Carla Leal, Isabel Gaivão, João Paulo Teixeira, António J A Nogueira, Paula Jorge. Cumulus cell DNA damage linked to fertilization success in females with an ovulatory dysfunction phenotype. Accepted for publication in Frontiers in Cell and Developmental Biology, October 24th, 2024. doi: 10.3389/fcell.2024.1448733

Bárbara Rodrigues, Vanessa Sousa, Carolyn M. Yrigollen, Flora Tassone, Olatz Villate, Emily G. Allen, Anne Glicksman, Nicole Tortora, Sarah L. Nolin, António J. A. Nogueira, Paula Jorge. *FMR1* allelic complexity in premutation carriers provides no evidence for a correlation with age at amenorrhea. Reprod Biol Endocrinol. 2024 Jun 21;22(1):71. doi: 10.1186/s12958-024-01227-5

Bárbara Rodrigues, Ana Gonçalves, Vanessa Sousa, Nuno Maia, Isabel Marques, Emídio Vale-Fernandes, Rosário Santos, António J A Nogueira, Paula Jorge. Use of the *FMR1* Gene Methylation Status to Assess the X-Chromosome Inactivation Pattern: A Stepwise Analysis. Genes (Basel). 2022 Feb 25;13(3):419. doi: 10.3390/genes1303041

Bárbara Rodrigues, Emídio Vale-Fernandes, Nuno Maia, Flávia Santos, Isabel Marques, Rosário Santos, António J A Nogueira, Paula Jorge. Development and Validation of a Mathematical Model to Predict the Complexity of *FMR1* Allele Combinations. Front Genet. 2020 Nov 13:11:557147. doi: 10.3389/fgene.2020.557147

* Authors contributed equally

Annexes

Annex 1. Informed consent, prepared by Bárbara Rodrigues and Emídio Vale-Fernandes, used throughout this project.

TERMO DE CONSENTIMENTO INFORMADO

Título do estudo de investigação. Impacto do perfil clínico e metabólico, do gene *FMR1* e outros modificadores genéticos na fertilidade feminina.

O estudo acima mencionado será realizado pelos alunos de doutoramento, **Emídio Vale-Fernandes** e **Bárbara Rodrigues**, orientados, respetivamente, por Mariana Monteiro (MD, PhD, ICBAS/UP) e Paula Jorge (PhD, ICBAS/UP), no âmbito dos projetos de doutoramento em Ciências Médicas e Ciências Biomédicas do Instituto Ciências Biomédicas Abel Salazar.

Eu, abaixo-assinado _____ (nome da participante).

Fui informada de que o estudo de investigação acima mencionado tem como objetivo entender o papel do gene *FMR1* e outros modificadores genéticos bem como identificar novos biomarcadores com impacto na fertilidade feminina.

Explicaram-me que para atingirem este objetivo serão analisadas amostras de ADN de sangue e material que sobra do procedimento de aspiração folicular. Sei que uma parte das amostras será utilizada de imediato para fazer alguns testes e que outra parte vai ser armazenada para ser utilizada posteriormente. Sei ainda que neste estudo está prevista a realização de testes genéticos, tendo-me sido explicado em que consistem.

Foi-me garantido que todos os dados relativos à identificação dos participantes neste estudo são confidenciais e que será mantido o anonimato. Compreendi a informação que me foi dada, tive oportunidade de fazer perguntas e todas as minhas dúvidas foram esclarecidas.

Sei que posso recusar ou interromper a qualquer momento a minha participação no estudo, sem nenhum tipo de penalização por este facto.

Assim, aceito participar de livre vontade no estudo acima mencionado. Concordo que sejam efetuados todos os testes necessários.

Autorizo a divulgação dos resultados obtidos no meio científico, garantindo o anonimato.

Nome da participante no estudo:

Data	Assinatura
___/___/_____	_____

Nome do investigador:

Data	Assinatura
___/___/_____	_____

Annex 2. Protocols for the processing and storage of biological materials

These protocols describe the procedures for the processing and storage of biological samples with the aim of preserving sample quality, integrity, and traceability. The protocols were designed to ensure the effective utilization of these samples in current and future research endeavors while adhering to relevant regulations and ethical standards.

Processing and storage

A. Follicular fluid (FF):

1. Centrifuge FF samples at 719 *g* for 10 minutes at 8°C.
2. Transfer 2 ml of FF to microtubes.
3. Centrifuge at 20817 *g* for 10 minutes at 21°C.
4. Transfer the supernatant to new microtubes.
5. Centrifuge again at 20817 *g* for 10 minutes at 21°C. If cell debris remains visible, repeat step 4.
6. Transfer supernatant to new microtubes and store at -70°C until use.
7. Record the number of FF aliquots stored and their physical location

B. Cumulus cells (CCs):

Note: CCs was transported in a flushing medium (Origio, Måløv, Denmark) together with DMEM/F-12 (Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12) medium with GlutaMAX™ supplementation (Gibco™, Thermo Fisher Scientific, Massachusetts, USA) and it is often possible to see a layer of liquid parafilm above this medium.

