



Paulo Filipe Freitas Rodrigues

**Evaluation on the impact of the X chromosome inactivation
profile on phenotypic manifestation of inherited metabolic
diseases in female patients**

Dissertação do 2º Ciclo de Estudos Conducente ao Grau de Mestre em Toxicologia
Analítica, Clínica e Forense

Trabalho realizado sob a orientação do Técnico Superior de Saúde Especialista de
Genética Francisco Laranjeira do CGM e da Doutora Professora Helena Carmo

November 2020

Scientific communications

In the ambit of this research, the following scientific communication were made:

- XLVIII Conferências De Genética Dr Jacinto Magalhães. Porto (Portugal) 31 Jan - 01Fev 2020. Rodrigues, P., Caseiro, C., Ribeiro, I., Carmo, H., Quelhas, D., Laranjeira, F. Biochemical and molecular characterization of carriers of pathogenic variants in the GLA gene, the cause of Fabry disease.
- ESHG 2020.2 – Living in Your Living Room Virtual Conference 6 – 9 June 2020. Laranjeira, F. E., Rodrigues, P., Caseiro, C., Ribeiro, I., Carmo, H., Quelhas, D. Biochemical and molecular landscape of Fabry disease in Portugal.
- 16th International Symposium of the Portuguese Society for Metabolic Disorders (Online) 11th-13th November 2020. Laranjeira, F., Rodrigues, P., Caseiro, C., Ribeiro, I., Carmo, H., Quelhas, D. Biochemical and molecular characterization of carriers of pathogenic variants in the GLA gene, the cause of Fabry disease.

Resumo

A Doença de Fabry é uma doença progressiva associada ao cromossoma X que resulta de défice da atividade enzimática da alfa-galactosidase A causada por mutações no gene *GLA*. As mulheres podem apresentar vários fenótipos, desde o assintomático ao severo, passando por formas sintomáticas intermédias. Esta variabilidade é atribuída ao facto de as mulheres possuírem dois cromossomas X, ao contrário dos homens que são hemizigóticos. Os objetivos deste trabalho experimental são (i) verificar existência de correlação entre acumulação de substrato e atividade enzimática em doentes com Doença de Fabry de ambos os géneros, (ii) observar qual o biomarcador que melhor permite prever a gravidade do fenótipo correspondente, (iii) analisar o polimorfismo CAG no gene *AR*, para avaliar se existe um número diferente de repetições deste polimorfismo possibilitando a distinção dos dois cromossomas X, e (iv) investigar o padrão de inativação do cromossoma X entre os pacientes, especialmente em mulheres, usando o ensaio HUMARA para detetar inativações preferenciais do cromossoma X.

A investigação foi dividida em 4 partes: (i) a recolha de dados bioquímicos de 145 homens, e 222 mulheres diagnosticados com doença de Fabry, (ii) a análise de amostras de sangue total relativamente ao polimorfismo CAG no gene *AR*, através de técnicas de PCR, (iii) a análise molecular da inativação do cromossoma X através do ensaio HUMARA, e (iv) a avaliação da relação existente entre o desvio preferencial de inativação e o fenótipo bioquímico.

Como resultados principais verificou-se que existem 37 variantes do gene *GLA*, na população estudada, sendo que as mais prevalentes são: a p.F113L (49%), a p.R118C (8%), a p.M290I (3%), e a p.D313Y (8%). Observou-se que existe correlação entre acumulação de substrato e atividade enzimática entre géneros. Posteriormente, constatou-se que o biomarcador mais apropriado é o Lyso-Gb3, dado que faculta uma melhor análise discriminativa entre pacientes diagnosticados com a doença de Fabry. Relativamente ao padrão de inativação do cromossoma X, observou-se que este influencia a severidade da doença, uma vez que quanto maior é a expressão do alelo mutado, maior é a severidade da doença. Contudo, no sexo feminino existe alguma disparidade de valores relativos à expressão do alelo normal em relação à atividade enzimática.

Em conclusão, assume-se que a gravidade da Doença de Fabry não está apenas associada ao tipo de variante que cada paciente possui, mas também que a severidade do fenótipo se encontra fortemente associada ao fenómeno molecular da inativação do cromossoma X.

Palavras chaves: Doenças Herdadas Metabólicas, Doenças Lisossomais, Inativação do Cromossoma X, Doença de Fabry, mulheres

Abstract

Fabry disease is a progressive disease associated to the X chromosome that results in a reduced enzyme activity of alfa-galactosidase A caused by mutations in the *GLA* gene. Female patients may present several phenotypes, from asymptomatic, to mildly symptomatic and even the severe disease phenotype. All this variability is due to the fact that women have two X chromosomes, unlike men who are hemizygous, with only one X chromosome. The aims of this experimental research are (i) to verify the existence of a correlation between substrate storage and enzyme activity for both genders, (ii) to observe which biomarker can better predict the disease phenotype, (iii) to analyse the CAG polymorphism in the *AR* gene to evaluate if a different number of repetitions of this polymorphism exist and enables the distinction of two X chromosomes, and (iv) to investigate the pattern of X chromosome inactivation procedure among the patients, especially on female patients performed with the HUMARA assay, which consists in detecting non-random X chromosome inactivation.

The research was divided in 4 parts: (i) the collection of data from 145 male and 222 female patients diagnosed with Fabry disease, (ii) the analysis of the CAG polymorphism in the *AR* gene by using PCR techniques, (iii) the molecular analysis of X chromosome inactivation by HUMARA assay, and (iv) the evaluation of the relation between the preferential X chromosome inactivation and the biochemical phenotype.

As principal results it was verified the existence of 37 *GLA* gene variants in the studied Portuguese population, being the more prevalent the following: p.F113L (49%), p.R118C (8%), p.M290I (3%), and p.D313Y (8%). It was checked the correlation between the substrate accumulation and enzyme activity among genders. Subsequently, it was established that the proper biomarker compound to be used is Lyso-Gb3 since it enables a better discriminative analysis in patients with Fabry disease. As for the patterns of X chromosome inactivation it was observed that it influences the severity of the disease, because the more the mutated allele is expressed, the more severe the disease is. However, in female gender a diversity of the values of normal allele expressed linked to the enzymatic activity can be found.

In conclusion, it is possible to assume that the Fabry Disease is not only linked to which variant the patient carries, but also to the molecular phenomenon of the X chromosome inactivation.

Keywords: Inborn Errors of Metabolism, Lysosomal Storage Disorder, X chromosome inactivation, Fabry disease, female patients

Index

Scientific communications	ii
Index of figures	x
Index of tables.....	xi
List of abbreviations	xii
1. Introduction	1
1.1. General Aspects.....	2
1.2. Inborn Errors of Metabolism	3
1.3. Lysosomal Storage Disorders.....	4
1.4. X Chromosome Inactivation.....	7
1.5. Fabry disease and its presentation in female patients.....	11
1.6. Biochemical and molecular landscape of Fabry Disease in Portugal.....	17
1.7. Aims of this thesis	19
2. Materials and Methods	20
2.1. Patients data collection	21
2.2. HUMARA test.....	21
2.3. Data analysis and treatment	23
3. Results.....	24
3.1 Data analysis.....	25
3.1.1. Patients molecular characterization.....	25
3.1.2. Substrate storage and enzymatic activity relation	27
3.1.3. Biochemical profile of <i>GLA</i> gene variants among genders	30
3.1.4. XCI analysis.....	32
4. Discussion	36
4.1. Fabry disease in Portugal.....	37
4.2. Molecular characterization	37
4.3. Genotype/Biochemical phenotype and clinical phenotype	39
4.4. XCI pattern.....	41

5. Conclusions	42
References	45

Index of figures

Figure 1 Proposed hypothesis driven algorithm for testing males and females with suspected Fabry Disease [Adapted from 58]	14
Figure 2. GLA gene variations among male and female patients of CGM-JM and its frequency.	25
Figure 3. Data collected from CGM-JM database of patients diagnosed with FD and enzyme activity in plasma and leukocytes.	27
Figure 4. Data collected from CGM-JM database of patients diagnosed with FD and substrate accumulation in serum and urine.	28
Figure 5. Relation between enzyme activity and substrate accumulation from patients diagnosed with FD from CGM-JM database (a – α -Gal activity measured in Leukocytes related with Gb3 storage; b - α -Gal activity measured in Plasma related with Gb3 storage; c - α -Gal activity measured in Leukocytes related with Lyso-Gb3 storage; d - α -Gal activity measured in Plasma related with Lyso-Gb3 storage)	29
Figure 6. Analysis of the prevalent variants of GLA gene for each gender of patients diagnosed with FD from CGM-JM (a – patient with atypical activity in p.E203 variant; b – patient with atypical activity in p.F113L variant; c – Klinefelter)	30
Figure 7. Enzyme activity analysis in male patients and different variants analysed in this study (a – activity in Leukocytes; b – Leukocytes and Plasma activity in patients with p.S276N variant; c – activity in Plasma).....	31
Figure 8. Enzyme activity in female patients and different variants analysed in this study (a – activity in Leukocytes, b – activity in Plasma)	32
Figure 9. Example of results of first-tier PCR for evaluation of informativeness of the VNTR in AR: A – male patient; B – female with equal number of repeats (non-informative); C – female with different number of repeats	32
Figure 10. Example of determination of the amount of each inactive allele using peak areas of PCR products with: A - undigested template; B – restriction enzyme digested template	33
Figure 11. Analysis of the correlation (N=33 female patients) between normal allele expression and the enzymatic activity in leukocytes.	34
Figure 12. Analysis of the correlation (N=14 female patients) between normal allele expression and the Lyso-Gb3 storage in plasma.....	34

Index of tables

<i>Table 1 Lysosomal Storage Disorders [adapted from ¹⁹]</i>	6
Table 2. Transcription factors which promote or suppress the activity of Xist [Adapted from ⁵³]......	10
Table 3. Composition of tubes in digestion procedure.....	22
Table 4. 37 GLA gene variants found among patients	26

List of abbreviations

AFD	Anderson-Fabry disease
AR	Androgen receptor gene
ASSC	Active-site-specific chaperones
BGU	Biochemical Genetics Unity
CAG	Polymorphism CAG
CGM-JM	Centro de Genética Médica Jacinto Magalhães
CKD	Chronic kidney disease
DBS	Dried Blood Spot
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated degradation
ERT	Enzyme replacement therapy
FD	Fabry disease
G6PD	Glucose-6-phosphate dehydrogenase
Gb3	Globotriaosylceramide
gDNA	Quantity of DNA
GFR	Glomerular filtration rate
GLA	Alpha-galactosidase A gene
HCM	Hypertrophic Cardiomyopathy
HGMD	Human Gene Mutation Database
HPRT	Hypoxanthine phosphoribosyltransferase
HUMARA	Human Androgen Receptor Assay
IEM	Inborn Errors of Metabolism
lncRNA	Long-non coding RNA
LSD	Lysosomal Storage Disorders
LVH	Left ventricular hypertrophy
Lyso-Gb3	Globotriaosylsphingosine
MPS	Mucopolysaccharidosis
NHS	National Health System
OMIM	Online Mendelian Inheritance in Man
VNTR	Variable Number of Tandem Repeat
WML	White Matter Lesions
XCI	X chromosome inactivation

XCR	X chromosome reactivation
Xi	Inactive X chromosome
XIC	X inactivation centre
XIP	X chromosome inactivation preferential
Xist	X-inactive specific transcript
α -Gal	Alpha-galactosidase A acid hydrolase

1. Introduction

1.1. General Aspects

Genetics is a branch of Biology focused on the study of genes, the patterns of inheritance of characteristics to the following generations, and the diversity that exists between organisms¹. According to classic genetics, the specific sequence of nucleic acids is called gene, being recognized as the fundamental unity that carries the inherited information¹. The main achievement of this scientific area is to answer how some nucleotide variations support the diversity among individuals and can lead to problems in human health or the development of genetic disorders¹.

The genetic diseases can be classified according to the patterns of inheritance as (i) autosomal dominant, (ii) autosomal recessive, (iii) linked to the sexual chromosomes, and (iv) related to mitochondrial DNA^{2,3}. As it can be seen on the Online Mendelian Inheritance in Man (OMIM) database, 94,9% of the entries for genetic disorders are for autosomal anomalies, 4,8% to anomalies that are linked to the sexual chromosomes, and 0,3% to those diseases that are related to the mitochondrial DNA⁴. The diseases that are sex-linked, can be divided on which chromosome does the genetic defect occur, like (i) X-linked diseases, if it is on the X chromosome, or (ii) Y-linked diseases, if the mutation happens on the Y chromosome.

The human species, as well as the other mammals, has two genders with different chromosomal material: (i) male (XY) that has one X chromosome and one Y chromosome, and (ii) female (XX) that carry two X chromosomes⁵. Therefore, in the case of autosomal disorders they can be classified as dominant or recessive, but for the X-linked diseases these terms may not be appropriate to use^{6,7}. Much debate has been held about the X-linked diseases. Since males have only one X chromosome, if the variation exists on this chromosome, these individuals are going to express the disorder entirely, but if the individual is a female, they may not express the disease due to the existence of another X chromosome, being in that case, carriers of the variation². Therefore, it is proposed that women's phenotypic diversity is due to heterogeneity in the X chromosome inactivation, regardless of the dominant or recessive trait, and, consequently, it is considered inappropriate to use these terms^{6,7}.

1.2. Inborn Errors of Metabolism

In living organisms, metabolism maintains cellular activities through complex biochemical processes, in order to sustain life⁸. These processes are organized into metabolic pathways that are dependent on certain compounds and specific enzymes to ensure daily life activities⁸. However, sometimes an enzyme activity deficiency in a metabolic pathway brings up disorders that are called Inborn Errors of Metabolism (IEM)^{8,9}. IEMs are a group of disorders that individually can be rare, but collectively are considered common, with an incidence of more than 1:1000⁸⁻¹⁰. More than 500 IEMs have been recognized, and 25% of them have onset in the neonatal period⁸. Neonates are usually healthy at the time of birth with signs developing in hours to days after birth⁹. The common way of inheritance for the IEMs is an autosomal recessive manner, but there are also some presenting X-linked, autosomal dominant or mitochondrial inheritance. X-linked conditions prevail on affected males with severe disease whereas female patients are either asymptomatic or have only a mild disorder, such as in the cases of the ornithine transcarbamylase deficiency, Fabry's disease, and Danon's disease, that frequently have symptomatic females^{8,9}.

Errors like this can be defined in two types of clinical categories based on their symptoms and frequent signs, main affected organs or systems, or acuity or chronicity of presentation^{8,11}. Type 1 contains disorders that involve only one functional system, or affect only an organ or anatomic system with symptoms that are uniform, and its correct diagnosis is normally easy to accomplish¹¹, or category 2 of IEMs includes illnesses in which the biochemical defects either affects one metabolic pathway common to a large number of cells or organs¹¹. These diseases can be divided in three groups like including: (i) IEMs causing intoxication due to error in the intermediary metabolic pathway, resulting in the accumulation of toxic substrates, (ii) those that result in decreased energy and mitochondrial respiratory chain defects, and (iii) diseases resulting in defects in the formation or degradation of complex molecules in specific organelles and cofactors, like lysosomal storage disorders^{8,9,12,13}.

During childhood until adolescence, individuals that have IEMs have a large spectrum of clinical features from normal physical appearance to having distinctive physical abnormalities⁸. Neonates that present IEMs can reveal one or more of the following clinical manifestations as: (i) neurological manifestations when deterioration of consciousness is the common feature, (ii) hepatic manifestation in case of presenting hypoglycemia for example, (iii) cardiac manifestations when associated with cardiac

diseases, (iv) unusual urine odour, (v) diversity in facial characteristics, and (vi) *hydrops fetalis*, generally in case of lysosomal storage disorders⁹.

1.3. Lysosomal Storage Disorders

Lysosomes are intracellular hydrolytic organelles in eukaryotic cells that are responsible for the macromolecules degradation, recycling and signalling, that facilitate the translocation of small molecules originated from this catabolism, which is carried out by more than 50 acid-dependent hydrolases contained in the lysosomal lumen^{14–18}. Morphologically, lysosomes are heterogeneous organelles with membrane and low internal pH, but they share resemblances with other organelles on the endocytic and secretory pathways. Their organelle compartments are also different from other organelles because they are smaller in size, highly enriched with transmembrane proteins, hydrolytic enzymes (proteases, glycosidases, nucleases, phosphatases, and lipases) or other proteins important for disabling substrate accumulation and lysosomal storage^{15,17,19–21}. If any impair or defect on the function of these organelles occurs, it will cause the accumulation of undigested or partially degraded macromolecules, which can result in cellular damage, namely lysosomal storage disorders (LSD)^{14,16,22}.

Lysosomal disorders comprise a family of more than 50 metabolic diseases, most of which are inherited as autosomal recessive traits, but three of them are X-linked²². These illnesses that result from inherited gene mutations are characterized by the intralysosomal accumulation of undegraded substrates²³. Those disturbances usually affect lysosomal enzymes, but other non-enzymatic lysosomal proteins defects can also lead to abnormal substrate accumulation^{17,20,21,24–27}. All the LSD have in common the accumulation of specific molecules or compounds on the organelles of the endosomal-autophagic-lysosomal system^{17,25}. LSDs, as a group, are common, with an incidence of approximately 1 in 5000 live births to 1 in 5500 for those that involve lysosomal enzyme or integral membrane proteins defects^{22,28}. Despite this, these numbers are probably higher given the high rate of undiagnosed or misdiagnosed cases^{20,22–24,29}. In LSDs the symptoms may appear at variable ages, in some cases it can be develop in utero or during the newborn period or becoming quite evident in late adult ages. Although, these infirmities are progressive and evolve over time^{22,23}.

It is now acknowledged that these disorders are not only consequence of storage defects, but also a result from the disturbance of complex cell signalling mechanisms. Examples are: (i) activator proteins needed for sphingolipid hydrolysis, (ii) integral

membrane proteins, (iii) protective protein/cathepsin A affected in galactosialidosis, or (iv) modifier proteins like the α -formylglycine-generating enzyme showing low activity leading to multiple sulphatase deficiencies^{25,30–33}. Despite of the distinctive types of storage material in the different LSD's, they share many common biochemical, cellular, and clinical features. Thereby, it must be highlighted that in most lysosomal disorders one type of substrate accumulates and in some diseases, for distinct reasons, the accumulated compounds can be heterogeneous^{18,21}.

The defects in proteins involved in lysosome regulation or function promote the accumulation of non-digested molecules that can subsequently alter or activate pathogenic cascades including, lysosomal pH regulation, inflammation, oxidative stress, altered lipid trafficking, synaptic release, endocytosis, vesicle maturation, autophagy, exocytosis, and altered Ca^{2+} homeostasis^{20,21}. These type of diseases tend to be multisystemic and are always progressive, but the rate of progression and severity are variable^{19,24,25}. Due to the many steps in the synthesis and processing of lysosomal hydrolases, it is expected that the lysosomal activity can at some point become dysfunctional. A common feature of these diseases is the presence of residual bodies that are vacuoles containing undigested material, and are the hallmark of primary storage^{19,34}. Additionally, the accumulation of substrate with the enlargement of the affected cell, can result in the enlargement of the respective organ¹⁹. A minority of these events is caused by defects in lysosomal membrane proteins, and some lysosomal storage diseases are even caused by the deficiencies in non-lysosomal proteins either from the endoplasmic reticulum, the Golgi apparatus, or the endosomal pathway^{18,21,34}.

These disorders are usually named according to the major stored compound. Therefore, in the case of glycosaminoglycan substrates accumulation they are classified as mucopolysaccharidoses, those where lipid storage prevails, are called lipidoses, sphingolipidoses, or mucolipidoses (**Table 1**)^{18,21,24,25,34}.

Table 1 Lysosomal Storage Disorders [adapted from ¹⁹]

Lysosomal Storage Disorders	Substrate	Diseases
	Mucopolysaccharidoses (MPS)	MPS I, II, IIIA, IIIB, IIIC, IIID, IVA, IVB, VI, VII
	Glycoproteinoses	Aspartylglucosaminuria, Fucosidosis, α -Mannosidosis, β -Mannosidosis, Mucopolipidosis I, Schindler disease
	Sphingolipidoses	Fabry's disease, Farber's disease, Gaucher's disease, GM1 gangliosidosis, Tay-Sachs disease, Sandhoff's disease, Krabbe's disease, Metachromatic leukodystrophy, Niemann-Pick disease, types A and B
	Other lipidoses	Niemann-Pick disease type C, Wolman's disease, Neuronal ceroid lipofuscinosis
	Glycogen	Pompe's disease
	Multiple enzyme deficiency	Multiple sulphatase deficiency, Galactosialidosis, Mucopolipidosis II/III, IV
	Lysosomal transport defects	Cystinosis, Sialic acid storage disease
	Lysosomal proteins defects	Danon disease, Hyaluronidase deficiency

As already mentioned, three LSDs are considered X-linked diseases and, in these cases, males are going to express the disease entirely because they only carry one X chromosome. In the case of heterozygous females, they are usually considered asymptomatic, but they can also express the disorder fully, mainly due to disturbances in the inactivation of the X chromosome.

1.4. X Chromosome Inactivation

The X chromosome inactivation (XCI) is the mechanism where a dosage equivalence takes place between eutherian mammals, and occurs randomly in females resulting in an average of 50:50 with Gaussian distribution³⁵⁻⁴⁴. Assuming a Normal distribution of maternal against paternal X inactivation, a major percentage of females will show skewed XCI, which means that a preferential inactivation of the same X chromosome occurs³⁵. Males have only one copy of each sex chromosome, an X, which is a large, gene-rich chromosome, and a Y, which is considered a gene poor and heterochromatic, and females have two copies of the X chromosome^{5,44}. Despite this, the presence of one or two X chromosomes would make a twofold difference in the concentrations of many gene products between genders⁵. Some theories have tried to explain how this mechanism of inactivation works, but they haven't so far reached any conclusion to make any of these theories reliable. Therefore, it was concluded that this dosage equivalence has nothing to do with the adjustment between sexual chromosomes and autosomal genes, but the main reason of this equilibrium is the harmonization in sexual chromosomes between genders^{5,45}. This compensation happens due to the XCI that refers to the transcriptional silencing and the heterochromatic behaviour of one of the two X chromosomes in females at the beginning of the development⁴⁶.

In females, one of the two X chromosomes in every cell becomes inactive in an early phase of embryonic development. The inactive X chromosome has some special properties, including, late S phase DNA replication, condensation during interphase, hypoacetylation of histone 4, and in eutherian somatic cells, the cytosine residues on the inactive X chromosome (Xi) are differentially methylated^{41,47,48}, as observed in heterochromatin. Even though the Xi shares the same characteristics as the DNA sequence, the structure of euchromatin resembles that of heterochromatin^{47,49}. Heterochromatin is considered the inactive or facultative chromatin, and it is distinguished by the different staining features and by its time of replication. Usually, this type of chromatin is condensed and darkly stained during interphase and at mitotic prophase. Despite all these characteristics, it is still possible to classify heterochromatin as facultative or constitutive. The constitutive form remains in this state in virtually all types of cells and contains a large repetitive DNA. On the other hand, the facultative heterochromatin refers to the chromosome material that can turn into the euchromatic or heterochromatic form^{47,49}.

An important fact on random XCI is the fact of female being mosaics, which means that in a proportion of female somatic cells, the paternal X chromosome is active, while in the

rest of cells the maternal X chromosome is active. As a consequence, expression of different allelic variants present on the two X chromosomes exists, leading to a mosaic pattern⁵⁰. A study using transgenic mice was made, in which gene expression from one X chromosome was highlighted by a green fluorescent reporter gene, and the other X chromosome had a red fluorescent reporter gene, providing an outstanding image of this mosaicism in cellular resolution^{50,51}. The X inactivation centre (Xic) is a critical region where the process is initiated and from which it spreads in both directions along the chromosome, and evidence from X-autosome translocations and deletions gives the impression that the Xic is necessary for the initiation of X inactivation in early embryogenesis^{42,52}.

The process of X inactivation has been divided into three phases: i) initiation, ii) spreading, and iii) maintenance⁴⁴. The initiation phase involves the choice of which of the two X chromosomes is going to be inactivated and requires the presence of a special locus on the X chromosome, the previously mentioned Xic^{37,39,41,43,44}. This choice of which X chromosome is going to be inactivated is ensured by a counting mechanism mediated by a limited blocking signal that is produced in limited quantities, but also sufficient, and it is going to make a bound with the Xic of a single X chromosome in each diploid cell, with all other X chromosomes inactivated⁴¹. The spreading phase is *cis*-limited and cannot affect physically detached segments^{42,47}. It consists on the proceeding of the inactivation of Xic resulting in the *cis*-limited inactivation of several genes on the X chromosome⁴². Finally, the maintenance mechanisms ensure that the Xi is copied and transmitted throughout successive cell divisions^{39,44,48,53}. The inactivation procedure is a highly controlled, multi-layered epigenetic event, and defects in the maintenance phase of this process have been associated with cancer in animals and humans⁵⁰.

During the earliest stages of female embryogenesis, both the X chromosomes that were donated by the sperm and the egg are active, and this can be proved by biochemical studies of X-linked genes such as glucose-6-phosphate dehydrogenase (G6PD), hypoxanthine phosphoribosyltransferase (HPRT), or α -galactosidase^{48,54}. At this stage, the sign of inactivation of the X chromosome is recognized because of an abnormal replication of one of the two X chromosomes during the cell cycle. In male germ cells, the X chromosome becomes condensed and transcriptionally inactive before the beginning of meiosis. During the S phase, a late replication occurs and the association with its pairing partner, the Y chromosome, which also becomes inactive at this phase, in order to form the cytologically identifiable sex vesicle⁴⁸. The reason for the inactivation of the X chromosome occurring during the male meiosis is because it may avoid the initiation of damage in the recombination that might happen as a result from the presence of unpaired

locals on the single X chromosome^{48,55}. In females, although the inactive state is stable and heritable for a lifetime, in the germline the Xi is reactivated at the onset of meiosis, leading to the hypothesis that this is a consequence which is linked to the requirement of a euchromatic active state to a normal meiotic chromosome pairing⁴⁸.

The XCI is not a permanent status, it can be reversed in the embryogenesis of mice by X chromosome reactivation (XCR) which can happen in the pluripotent cells and the germ cell lineage^{50,56,57}. The XCI and its XCR are linked by differentiation and pluripotency, while the molecular criss-cross between X chromosome dosage compensation and the cellular differentiation state is recently being revealed⁵⁰. In the heart of XCI/XCR regulation lies Xic, that contains several non-coding RNA genes, where the expression is controlled by pluripotency factors^{50,56}. In Xic, the long RNA is called X-inactive specific transcript (*Xist*) which is expressed from the future Xi⁴⁴. This non-coding RNA with complex repeat patterns is a key molecule that initiates chromosome-wide silencing and a lot of investigation was made to find out how *Xist* is regulated⁴⁴. It was discovered that, like transcription factors, RNA regulators of *Xist* lie within the Xic and are divided into two groups: activators and repressors (**Table 2**). The *Xist* coats and silence *in cis*, being denominated as *cis*-limited^{41,46,50,53,58–60}. *Xist* begins the XCI process by recruiting regulators and by changing the structure of the X chromosome leading to the establishment of unique epigenetic coating of the Xi, which includes chromosome-wide enrichment for repressive marks like DNA and histone methylations of X-linked gene promoters. All these events combined enable the maintenance of XCI until it is reversed again⁵⁰. The relationship between *Xist* and Xi reactivation in human cells is not as clear as it is in the mouse, because (i) there is no imprinted XCI phase to be reversed in preimplantation development^{46,61}, (ii) due to the reactivation of the Xi in germ cells being hard to overlook because of the obstacles that exist when obtaining appropriate tissue samples and the low amount of human germ cell culture system that run over Xi reactivation⁴⁶, and (iii) the reprogramming of human somatic cells under standard conditions does not lead to Xi reactivation^{46,62}. In the studies aiming at silencing some genic regions, the facultative heterochromatin formation occurs, however, there is not a lot of information about how this heterochromatin is formed and how chromatin changes are involved in the gene silencing⁴⁶.

Table 2. Transcription factors which promote or suppress the activity of *Xist* [Adapted from⁵³]

Transcription factors			
Activators		Repressors	
<i>Jpx</i>	lncRNA activator	<i>Tsix</i>	Antisense <i>cis</i> -acting repressor
<i>RNF12</i>	<i>Trans</i> -acting activator	<i>Xite</i>	Enhancer of <i>Tsix</i> transcription
RepA	RNA transcribed within <i>Xist</i> Exon 1 and activates <i>Xist</i> transcription	Pluripotency factors	Via <i>Xist</i> Repression
			Via <i>Xist</i> Activation
			Via <i>Rnf12</i> Repression

In all mammals, an interesting fact about X chromosome dosage compensation is the use of long-non coding RNA (lncRNA) like *Jpx*, *Ftx*, *Tsix*, and *XCAT*, most of them are found in the XIC of the X chromosome. The factors *Jpx* and *Ftx* work as activators of *Xist* expression and regulate in a certain way the XCI. The role of *Jpx* and *Ftx* in regulating *Xist* expression in humans is based on mouse studies, but, unlike mice, humans Xic does not present the *Xist* antisense *Tsix*⁴⁶. This factor was the first identified negative regulator of *Xist* and is transcribed in the antisense orientation and represses *Xist* transcription *in cis*⁶³. In some cases, it is possible the existence of translocations or deletions of the X chromosome with an autosomal chromosome, and due to this the inactivation do not occur such as the deletion of the *Tsix* suppressor results in the non-straight choice of which X chromosome is going to be inactivated, and, sometimes, it is activated by another noncoding locus called *Xite*^{43,47,53,64}. Another place of *Xist* that is required for silencing genes is the Repeat A region, that creates a short transcript that is not dependent on *Xist*, which is called RepA and it is important because it can facilitate *Xist* transcription^{52,53,65}. One of the positive regulators of *Xist* is the lncRNA *Jpx* that acts as an RNA associated with chromatin or DNA elements unlike other factors of the Xic. Finally, another positive

factor of the *Xist* activity, is the *Rfn12* that is located upstream of the *Xist*, and it is not in the traditional region considered the *Xic*. However, the levels of the RFN12 protein have been noted to increase at the beginning of XCI^{53,66–68}.