1. Centrifuge at 719 *g* for 10 minutes at 8°C.
2. Carefully remove the liquid parafilm and discard the supernatant.
3. Resuspend the pellet in 200 µL of DMEM/F-12 medium with GlutaMAX™ supplementation.
4. Transfer the cell suspension to a 1.5 mL cryovial (containing 1000 µl of DMEM/F-12 medium supplemented with GlutaMAX™. Add 300 µL of 100% dimethyl sulfoxide (DMSO) (Invitrogen™, Thermo Fisher Scientific). DMSO serves as a cryoprotectant but is cytotoxic at room temperature; therefore, cryovials must be immediately placed on ice.

5. Transfer the cryovial to a controlled-rate freezer -1°C/ minute cell (Mr. Frosty™ Freezing Container, Thermo Fisher Scientific) at -70°C for 24 hours. Subsequently, store the cryovials in the liquid phase of a nitrogen tank until use.
6. Record of the number of CCs aliquots stored and their physical location.

Record keeping: **Thawing of CCs frozen in liquid nitrogen**

1. Remove the cryopreserved vial containing CCs from the liquid nitrogen container and immediately transfer it to a designated ice bucket.
2. Allow the vial to thaw in a 37°C water bath, avoiding direct immersion of the vial in the water. Gently agitate the vial intermittently to facilitate thawing.
3. Transfer the thawed medium containing the CCs to a 2 mL microcentrifuge tube.
4. Centrifuge at 720 *g* for 5 minutes.
5. If a visible pellet is not observed, repeat the centrifugation at the same speed for an additional 5 minutes or until the pellet becomes apparent.
6. Discard the supernatant, retaining the pellet containing the CCs.
7. Wash the pellet with 200 µL of phosphate-buffered saline (PBS, 1x).
8. Centrifuge at 720 *g* for 5 minutes and discard the supernatant.
9. The resulting CCs pellet is now ready for subsequent procedures, such as the alkaline comet assay or RNA extraction.

C. Peripheral blood and plasma blood

Note: Peripheral blood samples shall be collected in ethylenediaminetetraacetic acid (EDTA)-containing tubes and stored in an upright position at room temperature to minimize erythrocyte hemolysis.

1. Centrifuge at 719 *g* at 8°C for 10 minutes.
2. Gently transfer the top layer of plasma into microtubes. Ensure a minimum collection of 500 µL of plasma.
3. Store plasma and blood aliquots at -70°C for future use.
4. Record the number of aliquots stored and their physical location
5. The remaining peripheral blood, devoid of plasma, is utilized for DNA extraction.

D. DNA extraction from peripheral blood using the Salting-Out method.

1. Transfer the EDTA-anticoagulated peripheral blood sample to a 50 mL tube and add erythrocyte lysis buffer (TLE) to achieve a final volume of 50 mL. Gently

homogenize the mixture by inverting the tube, and subsequently store the sample on ice for a minimum of 20 minutes.

2. Centrifuge at 720 *g* for 10 minutes at 8°C.
3. Discard the supernatant and thoroughly resuspend the leukocyte pellet in the residual liquid.
4. Add 10 mL of TLE and homogenize
5. Centrifuge at 720 *g* for 10 minutes at 8°C.
6. Discard the supernatant and thoroughly resuspend the pellet in the residual liquid.
7. Add 3 mL of nucleus lysis buffer (TLN) and thoroughly resuspend the pellet.
8. Add 100 µL of Proteinase K (10 mg/ml) and mix carefully by inverting the tube.
9. Add 300 µL of 10% (w/v) sodium dodecyl sulfate (SDS) and homogenize.
10. Incubate at 37°C overnight.
11. Add 1 mL of saturated NaCl solution and shake vigorously for 15 seconds.
12. Centrifuge at 1620 *g* for 15 minutes.
13. Carefully transfer the supernatant into 50 ml tubes. <u>If the supernatant is not clear, repeat steps 11 and 12.</u>
14. Precipitate the DNA by adding 2 volumes of absolute ethanol (twice the volume of supernatant obtained in step 13).
15. Carefully wind the precipitated DNA into the tip of a Pasteur pipette, forming a coil.
16. Transfer the DNA-laden pipette to a 12 mL tube containing approximately 5 mL of 70% (v/v) ethanol. Wash the spooled DNA thoroughly by immersing and rotating the pipette in the ethanol for a minimum of 10 minutes, repeating this process three to four times.
17. Transfer the coiled DNA to a 1.5 mL tube containing DNA rehydration solution. Resuspend the DNA by incubating the tube at 65°C for 30 minutes, followed by an overnight incubation at room temperature on a rotary shaker.
18. Quantify the isolated genomic DNA (gDNA) samples using a NanoDrop ND-1000 Spectrophotometer (version 3.3, Life Technologies™) at a wavelength of 260 nm.
19. Store the DNA at 4°C
20. Record of the physical location.