The mechanism underlying the XCI has been hypothesized as an important factor for some X-linked genetic diseases, such as the Fabry's disease which will be one of the topics of this project³⁵.

1.5. Fabry disease and its presentation in female patients

Fabry's disease (FD) or Anderson-Fabry disease (AFD) is an X-linked progressive disorder caused by mutations in the alpha-galactosidase A gene (*GLA*) causing a rare inborn error of the lysosomal acid hydrolase α -galactosidase (α -Gal) production resulting in an absent or low enzyme activity^{35,69,78–85,70–77}. In consequence, the accumulation of glycolipids occurs, mainly of globotriaosylceramide (Gb3) and its deacylated form, globotriaosylsphingosine (Lyso-Gb3) and other sphingolipids, suffering lysosomal substrate accumulation which leads to a lysosomal storage disorder, and later to life-threatening complications^{69,70,82,86,71–74,76,77,79,80}. Lyso-Gb3 is a sensitive biomarker for FD that can be measured in plasma and Dried Blood Spots (DBS) by HPLC-MS/MS^{84,87}. The FD is a pan-ethnic condition with no racial or ethnic predilection and its incidence rate varies between 1:117.000 and 1:476.000 live new-borns, but this prevalence might be higher because of late or incorrect diagnosis^{85,88,89}. Due to the lack of knowledge by the professionals and the low specificity of the symptoms of this disease, it becomes difficult to perform a specific statistical analysis about the prevalence^{74,77,79}.

This disorder is an interesting spotlight across the spectrum of medicine because it affects various tissues, mainly the cardiac, endothelial, neuronal and renal tissues resulting in end-stage organ damage and failure^{70,84,85}. The disease is difficult to diagnose due to the clinical variability of different mutations, and the diversity in disease severity and symptom onset⁷⁰. Hemizygous males that carry severe mutations in the *GLA* gene have no residual α -Gal activity and develop classic FD with the primary symptoms during childhood, like pain crises, dysesthesia, gastrointestinal disturbances, angiokeratomas, and autonomic dysfunction^{35,81,86}. Meanwhile, as age increases, they can develop life-threatening conditions involving vital organs, including progressive renal failure, stroke and hypertrophic cardiomyopathy with myocardial fibrosis and arrhythmias^{35,78,81}. Due to the X-linked inheritance, heterozygote females can present different disease severities ranging from asymptomatic carriers to a severe phenotype similar to that of males^{35,82,85,90,91}. These female patients tend to have low as males, or higher residual

enzyme activity as they are mosaic, as already mentioned, due to the heterogeneity of XCI, and they can display clinical features similar to those of males with major organ involvement and associated mortality^{35,69,70,82,92}.

In the case of FD, there have been doubtful results regarding the X inactivation profiles of heterozygous females, because two studies found that X inactivation is a major factor in determining the clinical severity of heterozygous females^{35,82,93,94}. Another two studies reached the conclusion that XCI was random in heterozygous females for FD, ruling out the importance of an explanation to the phenotypic diversity^{35,95,96}. Usually, classic phenotype patients are the ones presenting the mutant cell line predominance and have very low residual enzyme activity, although due to the way that the enzyme works, there may be some reported activity ranging between 1%-3%⁷⁰. Some of the patients that present these features display early symptoms, typically in their childhood, proceeding to organ complications in their 20s or 30s. Late patients, mostly in their 50s, are females or males with non-classical mutations, that may present damage in a single organ^{70,97}.

In classic FD, symptoms such as chronic neuropathic pain, episodes of severe pain, skin abnormalities, gastrointestinal disturbances, and asymptomatic corneal opacity usually appear during infancy, and generally a few years later in girls than in boys^{26,79,98–100}. Despite this, approximately 60% of females with pathogenic variation in the *GLA* gene have normal α -Gal activity⁸⁴. Symptomatic organ problems, in general, occur in the young adult phase in both genders, including chronic kidney disease (CKD) evolution to renal failure and left ventricular hypertrophy (LVH) associated with myocardial fibrosis and arrhythmias, hearing loss, ischemic attacks, strokes, lung manifestations, and premature death in some cases^{75,78,106,107,79,80,100–105}. In patients that have the later-onset phenotype, typical cardiac symptoms are frequent, and low glomerular filtration rate (GFR) appears in the 5th decade of life, representing a delayed progression of the disease^{26,86,90,108–110}.

The diagnosis of FD is usually delayed by a minimum of 3 years, and usually by more than 15 years, after the beginning of signs and symptoms^{85,107}. This delay is attributed to the rarity of the disease and to the diversity and lack of specificity of the symptoms^{26,111,112}. The first presentation is usually performed by a family doctor, or dermatologist, ophthalmologist, paediatrician, geneticist, neurologist, cardiologist or nephrologist²⁶. After this, a clinical examination where appropriate investigation is performed, according to the organ system involved, is required²⁶. When the clinical examination shows suspicion of FD it is required to do specific exams such as biochemical, and genetic confirmation^{26,113}. In some cases, the diagnosis is performed based on the α -Gal activity in leukocytes, DBS, plasma, and urinary Gb3 levels^{114–117}. It is important to consider that in females the level of Gb3 is, usually lower than in males, and in some patients, it is possibly not higher due to

mutations in the *GLA* gene. Therefore, it is important to create an hypothesis or algorithm (**Figure 1**) to diagnose the disease^{77,118}. The phase of diagnostic confirmation should rely upon genetic analysis, and diagnosis of FD in an individual shall lead to the testing of other family members and the support of genetic counselling²⁶. In patients with renal, cardiovascular or cerebrovascular disease the screening will also identify the FD and help them to benefit from the therapy²⁶. Another aspect for screening, is through the end-stage renal damage, because previously undiagnosed FD has been detected in males and females that were submitted to haemodialysis^{119–123}. It was also suggested that screening kidney transplant for FD is effective. Screening of risk groups is made by measuring the activity of α -Gal enzyme, but it can lead to inconclusive results in detecting FD, especially in females^{124,125}. Screening for FD has been performed in other groups, such as patients with unexplained atherosclerosis or cornea verticillata¹¹⁴. The finding of cornea verticillate has increasingly helped in the diagnosis, due to the fact that it can be observed since childhood and even in patients that have normal enzymatic activity⁸⁵. Any screening requires a solid, reliable, and cost-effective method²⁶. The prenatal diagnosis is also possible by determining the α -Gal activity in chorionic villi at 10 weeks of pregnancy or in cultured amniotic cells at 14 weeks of gestation¹²⁶.

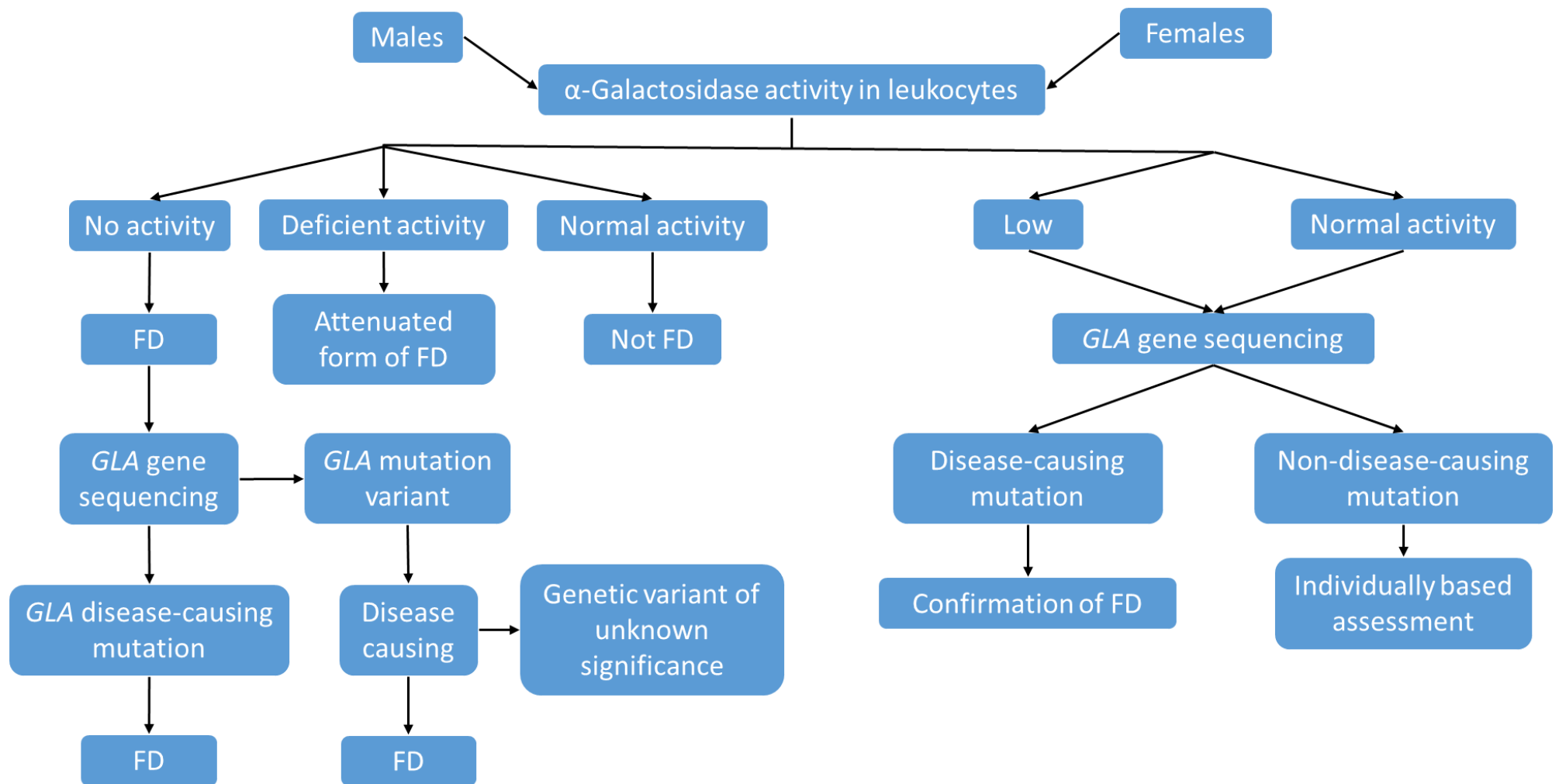


Figure 1 Proposed hypothesis driven algorithm for testing males and females with suspected Fabry Disease [Adapted from 58]

One of the most urgent needs is the investigation of reliable and validated biomarker(s) to assess the disease evolution and treatment response. Ideally, the measurement of such a marker would involve a non-invasive way of testing. Although various techniques have shown promising results, the clinical relevance for patients with FD has yet to be evaluated for its correlation with clinical endpoints¹⁰⁷. As previously mentioned, the current biomarkers that can monitor FD progression and response to the treatment include Gb3, and lyso-Gb3, measured in plasma, fibroblasts, and urine^{77,127}.

The treatment is extremely expensive (per patient and per year) making a careful diagnosis a matter of utmost importance, since in Portugal the therapy is fully supported by the National Health Services (NHS). Many alternative diagnostic methods such as increased glycosphingolipids in urine, or pathological evaluation of biopsied tissue from skin, heart, or kidney, are available, but the main diagnostic tools for FD is the deficiency of α -Gal activity in leukocytes or fibroblasts of male patients, or the detection of a mutation in *GLA* in female patients^{82,113}. Most of the literature emphasizes that enzyme replacement therapy (ERT) should be initiated upon confirmation of the diagnosis of FD, since an early ERT, conventional medical treatment and adjunctive therapy can pause the disorder progression and provide a better quality of life⁷⁷. The ERT appeared on the beginning of the 21st century, and it is based upon the recombinant human α -galactosidase A, that will replace the missing enzyme and catabolize the lipid deposit^{79,128,129}. The main lipid compounds such as Gb3, and lyso-Gb3 are used as biomarker of therapeutic efficacy^{85,130}. The recombinant human α -galactosidase A is commercially available by two pharmaceutical companies: (i) agalsidase alpha (Replagal®) and (ii) agalsidase beta (Fabrazyme®) which are administered to the patients by infusions throughout life^{77,81,82,113,131}. The Replagal® is produced from a stable transfected line of cultured human skin fibroblasts, and it is administered as an intravenous (iv) infusion at a dose of 0.2 mg/kg at every two weeks^{77,81,113,131,132}. As for the other form, the Fabrazyme® is obtained by the expression of human α -galactosidase DNA in Chinese hamster ovary cells, and it is administered as an (iv) infusion at a dose of 1.0 mg/kg, also at each two weeks^{77,81,113,131,133}. Despite this, the method for monitoring the therapeutic efficiency is limited due to Gb3 being a subclinical marker, which has levels at baseline that do not correlate with the disease severity, although changes in these levels serve as an indicator that the infused enzyme is active^{89,134}. Ideally, ERT should begin as soon as the symptoms of the disease are noted, because this therapy may slow down or prevent irreversible modifications in the cardiac or renal systems if the onset is at an earlier age, but in more severe cases the efficacy of the ERT is limited¹³⁵. In addition, patients shall receive treatment for the organs that are involved to control symptoms like: (i)

nephroprotection with the use of angiotensin-converting enzyme inhibitors, or angiotensin receptor blockers to control proteinuria, (ii) analgesics and opiates to pain control, (iii) avoid the use of anti-inflammatory agents, (iv) change of habits and lifestyle, including not smoking, reduction of sodium ingestion, and practice of physical activity^{85,136}. The positive response with this earlier therapy is supported by a study indicating that long-term ERT in young patients can make a complete Gb3 clearance of the glomerular endothelial cells of the kidney with dose dependent clearance of the renal podocyte inclusions^{77,137,138}. Clinical studies have shown a decrease in the frequency of pain episodes, cardiac mass, and storage of Gb3 in the skin, and kidneys, with even improvement of renal function in some patients, but there are still no studies that demonstrate lower values in mortality^{85,139}.

In FD, a significant number of disease-causing mutations are considered missense, which cause an unstable lysosomal protein, but still competent in catalysis¹⁰⁷. Studies of trafficking and deterioration of several mutant forms of α -Gal indicate that the mutant type of enzymes is gathered in the endoplasmic reticulum (ER) and degraded by endoplasmic reticulum-associated degradation (ERAD) because of their abnormal conformation¹⁴⁰. An intervention therapy for this is the use of small molecules active-site-specific chaperones (ASSC) to increase residual enzyme activity by reducing misfolding of the non-desired protein in order to avoid the premature degradation by ERAD^{85,141–144}. A drug acting as a chaperone is migalastat, which is administered orally, in order to avoid the need for infusions every 15 days^{85,144}. Despite this, this drug can only be used in 30-50% of patients with FD that have specific variations⁸⁵. Studies have been made where a first clinical trial on migalastat presented an effect comparable to that of ERT on renal and cardiac features^{85,144}. Another study did the analysis of the combination of the chaperone and ERT techniques and showed positive results, with an increase up to 1.2 and 5.1 times in enzymatic activity, when compared to ERT alone^{85,144}. An alternative method of treatment is substrate deprivation which is based not on the substitution of the defective glycosidase, but on the inhibition of an earlier step in the synthesis of the accumulating glycosphingolipid, decreasing the production of Gb3 and its accumulation¹⁴⁵. Finally, FD appears to be a disorder suitable for gene therapy^{85,146}. This technique consists in the addition of a normal gene of α -Gal to the DNA of the patient^{85,146}. This addition has the purpose to start producing a normal enzyme^{85,146}.

All the patients that manifest the FD should be followed by a metabolic specialist and a genetic counsellor, and after the diagnosis is confirmed the patients should have an annual check-up that includes (i) routine haematology, (ii) urinary protein and (iii) creatinine tests⁷⁷. In addition to this, it is recommended to do an electrocardiogram, echocardiography, and magnetic resonance of the brain, heart and kidney in order to

document the disease development and to verify if the treatment is going on a right course⁷⁷.

1.6. Biochemical and molecular landscape of Fabry Disease in Portugal

According to the Human Gene Mutation Database (HGMD) this particular disease has more than 1000 *GLA* gene variations reported ¹⁴⁷.

In the research of Pinto *et al.* (2004) the prevalence of LSD's in Portugal was studied and it was found that only 4 patients were diagnosed with FD. In the study by Azevedo *et al.* (2020) it was found that at least 21 families diagnosed with FD were carriers of the variant p.F113L, being considered the most prevalent variation in the Portuguese population.

The p.F113L is a pathogenic variant that results in a non-functional enzyme which is unstable on the ER, where the pH is neutral, and it will be degraded leading to a late-onset of FD^{140,148}. It is a variation mostly associated to hypertrophic cardiomyopathy (HCM) involvement, but its natural history is not fully known since in some cases renal involvement and sensorineural deafness were detected in the absence of a cardiac role^{149,150}. It was mentioned a high prevalence of FD among patients with HCM with a founder effect of the disorder due to this type of variant in Guimarães, Portugal¹⁴⁸. After analysis, such as Pedigree of this variant, results were obtained that led to the diagnosis of a large number of patients with this mutation, not only in this Portuguese region, but also in several parts of the world¹⁴⁸. The study by Azevedo *et al.* (2020) acknowledges for the first time the prevalence of each rhythm and conduction disease in a cardiac late-onset phenotype of FD, which means that cardiac disturbances have a fundamental impact on the prognosis of patients with the p.F113L mutation¹⁴⁸. Despite the cardiac features having the highest impact on the prognosis, the extracardiac manifestations like albuminuria, brain white matter lesions (WML) and sensorineural deafness were also as common as LVH¹⁴⁸.

In the study by Baptista *et al.* (2010) it was debated that family members who carry the variant p.R118C presented some features of FD, but no sign of storage accumulation in the kidney, creating the doubt among the researchers of its pathogenicity.

The p.R118C is a variation that produces a functional enzyme but with a lower activity than normal, which may occur due to an instability of the enzyme¹⁵¹. It is a missense mutation with a high prevalence in the Iberic Peninsula for which the possibility of a polymorphism was excluded since this variation was not found in 240 alleles from control

individuals^{151–153}. Despite the enzymatic activity encountered in patients being usually relatively high, there was a female patient that had reduced enzyme activity which raises the possibility of this variation being pathogenic¹⁵². In the research of Baptista et al. (2010) it was described that this variation has also been reported to have a similar behaviour to other *GLA* gene mutations associated with late-onset¹⁵⁴. The data of this experiment showed that in a Portuguese family a mild clinical phenotype was present, that included a female patient that manifested angiokeratoma and albuminuria¹⁵⁴. Other studies did not find histopathological evidence of FD with cardio or kidney involvement, and the observation of angiokeratoma was inconclusive due to being an isolated feature, which seems to be random in patients affected with other disorders^{153,155}. The variant p.R118C of the *GLA* gene was first mentioned to be pathogenic mainly due to theoretical criteria, but the mild enzymatic activity may not be enough to cause major signs of FD, which suggest a low level of pathogenicity and to some authors it is considered a neutral mutation^{153,156–158}. Despite the deliberation over the clinical severity associated to this variant, it is believed that it may cause specific organ manifestations, but the characteristic storage of Gb3 in relevant organs is not found^{154,156–158}. A recent research did the clinical testing in patients, in order to discard storage of Gb3 in kidneys associated with the p.R118C mutation¹⁵⁷. They found Gb3 deposits, structural changes on podocytes and renal tissues, that are typical features of FD, supporting that this mutation is pathogenic according to a pathological criterion¹⁵⁷. The study of this mutation and the p.D313Y variant show contradictory results in the literature, but the investigators believed that they are pathological mutations¹⁵⁶.

A similar behaviour occurs with the p.D313Y variant that was found in some patients that presented signs of FD^{151,159–161}. This variant, like the p.R118C, was described as non-pathogenic because it does not lead to severe organ manifestations like those observed in classic FD^{159,162,163}. This type of variant is also known as “pseudodeficiency allele” that is a sequence that produces an enzyme that is transported to the lysosomes, where it has a relatively reduced enzymatic activity and instability at neutral pH^{159,164,165}. Despite this, in several studies regarding this variant in patients with LVH, this mutation was considered either pathogenic or non-pathogenic^{159,166–168}. Both sides have low reliability which suggest that this variant is only mildly problematic depending on other factors such as the XCI¹⁶⁵. Regardless of this, none of these studies obtained histological evidences confirming or excluding the diagnosis of patients with cardiomyopathy, and the main issue if the p.D313Y is pathogenic still remains controversial¹⁵⁴.

1.7. Aims of this thesis

The X chromosome inactivation pattern is postulated as a determinant factor in the prediction of the FD phenotype. The purpose of this project is to investigate male and female patients that are carriers of pathogenic variants of an X chromosome-linked disease, the FD, by comparing a set of characteristic biochemical parameters among different phenotypes with the aim of finding a possible correlation between the genetic variants and the biochemical characteristics. Such an evaluation can guide therapeutic interventions in highly variable phenotypic group of patients, as in the case of female patients.

To achieve the purpose of this study, it is necessary to:

- i. Verify the existence of a correlation between substrate storage and enzyme activity for both genders;
- ii. Which biomarker compound gives more concrete information on phenotype severity;
- iii. Analyse the CAG polymorphism in the *Androgen Receptor (AR)* gene to evaluate if a different number of repetitions of this polymorphism enables the distinction of two X chromosomes;
- iv. The pattern of X chromosome inactivation procedure among the patients, especially on female patients.

2. Materials and Methods

This project was developed at Centro de Genética Médica Jacinto Magalhães (CGM-JM) in Porto, more precisely under the surveillance of the technical team of the Biochemical Genetics Unity (BGU). The investigation was divided in 4 parts: (i) patients data collection, (ii) analysis of the samples to check if the samples are informative to CAG polymorphism in the *AR* gene, (iii) molecular analysis of X chromosome inactivation by the HUMARA assay, and (iv) evaluation of the relation between preferential X chromosome inactivation (PXI) and biochemical phenotype.

2.1. Patients data collection

Information of the patients from the BGU database was collected, namely of data of features with interest to this project, including (i) *GLA* gene variants, (ii) gender of the patients, (iii) type of sample (DBS or whole blood), and (iv) biochemical parameters such as metabolites (Gb3 and Lyso-Gb3) accumulation and enzymatic activity in leukocytes or plasma. There were some patient's data that was not possible to obtain from the informatic database and it was necessary to verify the clinical files of those individuals. The first part of this work focused on the biochemical and molecular characterization of carriers of pathogenic variants in the *GLA* gene, searching for relationships among them.

2.2. HUMARA test

The Human Androgen Receptor Assay (HUMARA) was originally described as a sensitive means for detecting non-random X inactivation^{169,170}. This assay amplifies a highly polymorphic CAG repeat – variable number of tandem repeat (VNTR) - with 87% of heterogeneity in human *AR* gene at Xq12, a member of the superfamily of nuclear receptors of transcriptional factors that are activated by hormones^{171,172}. This region is sensitive to methylation¹⁶⁹ and when a X chromosome is active, the restriction sites of several endonucleases are not methylated and, therefore, are sensitive to enzymatic cleavage. In consequence, the fragmentation of this region's template avoids its amplification by Polymerase Chain Reaction (PCR)¹⁷³. On the other hand, when the X chromosome is inactive, these regions are methylated and due to this feature they resist to enzymatic digestion, and subsequently produce amplicons¹⁷³.

The first step was the evaluation of samples of all females to verify if they are informative to *AR* gene VNTR, that is, if the two alleles have different number of repeats so they can be distinguished. The DNA quantification of each sample was made using

NanoDrop® ND-1000 Spectrophotometer, and dilutions were prepared in order to achieve a concentration of 25.0 ng/μL for each sample (gDNA). The PCR amplification was performed in a final volume of 25.0 μL which contains 9.5 μL of H₂O, 12.5 μL of Promega PCR Master Mix (Taq DNA Polymerase, deoxynucleotide (dNTPS), magnesium chloride (MgCl₂) and reaction buffers), 1.0 μL of each HUMARA-1F and HUMARA-2R primers (10 pmol/μL), and 1.0 μL of gDNA, except for the negative control sample which is prepared with the same volume of H₂O. The PCR program was performed in a Veriti® Thermal Cycler and consisted of an initial denaturing step of 5 minutes at 95°C, followed by 27 cycles of 45 seconds at 95°C, 30 seconds at 65°C, and 30 seconds at 72°C, and a final extension of 15 minutes at 72°C. The amplicons were analysed by capilar electrophoresis in ABI 3130XL® Genetic Analyser and GeneMapper® software ([Applied Biosystems](#)).

Only those subjects with different number of repeats between alleles were selected for the next step, where the DNA templates were digested by *RsaI* and *HhaI* restriction enzymes prior to the amplifications by PCR.

Table 3. Composition of tubes in digestion procedure.

Reagents	Tube A	Tube B	Tube C
H ₂ O	19μL	17μL	13μL
CutSmart® buffer	3μL	3μL	3μL
gDNA (25ng/μL)	8μL	8μL	8μL
<i>RsaI</i>	-	2μL	2μL
<i>HhaI</i>	-	-	Add 4μL after 3h of incubation at 37°C just with <i>RsaI</i>

The incubation procedure was performed in Biometra® TGradient Thermocycler. The final PCR was developed in a volume of 25.0 μL which contained 0.5 μL of H₂O, 12.5 μL of Promega PCR Master Mix, 1.0 μL of HUMARA-1F and HUMARA-2R primers (10 pmol/μL respectively), and 10.0 μL of digested DNA, except for the control sample which is prepared with the same volume of H₂O. The PCR program was executed in a Veriti® Thermal Cycler and consisted of an initial denaturing step of 5 minutes at 95°C, followed by 27 cycles of 45 seconds at 95°C, 30 seconds at 65°C, and 30 seconds at 72°C, and a final extension of 15 minutes at 72°C. Once the amplification was complete, the samples were submitted to a capillary electrophoresis, and the results were further analysed resorting to GeneMapper.

After gathering all the results concerning to the double digestion of the alleles, the analysis of the XCI was done using the following **Equation 1**:

$$\frac{2nd\ digested\ allele\ area \times 100}{\frac{1st\ digested\ allele\ area \times 2nd\ non\ digested\ allele\ area}{1st\ non\ digested\ allele} + 2nd\ digested\ allele\ area}$$

Equation 1

This equation will allow us to calculate the extension of the methylation of the second allele (inactive), which means that this percentage corresponds to the percentage of the active first allele.

2.3. Data analysis and treatment

To perform the analysis, we have to take into account the existence of difference reference ranges for the biochemical parameters that must be evaluated. When focusing on substrate accumulation, the reference range for Lyso-Gb3 is from 0 to 0.5 nmol/l, and between 7 and 13 µg/mmol creatinine for Gb3 storage. In these parameters it is assumed the value of 100% of storage as the maximum value of the reference range. As for α-Gal activity, in leukocytes the reference values range from 36 to 80 nmol/h/mg protein, and for plasma from 6 to 19 nmol/h/ml plasma. Therefore, it was established that the value of 100% of the enzyme activity corresponds to the minimum value of the reference range. The values are all in the format of percentage, in order to harmonize the measured parameters and allow for comparisons between them.

In the dot plots, logarithmic scales were used because it provides a better distribution of the patients enabling to distinguish some clusters according to the variant that is expressed.

The box plot was made according to 4 values: (i) the minimum and maximum registered value of α-Gal activity, (ii) the 1st quartile where 25% of the measured values are less or equal to this percentage and 3rd quartile where 75% of the α-Gal activity values are until this range of percentage, (iii) another value that appear in the middle of the box is represented by the median.

3. Results

3.1 Data analysis

3.1.1. Patients molecular characterization

The data concerning the biochemical parameters for 145 male and 222 female patients were obtained. A correlation between these data and the molecular defect was evaluated. Thirty seven genetic variants of the *GLA* gene were found in the Portuguese population of this study with some of them being more prevalent as depicted in **Figure 2**.