E. Preparation of the solutions required for the different protocol

1. Phosphate Buffered Saline solution (PBS, 1 x), pH = 7: NaCl 0.138 M and KCl 0.0027 M

Components	
PBS Tablets (Calbiochem, Merck, Darmstadt, Germany)	1 x
dH_2O	Up to 1000 mL

- Autoclave and store at room temperature.
- Filter before use with 0.2 µm filters.

2. Erythrocyte lysis buffer (TLE): NH4Cl 155 mM, KHCO3 10 mM and EDTA-NaOH 1 mM

Components	
1 M NH_4Cl	155 mL
1 M $KHCO_3$	10 mL
0.5 M EDTA-NaOH pH = 7.4	2 mL
dH_2O	Up to 1000 mL

- Store at room temperature.

3. Nucleus lysis buffer (TLN): Tris-HCl 10 mM, NaCl 400 mM and EDTA-NaOH 2 mM

Components	
1 M Tris-HCl pH = 8.0	10 mL
5 M NaCl	80 mL
0.5 M EDTA-NaOH pH = 8.0	4 mL
dH_2O	Up to 1000 mL

- Autoclave and store at room temperature.

4. Sodium dodecyl sulfate (SDS, 10%)

Components	
SDS	50 g
dH_2O	Up to 500 mL

- Store at 4 °C.

5. Saturated NaCl solution (NaCl, 5 M)

Components	
NaCl	43.5 g
dH_2O	Up to 100 mL

- Autoclave and store at room temperature.

6. Ethanol 70%

Components	
Absolute ethanol	350 mL
dH_2O	Up to 150 mL

- Store at room temperature.

Annex 3. Proposed form for future investigation in ovulatory dysfunction

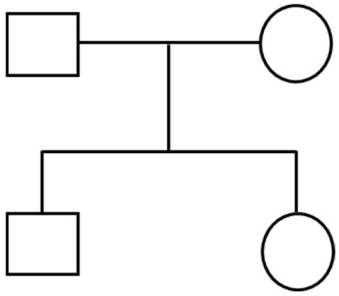
Objective: To establish a standardized protocol for genetic investigations aimed at elucidating the underlying genetic factors contributing to ovulatory dysfunction.

Protocol

Sample identification: _____

Age (years) _____Weight (Kg) _____Height (m) _____

1. Family history



Origin _____

Consanguinity ☐

Other diseases in the family:

_____

2. Clinical data

Infertility

- Primary ☐ Secundary ☐ Duration (months): _____
- Smoking habits: Yes ☐ No ☐ Ex-smoker ☐

Other diseases? _____

1. Anovulation or Oligoovulation

- Irregular menstruation ☐ Since _____ (years)
- Amenorrhea ☐ or Oligomenorrhea ☐ Since _____ (years)

Hormonal treatments: ☐ If yes, when? _____

---

2. Hyperandrogenism

- Acne ☐
- Hirsutism ☐

3. Body Mass Index (Kg/m^2)

- < 18.5 (Low weight) ☐
- Between 18.5 and 24.9 (Normal) ☐
- Between 25 e 29.9 (Overweight) ☐
- > 30 Obesity ☐

Primary obesity: Yes ☐ No ☐ Other? _____

4. Insulin resistance – fasting glucose

- Between 100 – 125 mg/dL ☐
- Diabetes > 126 mg/dL ☐
- Glucose intolerance >140 mg/dL ☐

5. Pelvic ultrasound

Baseline ultrasound - Number of antral follicles (AFC)

- Right ovary (AFC) _____ Left ovary (AFC) _____
- Polycystic/ micropolycystic/ multifollicular ovaries ☐
- Blocked fallopian tubes ☐
- Small follicles ☐

6. Biochemical Profile - Hormonal levels

Name	Unit	Measurement	Reference values
Anti-Müllerian hormone (AMH)	ng/mL		
Estradiol (E2)	pmol/L		
Follicle-stimulating hormone (FSH)	IU/L		
Luteinizing hormone (LH)	IU/L		
Progesterone	nmol/L		
Prolactin	mU/L		
Testosterone	nmol/L		
17-Hydroxyprogesterone (OHP)	nmol/L		
Androstenedione	ng/mL		
Dehydroepiandrosterone sulfate (DHE-S)	µg/dL		
Dihydrotestosterone (DHT)	nmol/L		
Free Thyroxine (T4L)	ng/dL		
Free Triiodothyronine (T3L)	ng/dL		
Stimulating hormone (TSH)	mIU/mL		

---

7. Proposed Treatment

- Intrauterine insemination (IUA) ☐
- Conventional *in vitro* fertilization (IVF) ☐
- Intracytoplasmic sperm injection (ICSI) ☐

Cycle number: _____

8. Therapeutic regimen

Antagonist cycle ☐ Agonist cycle ☐

Drug: _____ Total dose of gonadotrophins: _____(mUI/ml)

Duration stimulation _____ (days)

Trigger: _____

Number of follicles in trigger:

Size	Right ovary	Left ovary
< 10 mm		
10 – 14 mm		
≥ 15 mm		

9. Genetic studies

FMR1 gene screening:

- Result: _____

Karyotype analysis:

- Result: _____

Other genetic test performed: _____