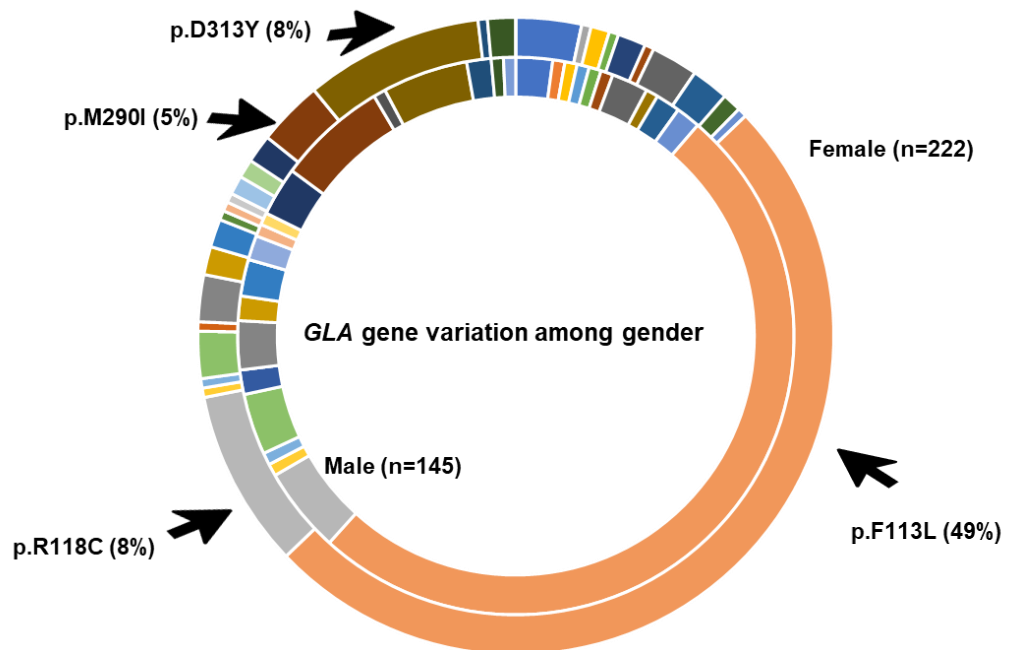


Figure 2. GLA gene variations among male and female patients of CGM-JM and its frequency.

According to this, it was verified that the prevalent variants among the Portuguese population investigated in the present study were in the majority missense types, like p.F113L, p.R118C, p.M290I, p.N215S and p.D313Y with cardiac and renal involvement. However, one of them stands out for being a non-sense variant, the p.E203* with the same organ manifestation as other variants. Among those which were more prevalent, one of them, the p.F113L stands out from the others with a prevalence around 50% of males and females. The variants are represented also in **Table 4**.

Table 4. 37 GLA gene variants found among patients

GLA gene variants	N° Male	N° Female	Type of mutation
c.104G>A (p.G35E) – exon 1	3	7	Missense
c.1067G>A (p.R356Q) – exon 6	1	0	Missense
c.1072_4del (p.E358del) - exon 7	0	1	Small deletion
c.443G>A (p.S148N) - exon 3	1	2	Missense
c.1176_9del (p.R392fs*1) - exon 7	1	0	Truncating
c.1229C>T (p.T410I) - exon 7	1	1	Missense
c.142G>A (p.E48K) - exon 1	0	3	Missense
c.154T>G (p.C52G) - exon 1	1	1	Missense
c.195-1G>A - intron 1	3	5	Splicing
c.274G>T (p.D92Y) - exon 2	1	0	Missense
c.281G>C (p.C94S) - exon 2	2	4	Missense
c.317_27del (p.L106fs*13) - exon 2	0	2	Truncating
c.335G>A (p.R112H) - exon 2	2	1	Missense
c.337T>C (p.F113L) - exon 2	71	105	Missense
c.352C>T (p.R118C) - exon 2	7	19	Missense
c.376A>G (p.S126G) - exon 3	1	1	Missense
c.386T>G (p.L129R) - exon 3	1	1	Missense
c.427G>A (p.A143T) - exon 3	5	5	Missense
c.457_9del (p.D153del) - exon 3	2	0	Small deletion
c.548G>T (p.G183V) - exon 4	0	1	Missense
c.607G>T (p.E203*) - exon 4	4	5	Truncating
c.644A>G (p.N215S) - exon 5	2	3	Missense
c.658C>T (p.R220*) - exon 5	3	3	Truncating
c.680G>A (p.R227Q) - exon 5	0	1	Missense
c.683A>G (p.N228S) - exon 5	2	0	Missense
c.716_8del (p.I239del) - exon 5	1	1	Small deletion
c.718_9del (p.K240fs*9) - exon 5	0	1	Truncating
c.769G>A (p.A257T) - exon 5	1	0	Missense
c.785G>T (p.W262L) - exon 5	0	2	Missense
c.794C>G (p.P265R) - exon 5	0	2	Missense
c.827G>A (p.S276N) - exon 6	4	3	Missense
c.870G>C/A (p.M290I) - exon 6	9	7	Missense
c.928C>T (p.L310F) - exon 6	1	0	Missense
c.937G>T (p.D313Y) - exon 6, pseudodeficiency.	7	19	Missense
IVS1_IVS4dup	2	1	Large Duplication

3.1.2. Substrate storage and enzymatic activity relation

It is known that the cause of FD is a defect of the function of α -Gal which raises an abnormal storage of typical compounds like Gb3 or Lyso-Gb3. Usually α -Gal activity is measured in leukocytes and plasma and substrates accumulation is measured in urine (Gb3) and serum (Lyso-Gb3), as some patients have no urinary excretion of these substances. The first step was to compare the enzyme activity in different sample types (plasma and leukocytes). These data were plotted as dot plot graphics in **Figure 3**.

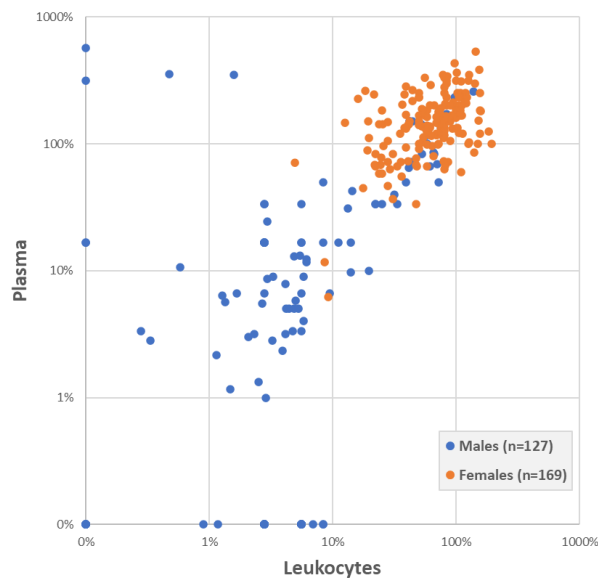


Figure 3. Data collected from CGM-JM database of patients diagnosed with FD and enzyme activity in plasma and leukocytes.

It is possible to observe that male patients diagnosed with FD tend to present low α -Gal activity in leukocytes and in plasma. A confusing event that can be visualized in this gender, is due to the fact that some patients present 0% of α -Gal function in leukocytes, but in plasma they are close to the normal values of the enzyme function. As for female patients, it is depicted that enzymatic activity ranges around the normal values (100% of enzyme function).

The data for the storage lipids, Lyso-Gb3 and Gb3, can be depicted in **Figure 4**.

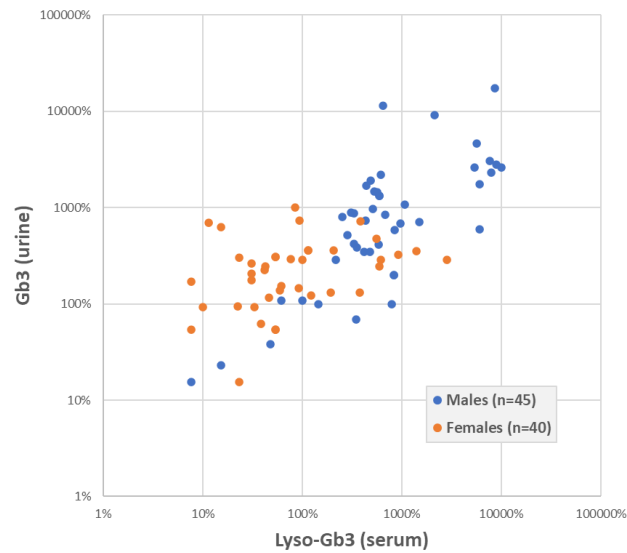


Figure 4. Data collected from CGM-JM database of patients diagnosed with FD and substrate accumulation in serum and urine.

For male patients, it is possible to observe that they usually have higher values of storage, for both Gb3 and Lyso-Gb3. In female patients, it is observed that they show a behaviour opposite to men, which means that they tend to have less substrate accumulation, near the normal higher levels. With this type of evaluation, it is possible to assume that a correlation exists between substrate accumulation and the activity of the enzyme. Therefore, an analysis was made with the accumulation of substrates in function of the α -Gal activity in serum and leukocytes as seen in **Figure 5**.

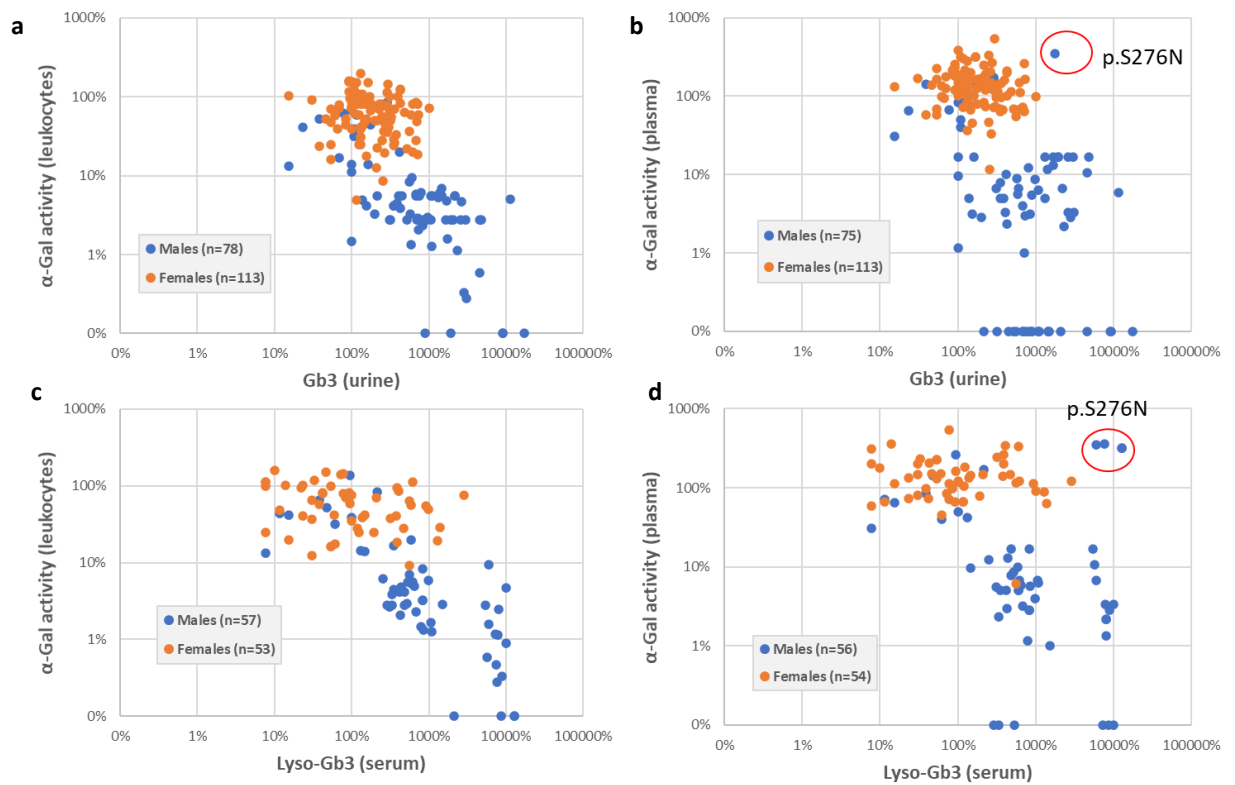


Figure 5. Relation between enzyme activity and substrate accumulation from patients diagnosed with FD from CGM-JM database (a – α -Gal activity measured in Leukocytes related with Gb3 storage; b - α -Gal activity measured in Plasma related with Gb3 storage; c - α -Gal activity measured in Leukocytes related with Lyso-Gb3 storage; d - α -Gal activity measured in Plasma related with Lyso-Gb3 storage)

It can be observed that the enzyme activity is usually slightly higher in plasma when compared to the leukocytes. It must be highlighted that for some male patients, no enzymatic activity was detected in plasma, which only occurs more rarely in leukocytes

In the analysis using the storage of Lyso-Gb3 as substrate parameter it was evident the existence of some distinctive clusters among male patients, which were not observed when using Gb3 as a parameter. As for female patients, using serum Lyso-Gb3 storage, a broader distribution is observed, when compared to urinary Gb3, but still it is not possible to distinguish subsets.

As it was mentioned before, recent studies have demonstrated that among the storage compounds, plasmatic Lyso-Gb3 is the best biomarker to be used. Using this compound in male patients, distinctive clusters can be visualized, which means that typical (low enzyme activity associated with substrate accumulation) variants of *GLA* gene likely exist, but, in other cases, there may be variants displaying atypical behaviour that does not follow the expected biochemical parameters. In female patients, as predicted, they present more diversity, which does not allow to depict clusters as in the case of the male individuals, but there are several female patients with a normal enzyme activity associated

with higher values of substrate accumulation. Therefore, it is reasonable to state that FD may not be determined only by the type of mutation, but also by another molecular phenomenon such as the XCI. Thus, with these interferences related to Gb3, it may be assumed that using the Lyso-Gb3 storage as a biomarker for FD conducts to a reliable analysis as it can be depicted in **Figure 5**.

3.1.3. Biochemical profile of *GLA* gene variants among genders

Afterwards, a detailed evaluation focusing on finding which variants were associated to each cluster previously identified was conducted, and the results are represented in **Figure 6**.

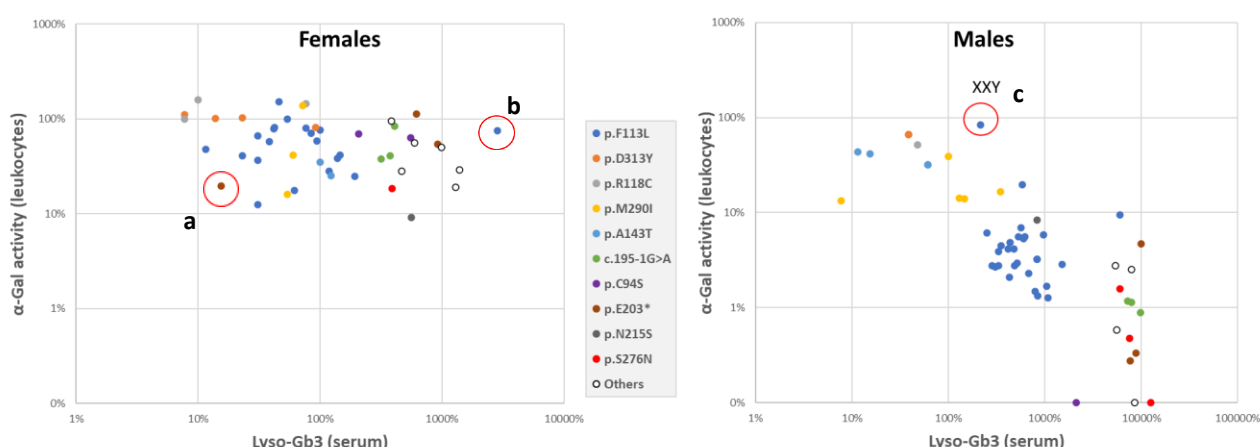


Figure 6. Analysis of the prevalent variants of *GLA* gene for each gender of patients diagnosed with FD from CGM-JM (a – patient with atypical activity in p.E203 variant; b – patient with atypical activity in p.F113L variant; c – Klinefelter)

It was observed that two categories can be found in male patients: (i) typical *GLA* gene variants such as p.F113L, p.E203*, p.S276N (in leukocytes), c.195-1G>A, p.N215S, and others, and (ii) atypical variants like p.R118C, p.M290I, p.D313Y, and p.A143T. The typical variants present the exact characteristics that are supposed to appear in patients with classical FD, which means higher values for substrate storage and low levels of enzyme function. The ones that are called atypical do not follow the pattern of a classical FD clinical board, because in those cases normal levels of substrate accumulation and nearly normal α -Gal activity can be found.

It is important to highlight, in the case of the male gender, that the C patient is a carrier of a typical variant, p.F113L, but it is not on the same cluster as the other patients, due to the fact that it is a Klinefelter syndrome case.

For female patients, as already stated heterogeneous results can be observed, but despite this, it is still possible to do the distinction of these categories according to the type of variants.

Patient B is an exception, having a much higher Lyso-Gb3 accumulation than the rest of the carriers of the p.F113L variant; on the other hand, patient A shows a behaviour opposite to the other p.E203* variant carriers, with very low Lyso-Gb3 accumulation.

In the previous plots, only patients for whom data were available concerning both biochemical parameters are represented. However, there are a large number of patients that do not fall in this situation. So, in order to get more representative data, the enzyme activity for the most prevalent variants, in both genders was also plotted as depicted in **Figure 7**.

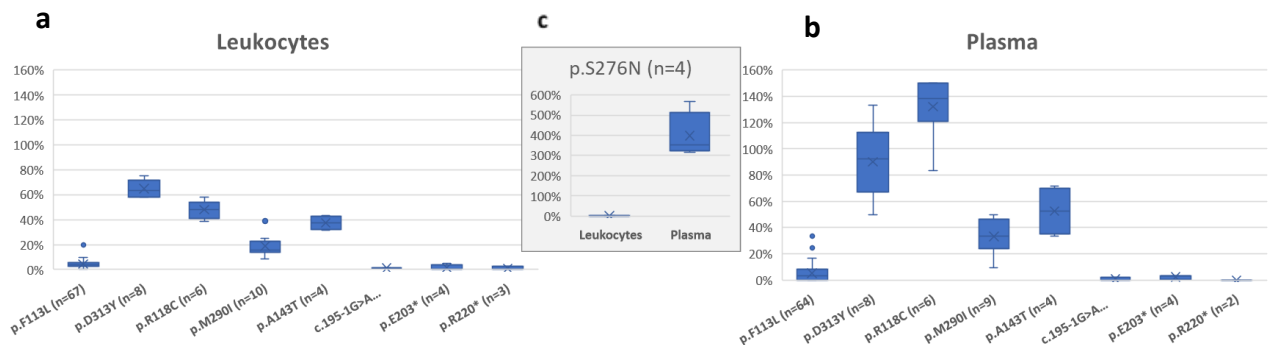


Figure 7. Enzyme activity analysis in male patients and different variants analysed in this study (a – activity in Leukocytes; b – Leukocytes and Plasma activity in patients with p.S276N variant; c – activity in Plasma)

As it is shown, in the case of enzymatic activity in leukocytes it is possible to observe that all *GLA* gene variants have reduced α -Gal function. In addition to this, atypical variant types can be distinguished, for which this biochemical feature can range from the normal to medium values, but still with low or reduced activity. As for the typical variants, as expected, a low or null enzymatic activity is observed. However, the enzyme activity in plasma allows to observe that the distribution of the variants is slightly higher when compared to the leukocytes, but still enabling to divide the variants into the different categories.

Another curious fact that can be observed is related to the p.S276N variant, that shows different enzyme activity in leukocytes, with low values indicating the classical FD phenotype, and normal values of the enzyme function in plasma, representative of healthy individuals.

Regarding female patients, the situation is not as clear as it was for the male patients, as it is possible to observe in **Figure 8**.

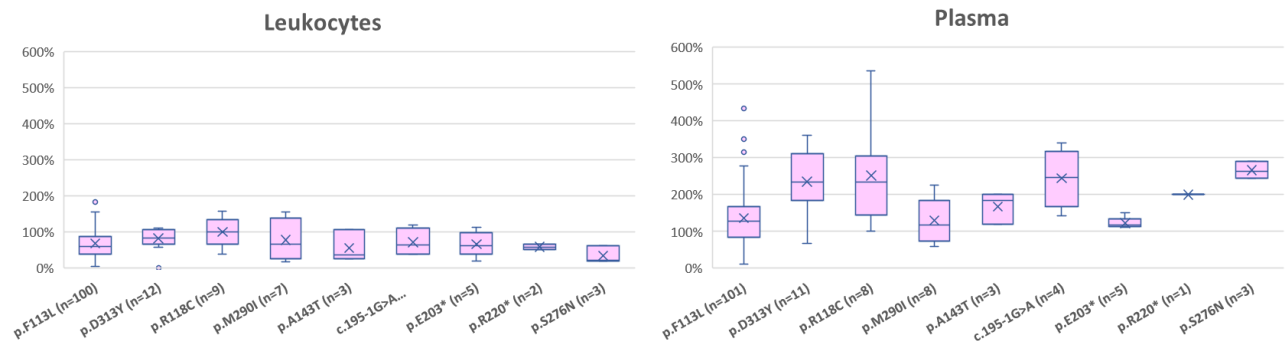


Figure 8. Enzyme activity in female patients and different variants analysed in this study (a – activity in Leukocytes, b – activity in Plasma)

In the case of female patients, it is not possible to see the behaviour of the different types of variants as in male individuals, because they are all overlapped and it is therefore impossible to efficiently separate the variants in distinct clusters. This aspect reinforces the idea that the female phenotype of the disease is not only influenced by the type of *GLA* gene variants, but it is also linked to the modulator effect of the XCI.

3.1.4. XCI analysis

Of the 53 female patients analysed only 47 were informative relative to the VNTR of the *AR* gene.

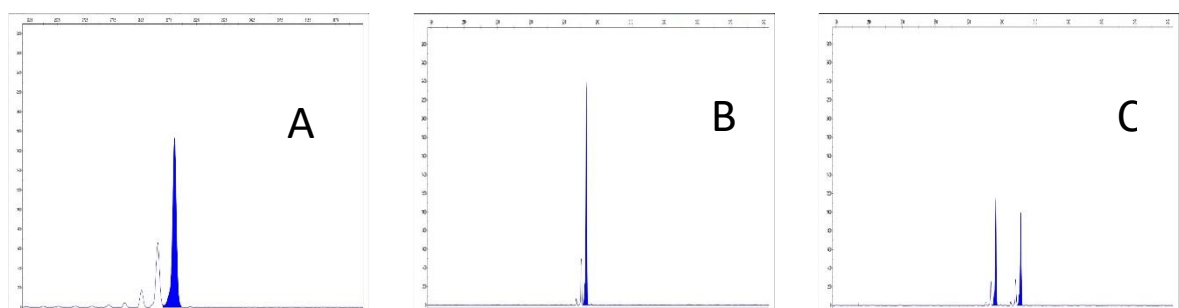


Figure 9. Example of results of first-tier PCR for evaluation of informativeness of the VNTR in *AR*: A – male patient; B – female with equal number of repeats (non-informative); C – female with different number of repeats

In order to eliminate the specific activity effect of each variant only female patients with the same variant, p.F113L, were analysed regarding the XCI study, therefore only 36 patients were considered. Three of them were non-informative so only 33 proceeded for the next phase.

Based on the data displayed in **Figure 10**, it was determined the percentage of the inactivation of each allele.

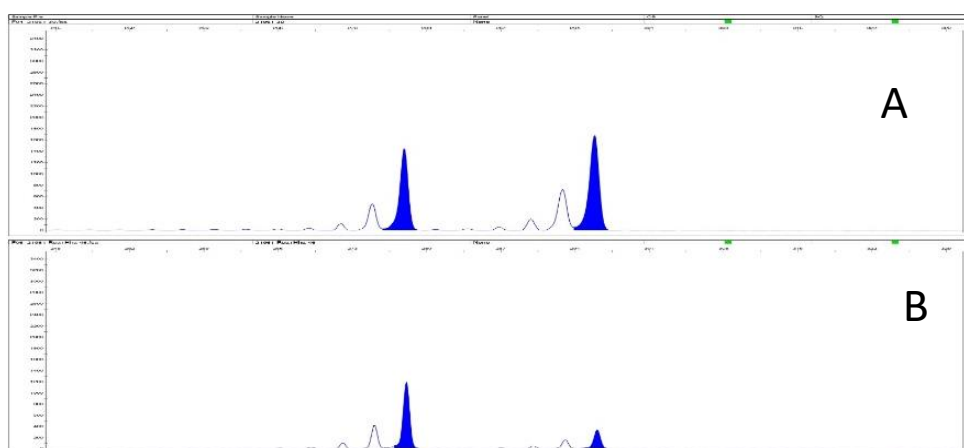


Figure 10. Example of determination of the amount of each inactive allele using peak areas of PCR products with: A - undigested template; B – restriction enzyme digested template

In the A example we have two distinctive alleles with a slight difference on the peak areas representing the initial PCR with the non-digested DNA. As for the B example, we have the final PCR where the DNA analysed suffered the digestion process and one allele that did not differ that much from the initial PCR can be distinguished, but there is another allele that suffered the action of the restriction enzyme. To accomplish the goal of this work it was necessary to figure out which allele was carrying the mutation and the easiest way to evaluate that is by studying a male relative. As this was not possible for all females of this set, conclusive results were achieved for only 14 female patients.

After these calculations, it was evaluated the relation of percentage of normal allele and enzymatic activity in leukocytes and Lyso-Gb3 substrate accumulation in plasma. Assuming that enzymatic activity is due, mostly, to normal enzyme molecules, derived from non-mutated allele, it was analysed the relation of leukocyte's enzyme activity with percentage of the normal allele expression, as it is possible to see in **Figure 11**. Therefore, as the enzymatic increases, the percentage of normal allele expressed will

increase as well. With this dot plot it is possible to infer that the inactivation pattern is associated with the enzymatic activity.

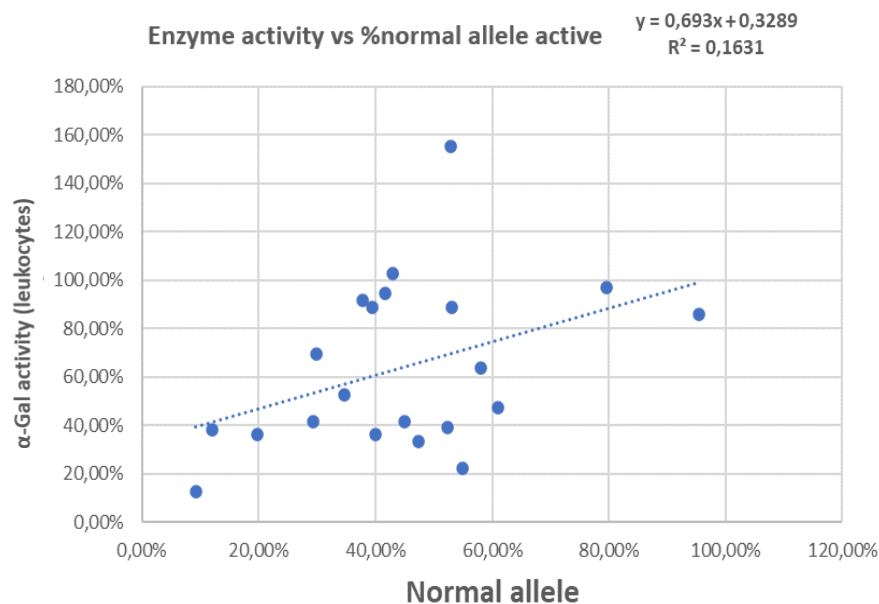


Figure 11. Analysis of the correlation (N=33 female patients) between normal allele expression and the enzymatic activity in leukocytes.

A similar analysis was made according to the substrate accumulation as it can be depicted in **Figure 12**.

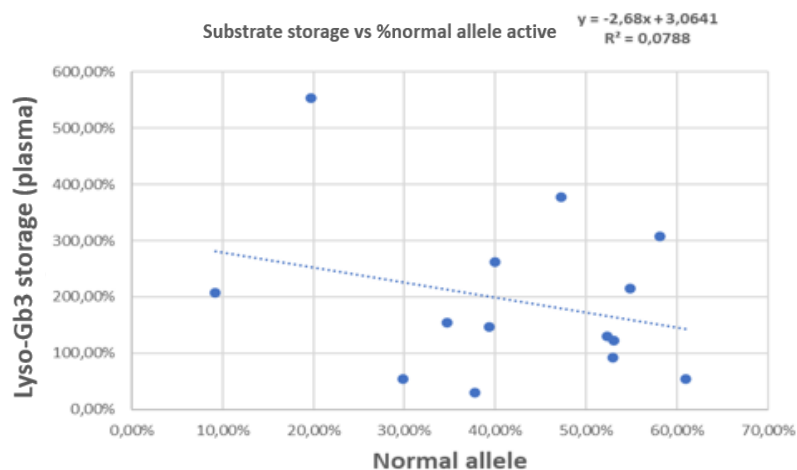


Figure 12. Analysis of the correlation (N=14 female patients) between normal allele expression and the Lyso-Gb3 storage in plasma.

It can be seen that the more the normal allele is expressed, the lower values of substrate accumulation occur.

4. Discussion

4.1. Fabry disease in Portugal

FD was not always known to be representative by several organ manifestations like cardio, or renal, or dermatological as a whole clinical board to be evaluated. In the study of Pinto, *et al.* (2004), only 4 patients with renal involvement were diagnosed with this disorder. However, FD is a more common disease in Portugal than previously acknowledged. In fact, in recent years increasing number of patients have been diagnosed as noted in the study by Baptista *et al.* (2010) that reported 125 patients with not only a compromised renal function but also showing signs of cardiomyopathy. In the recent study of Azevedo *et al.* (2020), approximately 360 patients carrying *GLA* gene variants with cardiac, or renal, or cerebrovascular symptoms were described.

Therefore, a FD diagnosis was initially based on the signs of renal disturbances with storage of specific compounds but now it is recognized that the clinical board of this disorder may not be associated only to the renal system, but also affects other organs. Thus, the requirements for correct diagnosis of these patients evolved and consequently the number of diagnosed patients with FD became higher.

4.2. Molecular characterization

According to HGMD more than 700 variants of *GLA* gene are missense mutations, which means that these type of *GLA* gene variations are more prevalent than other type of mutations

147.

In our results we have found 37 genetic variations, although two of them - c.870G>A and c.870G>C - generate the same amino acid exchange – p.M290I. Therefore, the data from patients carrying both these variants are analysed together and only 36 variants are represented.

Regarding the results obtained in this work, it was possible to check that according to the type of mutation it was obtained 67.5% of missense, 2.7% duplications and splicing and 8.1% of deletions and 13.5% truncating (frameshift or non-sense) according to **Table 4**. In **Figure 2**, it is observed that some variants are more prevalent than others such as: (i) p.F113L in 49% of both genders, (ii) 8% with p.R118C, (iii) 8% with p.D313Y, and (iv) 5% with p.M290I. Based on the database of mutations, it was expected that approximately 70% of the variants were missense mutations, and according to our results it was observed that 72% of the patients were carriers of missense variants of the *GLA* gene, as expected. The p.F113L is a particular variant that was extensively studied in Azevedo *et al.* (2020) where they conclude that in the Northern region of Portugal a founder effect

exists in the Portuguese population that are diagnosed with FD related to this variation, and this is in concordance with our findings¹⁴⁸.

It is important to reinforce that one of the major issues that has been more frequently subject of debate refers to the pathogenicity of some variants. The p.R118C and p.D313Y are *GLA* gene variants described as non-pathogenic because they do not lead to a severe organ manifestation as those observed with the classical FD genotype^{159,162,163}. According to the literature, the major issue is related to the biochemical parameters in these two variants, because they produce an enzyme with slightly reduced activity, which means that in cases of instability or pH variation, the enzyme does not develop its essential function, which is the deterioration of the compounds that will be deposited in the most affected organs^{151,159,164}. A new study has been presented where the p.R118C is considered a pathological Fabry variant¹⁷⁴. In that work, the authors describe 3 carriers of the p.R118C variant of the same family where one of them had elevated levels of Lyso-Gb3, and a cardiac biopsy showed FD-related changes in another member¹⁷⁴. In this family, the disease severity was mild but a clinically significant FD diagnosis was established due to the cardiac manifestation¹⁷⁴.

There is one variant that behaves in an atypical way, the p.A143T. This mutation is considered a variant of unknown meaning with its corresponding phenotype ranging from classical FD to a normal phenotype of unaffected patients with normal enzyme activity¹⁷⁵. It is a variant whose pathogenicity is debated, since some authors hypothesised that this mutation is possibly associated with an atypical late-onset cardiac phenotype of FD^{175–177}. Despite these features, some investigators debate that the association between cardiac symptoms and the p.A143T may be just coincidental and regard this mutation as non-pathogenic^{175,178}. In fact, in comparison with patients with classical FD, those individuals that present the p.A143T variant have normally a substantial residual α -Gal activity, have no increase in plasma Gb3 and Lyso-Gb3, and no deposits of Gb3 in the organs that are usually affected^{175,177}. Altogether, there is no direct evidence whether it is a pathogenic variant or not, due to the wide variability in the clinical presentation among different cases¹⁷⁵.

Cardiomyopathy is more common when male patients with later-onset present the p.N215S mutation, but in some cases renal failure also exists^{179,180}. This variant predictions for α -Gal activity are unclear, but analysis of the expression of the mutation showed a mild enzyme activity compared to the classical FD, being considered pathogenic and found to have more incidence in different groups of the population with late-onset FD¹⁸¹. Despite the isolated manifestations associated with this mutation, cardiac outcomes for patients resemble the classical patients, but some patients with the

p.N215S variant have shown more severe disease features¹⁸⁰.

The majority of variants are missense, but it is known that a few of them are nonsense, like the p.E203* for which the patients carrying the mutation develop classical FD¹⁸². In this type of mutation, the diagnosis of a severe phenotype of FD is supported by the observation that all male patients carriers have a substantially reduced or absent enzymatic activity and clinical features in young age¹⁸². In some research works, increased storage of Gb3 and Lyso-Gb3 were observed in all patients with this mutation, which is a finding consistent with a classical FD with later-onset, but the amount of deposits were higher in male than in female patients¹⁸². Interestingly, broad clinical spectrum is associated with p.E203* variant, as illustrated by two sisters

4.3. Genotype/Biochemical phenotype and clinical phenotype

With the analysis of the biochemical parameters of substrate accumulation and enzymatic activity (**Figure 3** and **Figure 4**) it is analysed firstly the α -Gal activity and it is possible to observe that a correlation exists between the measured values for both male and female patients. As anticipated, in theoretical terms, the Gb3 and Lyso-Gb3 storage tend to be higher in male patients than in female, and in contrast, the tendency regarding the α -Gal A activity is the opposite^{77,118}. Despite this, it is not possible to make a prognosis according to these biochemical parameters due to the diversity that exists in female patients, as it can be clearly observed in **Figure 5**. This heterogeneity among females with FD has been considered a challenge to clarify, because there are some investigators that support the idea of the preferential XCI of the mutated allele, and others that believe that this inactivation of the X chromosome follows a random pattern of selection⁹⁵. Female patients that are diagnosed with FD can be divided into three groups according to the degree of severity of the symptoms: asymptomatic carriers, mildly symptomatic carriers, and the ones who develop the disease classically like male patients^{35,82,85,90,91}.

A subject of much debate is which storage compound should be defined as a biomarker of FD⁸⁷. In the study by Togawa et al. (2010) it was concluded that Gb3, a major substrate of GLA, may not be the best biomarker of this disease. In this research the results showed that the Gb3 levels in plasma and urine were increased in hemizygous males with classical phenotype, but that male patients with variant FD and heterozygous female individuals do not have high levels of this major storage material. In female patients, both for the asymptomatic and for the symptomatic ones, the values of Lyso-Gb3 were mildly increased⁸⁷. With these features it was suggested that an elevated storage of Lyso-Gb3 is a potential biomarker of FD.

In our study, we revealed that the levels of Lyso-Gb3 are relatively higher in both genders than the Gb3 values among gender, which supports the conclusion of the Togawa et al. (2010) study. In our analysis, represented in **Figure 5**, it is possible to see the differences when we compared the distribution of the patients according to each type of substrate in relation to the enzyme activity. By using the Lyso-Gb3, it is easier to visualize the clusters that represent patients with different types of *GLA* variants, which helps in the analysis of which variant the patient can possess according to the values of the biochemical parameters. The Gb3 storage follows the patterns of the FD biochemical parameters, with high storage linked to a lower or reduced enzymatic activity, although it is not easier to distinguish the different variants or the visualization of the clusters that would help on the prediction of which type of variant the patient carries. Therefore, it represents an obstacle to the analysis which reinforces the idea that Gb3 is not the best biomarker to be used in the diagnosis of FD patients.

In **Figure 5**, the correlation that exists between the measured values of substrate accumulation with enzyme activity was analysed and there is a clear difference between genders. Male patients are hemizygous, therefore if they have the mutated allele that leads to the FD disorder, they will express the disease entirely, as commonly designated in the literature^{35,81,86}. It should be highlighted that men are in the same range of profiles with high storage and low or residual enzyme activity in the results, unlike women that have more variability according to these biochemical parameters as it was expected. With the ability to foresee which *GLA* gene variant exists, it was intended to check if the distribution of the patients is related or follows a pattern according to which *GLA* gene variant is present. In this way, the depicted clusters can help professionals during the diagnosis based on blood analysis that can support the prediction of what kind of variant they are dealing with.

With this analysis, as shown in **Figure 6**, it was possible to separate, in male patients, the *GLA* gene variants considering the following ranges: (i) from normal levels to 10-fold increase of substrate accumulation with low enzyme activity that are characteristic of the p.M290I and p.A143T variants, (ii) from normal levels to 10-fold increase of substrate and almost normal enzyme activity that are characteristic of the p.R118C and p.D313Y variants, (iii) from 10-fold to 80-fold increase of storage and low to residual enzyme activity that are characteristic of the most prevalent variant in this group, the p.F113L variant, and (iii) from 100-fold to 1000-fold increase of substrate accumulation but also low α -Gal activity that are seen in patients with the p.E203* and other variants significantly disturbing of the enzyme functioning. As for female patients, they share with male patients some distinction among variants but with different ranges and with larger variability within

groups.

Regarding the enzyme activity in serum or leukocytes based on the prevalent *GLA* gene variants (**Figure 6**) it was not possible to distinguish clusters in female patients, because they do not share the same distribution or same sort as male patients. Therefore, it is easier to verify the distribution of the variants in female gender in **Figure 7** and **Figure 8**. Even with this attempt, it was not possible to check these aspects on the female gender due to the overlapped values for the different *GLA* gene variants, which indicates that FD is not only linked to which variant the patient carries, but by another molecular process like XCI.

4.4. XCI pattern

In the analysis of the inactivation performance the data of enzymatic activity in leukocytes was used because this is the cell type analysed for the biochemical analysis and DNA extraction performed at the CGM-JM. The results showed that with increasing normal allele expression, the enzyme activity also increases, as expected. This feature indicates that the XCI is a phenomenon that modulates the severity of the FD, despite the R^2 being low, as it is possible to see in **Figure 11**. In **Figure 12** a similar graphic is created but using the substrate accumulation. This method is not as direct as with the enzymatic activity because the Gb3 or Lyso-Gb3 levels of storage are not measured at leukocytes exclusively. The storage compound data results from the sum of the entire tissues in the organism that excrete the substrate to the blood circulation. Therefore, using the α -Gal activity provides more reliable information about the inactivation pattern and its respective role in the severity of FD.

Thus, it has to be considered for future investigations a larger sample, and a statistical analysis of the data to bring more confidence to these results. Despite this, we are faced with the evidence that among the carriers of the p.F113L variant, although they have a common ancestor, there is still a lot of variability in haplotypes, which was found also in the *AR* gene, illustrating the occurrence of recombination in the X chromosome, that may change the genetic environment of the *GLA* gene in regard to modifying factors¹⁴⁸.

5. Conclusions

In this project it was possible to highlight some aspects relative to FD and its expression based on the biochemical parameters.

First, an up-to-date picture of FD genotypes in Portugal was obtained, as CGM-JM is the nation-wide reference laboratory for this disease. Second, it was possible to observe a general agreement between the biochemical markers of FD in different types of biological samples. According to our results, it is reasonable to affirm that Lyso-Gb3 is the most adequate compound to be used as a biomarker of this disorder because, while overcoming the limitation caused by the fact that some FD patients do not excrete Gb3 in urine, it shows a more mutation-dependent level, allowing for improvement in genotype-phenotype correlations.

Regarding to the *GLA* gene variants in the Portuguese population it is possible to divide them in two groups, regarding to their biochemical phenotype (i) the ones presenting typical phenotypes, with low enzyme activity and high substrate accumulation, associated with the classical FD phenotype, as p.F113L and p.E203*, and (ii) the atypical variants, with less pronounced effect on those parameters that were demonstrated to correspond to the p.M290I, p.R118C, p.A143T, and p.D313Y. Within this last group, the p.R118C and p.D313Y variants have been in the centre of controversy regarding its pathogenicity and are precisely the ones with a more moderate effect on biochemical parameters. Analysing the biochemical parameters, like substrate accumulation and enzyme activity respectively, it was obvious the differences that exist between female and male gender. For the male patients, as carriers of only one X chromosome, the clinical picture is characterized as classical FD phenotype with higher values of substrate storage linked to a reduced or null enzymatic activity. As for female patients, the clinical presentation is diverse, creating the 3 sub-group phenotypes with asymptomatic carriers, mildly symptomatic carriers and severe symptomatic carriers.

In what concerns to the main purpose of this investigation, our results strongly suggest that, although the biochemical phenotype in female FD patients is derived from the *GLA* gene variants, it is also modulated by the phenomenon of XCI. Even though being strongly associated with the inactivation procedure, among female patients there still is a lot of diversity with the correlation between the enzyme activity and the expression of the normal allele.

Taking into consideration the pandemic situation of COVID-19, the project suffered substantial interferences on its accomplishment, being impossible to achieve a reliable statistical result. Therefore, this study should be extended to a larger number of female patients, and an appropriated statistical analysis must be conducted to reinforce the results of this research. With this, in the case of finding new variants of the *GLA* gene, the

biochemical phenotype can be evaluated to verify in which cluster they appear. The cluster information will allow to postulate that those new variants that may appear in the future will have similar behaviours to those of the already known variants. All of this with the main goal of providing more accurate prognosis of patients and allow for better decision concerning treatment options.

References

1. Wilson, G. N. *Clinical Genetics*. (Wiley-Liss, 2000).
2. Trent, R. J. *Genes, environment and inheritance. Molecular Medicine* (2012). doi:10.1016/b978-0-12-381451-7.00002-5.
3. Levine, F. *Basic genetic principles. Fetal and Neonatal Physiology: Third Edition* vols 1–2 (Elsevier Inc., 2003).
4. Online Mendelian Inheritance in Man (OMIM). <https://www.ncbi.nlm.nih.gov/omim>.
5. Brockdorff, N. & Turner, B. M. Dosage compensation in mammals. *Cold Spring Harb. Perspect. Biol.* (2015).
6. Dobyns, W. B. *et al.* Inheritance of most X-linked traits is not dominant or recessive, just X-linked. *Am. J. Med. Genet.* **129A**, 136–143 (2004).
7. Antonarakis, S. E. & Cooper, D. N. *Human gene mutation in inherited disease. Emery and Rimoin's Principles and Practice of Medical Genetics* (Elsevier, 2013). doi:10.1016/b978-0-12-383834-6.00007-0.
8. Agana, M., Frueh, J., Kamboj, M., Patel, D. R. & Kanungo, S. Common metabolic disorder (inborn errors of metabolism) concerns in primary care practice. *Ann. Transl. Med.* **6**, 469–469 (2018).
9. El-Hattab, A. W. Inborn Errors of Metabolism. *Clin. Perinatol.* **42**, 413–439 (2015).
10. Campeau, P. M., Scriver, C. R. & Mitchell, J. J. A 25-year longitudinal analysis of treatment efficacy in inborn errors of metabolism. *Mol. Genet. Metab.* **95**, 11–16 (2008).
11. Saudubray, J. M. & Garcia-Cazorla, À. Inborn Errors of Metabolism Overview: Pathophysiology, Manifestations, Evaluation, and Management. *Pediatr. Clin. North Am.* **65**, 179–208 (2018).
12. Leonard, J. V. & Morris, A. A. M. Diagnosis and early management of inborn errors of metabolism presenting around the time of birth. *Acta Paediatr. Int. J. Paediatr.* **95**, 6–14 (2006).
13. Saudubray, J. M., Van Den Berghe, G. & Walter, J. H. *Inborn metabolic diseases diagnosis and treatment. Inborn Metabolic Diseases: Diagnosis and Treatment* (2012). doi:10.1007/978-3-642-15720-2.

14. Saftig, P. & Klumperman, J. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nat. Rev. Mol. Cell Biol.* **10**, 623–635 (2009).
15. Dell’Angelica, E. C., Mullins, C., Caplan, S. & Bonifacino, J. S. Lysosome-related organelles. *Encycl. Cell Biol.* **14**, 1265–1278 (2000).
16. Ruivo, R., Anne, C., Sagné, C. & Gasnier, B. Molecular and cellular basis of lysosomal transmembrane protein dysfunction. *Biochim. Biophys. Acta - Mol. Cell Res.* **1793**, 636–649 (2009).
17. Platt, F. M., Boland, B. & van der Spoel, A. C. Lysosomal storage disorders: The cellular impact of lysosomal dysfunction. *J. Cell Biol.* **199**, 723–734 (2012).
18. Ballabio, A. & Gieselmann, V. Lysosomal disorders: From storage to cellular damage. *Biochim. Biophys. Acta* **1793**, 684–696 (2009).
19. Pastores, G. M. & Hughes, D. A. Lysosomal Storage Disorders. *Encycl. Neurol. Sci.* 944–951 (2014) doi:10.1016/B978-0-12-385157-4.00096-8.
20. Schultz, M. L., Tecedor, L., Chang, M. & Davidson, B. L. Clarifying lysosomal storage diseases. *Trends Neurosci.* **34**, 401–410 (2011).
21. Vitner, E. B., Platt, F. M. & Futerman, A. H. Common and uncommon pathogenic cascades in Lysosomal Storage Diseases. *J. Biol. Chem.* **285**, 20423–20427 (2010).
22. Platt, F. M., D’Azzo, A., Davidson, B. L., Neufeld, E. F. & Tifft, C. J. Lysosomal storage diseases. *Nat. Rev. Dis. Prim.* **4**, (2018).
23. Parenti, G., Andria, G. & Ballabio, A. Lysosomal Storage Diseases: From Pathophysiology to Therapy. *Annu. Rev. Med.* **66**, 471–486 (2015).
24. Anderson, S. Newborn screening for lysosomal storage disorders. *J. Pediatr. Heal. Care* **32**, 285–294 (2018).
25. Kuech, E. M., Brogden, G. & Naim, H. Y. Alterations in membrane trafficking and pathophysiological implications in lysosomal storage disorders. *Biochimie* 152–162 (2016) doi:10.1016/j.biochi.2016.09.011.
26. Mehta, A. *et al.* Fabry disease: a review of current management strategies. *Q J Med* **103**, 641–659 (2010).
27. Rozenfeld, P. & Feriozzi, S. Contribution of inflammatory pathways to Fabry disease pathogenesis. *Mol. Genet. Metab.* **122**, 19–27 (2017).

28. Meikle, P. J. Prevalence of Lysosomal Storage Disorders. *J. Am. Med. Assoc.* **281**, 1–6 (1999).
29. Fuller, M., Meikle, P. J. & Hopwood, J. Epidemiology of lysosomal storage disorder diseases: an overview. *Oxford PharmaGenesis* <https://www.ncbi.nlm.nih.gov/books/NBK11603/> (2006).
30. Dierks, T. *et al.* Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C α -formylglycine generating enzyme. *Cell* **113**, 435–444 (2003).
31. Kolli, N. & Garman, S. C. Proteolytic activation of human cathepsin a. *J. Biol. Chem.* **289**, 11592–11600 (2014).
32. Nishino, I. *et al.* Primary LAMP-2 deficiency causes X-linked vacuolar cardiomyopathy and myopathy (Danon disease). *Nature* **406**, 906–910 (2000).
33. Futerman, A. H. & Van Meer, G. The cell biology of lysosomal storage disorders. *Nat. Rev. Mol. Cell Biol.* **5**, 554–565 (2004).
34. Kolter, T. & Sandhoff, K. Lysosomal degradation of membrane lipids. *FEBS Lett.* **584**, 1700–1712 (2010).
35. Echevarria, L. *et al.* X-chromosome inactivation in female patients with Fabry disease. *Clin. Genet.* **89**, 44–54 (2016).
36. Patel, S. *et al.* Human embryonic stem cells do not change their X inactivation status during differentiation. *Cell Rep.* **18**, 1–14 (2017).
37. Hwang, J. Y., Oh, J. N., Park, C. H., Lee, D. K. & Lee, C. K. Dosage compensation of X-chromosome inactivation center-linked genes in porcine preimplantation embryos: Non-chromosome-wide initiation of X-chromosome inactivation in blastocysts. *Mech. Dev.* **138**, 246–255 (2015).
38. Couldrey, C. *et al.* Bovine mammary gland X chromosome inactivation. *J. Dairy Sci.* **100**, 5491–5500 (2017).
39. Heard, E. Recent advances in X-chromosome inactivation research. *Curr. Opin. Cell Biol.* **16**, 825–832 (2004).
40. Migeon, B. R. Choosing the active X: The human version of X inactivation. *Trends Genet.* **33**, 899–909 (2017).
41. Avner, P. & Heard, E. X-chromosome inactivation: counting, choice and initiation.

- Nat. Rev. Genet.* **2**, 59–67 (2001).
42. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–137 (1996).
 43. Ogawa, Y. & Lee, J. T. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **11**, 731–743 (2003).
 44. Lu, Z., Carter, A. C. & Chang, H. Y. Mechanistic insights in X-chromosome inactivation. *Philos. Trans. B* **372**, (2017).
 45. Deng, X., Berletch, J. B., Nguyen, D. K. & Distèche, C. M. X chromosome regulation: diverse patterns in development tissues and disease. *Nat. Rev. Genet.* **15**, 367–378 (2014).
 46. Sahakyan, A., Yang, Y. & Plath, K. The role of Xist in X-chromosome dosage compensation. *Trends Cell Biol.* **28**, 999–1013 (2018).
 47. Lyon, M. F. X-Chromosome inactivation: a repeat hypothesis. *Cytogenet. Genome Res.* **80**, 133–137 (1998).
 48. Heard, E., Clerc, P. & Avner, P. X-chromosome inactivation in mammals. *Annu. Rev. Genet.* **31**, 571–610 (1997).
 49. Lyon, M. X-Chromosome inactivation and developmental patterns in mammals. *Biol. Rev.* **47**, 1–35 (1972).
 50. Payer, B. Developmental regulation of X-chromosome inactivation. *Semin. Cell Dev. Biol.* **56**, 88–99 (2016).
 51. Wu, H. *et al.* Cellular resolution maps of X chromosome inactivation: Implications for neural development, function, and disease. *Neuron* **81**, 103–119 (2014).
 52. Wutz, A., Rasmussen, T. P. & Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat. Genet.* **30**, 167–174 (2002).
 53. Jeon, Y., Sarma, K. & Lee, J. T. New and Xisting regulatory mechanisms of X chromosome inactivation. *Curr. Opin. Genet. Dev.* **22**, 62–71 (2012).
 54. Grant, S. G. & Chapman, V. M. Mechanisms of X-chromosome. *Annu. Rev. Genet.* **22**, 199–233 (1988).
 55. Jablonka, E. & Lamb, M. J. Meiotic pairing constraints and the activity of sex chromosomes. *J. Theor. Biol.* **133**, 23–36 (1988).

56. Payer, B. & Lee, J. T. Coupling of X-chromosome reactivation with the pluripotent stem cell state. *RNA Biol.* **11**, 798–807 (2014).
57. Pasque, V. & Plath, K. X chromosome reactivation in reprogramming and in development. *Curr. Opin. Cell Biol.* **37**, 75–83 (2015).
58. Brown, C. J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44 (1991).
59. Simon, M. D. *et al.* High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **000**, 1–5 (2013).
60. van Bemmelen, J. G., Mira-Bontenbal, H. & Gribnau, J. Cis- and trans-regulation in X inactivation. *Chromosoma* **125**, 41–50 (2016).
61. Okamoto, I. *et al.* Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* **472**, 370–374 (2011).
62. Tchieu, J. *et al.* Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell* **7**, 329–342 (2010).
63. Maclary, E., Hinten, M., Harris, C. & Kalantry, S. Long noncoding RNAs in the X-inactivation center. *Chromosom. Res.* **21**, 601–614 (2013).
64. Lee, J. T. & Lu, N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* **99**, 47–57 (1999).
65. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* (80-.). **322**, 1–11 (2008).
66. Jonkers, I. *et al.* RNF12 is an X-encoded dose-dependent activator of X chromosome inactivation. *Cell* **139**, 999–1011 (2009).
67. Navarro, P., Moffat, M., Mullin, N. P. & Chambers, I. The X-inactivation trans-activator Rnf12 is negatively regulated by pluripotency factors in embryonic stem cells. *Hum. Genet.* **130**, 255–264 (2011).
68. Barakat, T. S. *et al.* RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet.* **7**, 1–12 (2011).
69. Ortiz, A. *et al.* Fabry disease revisited: Management and treatment recommendations for adult patients. *Mol. Genet. Metab.* **123**, 416–427 (2018).
70. Cairns, T. *et al.* Hot topics in Fabry disease. *Postgrad. Med. J.* **94**, 709–713 (2018).

71. Lee, T. H. *et al.* Genomic screening of Fabry disease in young stroke patients: the Taiwan experience and a review of the literature. *Eur. J. Neurol.* **26**, 553–555 (2019).
72. Marian, A. J. Challenges in the diagnosis of Anderson-Fabry disease. *J. Am. Coll. Cardiol.* **68**, 1051–1053 (2016).
73. Olivera-González, S., Josa-Laorden, C. & Torralba-Cabeza, M. A. The pathophysiology of Fabry disease. *Rev. Clínica Española* **218**, 22–28 (2018).
74. Rodríguez, M. L. Treatment in Fabry disease. *Rev. Clínica Española (English Ed.* **218**, 489–495 (2018).
75. Esposito, R. *et al.* Prominent longitudinal strain reduction of left ventricular basal segments in treatment-naïve Anderson-Fabry disease patients. *Eur. Hear. J. - Cardiovasc. Imaging* **20**, 438–445 (2019).
76. Körver, S. *et al.* Predictors of objective cognitive impairment and subjective cognitive complaints in patients with Fabry disease. *Sci. Rep.* **9**, 1–12 (2019).
77. El-Abassi, R., Singhal, D. & England, J. D. Fabry's disease. *J. Neurol. Sci.* **344**, 5–19 (2014).
78. Moiseev, S. *et al.* The prevalence and clinical features of Fabry disease in hemodialysis patients: Russian nationwide Fabry dialysis screening program. *Nephron Clin. Pract.* **141**, 249–255 (2019).
79. Seydelmann, N., Wanner, C., Stork, S., Ertl, G. & Weidemann, F. Fabry disease and the heart. *Best Pract. Res. Clin. Endocrinol. Metab.* **29**, 195–204 (2015).
80. Waldek, S. & Feriozzi, S. Fabry nephropathy: a review - how can we optimize the management of Fabry nephropathy? *BMC Nephrol.* **15**, 1–13 (2014).
81. Biegstraaten, M. *et al.* Recommendations for initiation and cessation of enzyme replacement therapy in patients with Fabry disease: The European Fabry Working Group consensus document. *Orphanet J. Rare Dis.* **10**, (2015).
82. Germain, D. P. *et al.* The effect of enzyme replacement therapy on clinical outcomes in female patients with Fabry disease – A systematic literature review by a European panel of experts. *Mol. Genet. Metab.* **126**, 224–235 (2019).
83. Doheny, D. *et al.* Fabry Disease: Prevalence of affected males and heterozygotes with pathogenic GLA mutations identified by screening renal, cardiac and stroke

- clinics, 1995-2017. *J. Med. Genet.* **55**, 261–268 (2018).
84. Baydakova, G. V. *et al.* α -Galactosidase A/lysoGb3 ratio as a potential marker for Fabry disease in females. *Clin. Chim. Acta* **501**, 27–32 (2020).
 85. Bernardes, T. P., Foresto, R. D. & Kirsztajn, G. M. Fabry disease: genetics, pathology, and treatment. *Rev. Assoc. Med. Bras.* **66**, s10–s16 (2020).
 86. Schiffmann, R. *et al.* Fabry disease: progression of nephropathy, and prevalence of cardiac and cerebrovascular events before enzyme replacement therapy. *Nephrol. Dial. Transplant.* **24**, 2102–2111 (2009).
 87. Togawa, T. *et al.* Plasma globotriaosylsphingosine as a biomarker of Fabry disease. *Mol. Genet. Metab.* **100**, 257–261 (2010).
 88. Hsu, T. R. & Niu, D. M. Fabry disease: Review and experience during newborn screening. *Trends Cardiovasc. Med.* **28**, 274–281 (2017).
 89. Spada, M. *et al.* High incidence of later-onset Fabry disease revealed by newborn screening. *Am. J. Hum. Genet.* **79**, 31–40 (2006).
 90. MacDermot, K. D., Holmes, A. & Miners, A. H. Anderson-Fabry disease: clinical manifestations and impact of disease in a cohort of 60 obligate carrier females. *J. Med. Genet.* **38**, 769–75 (2001).
 91. Whybra, C. *et al.* Anderson-Fabry disease: Clinical manifestations of disease in female heterozygotes. *J. Inherit. Metab. Dis.* **24**, 715–724 (2001).
 92. Deegan, P. B. *et al.* Natural history of Fabry disease in females in the Fabry Outcome Survey. *J. Med. Genet.* **43**, 347–352 (2006).
 93. Dobrovolny, R. *et al.* Relationship between X-inactivation and clinical involvement in Fabry heterozygotes. Eleven novel mutations in the α -galactosidase a gene in the Czech and Slovak population. *J. Mol. Med.* **83**, 647–654 (2005).
 94. Morrone, A. *et al.* Inactivation analysis in manifesting carriers. 1–6 (2003).
 95. Maier, E. M. *et al.* Disease manifestations and X inactivation in heterozygous females with Fabry disease. *Acta Paediatr. Int. J. Paediatr.* **95**, 30–38 (2006).
 96. Elstein, D., Schachamrov, E., Beerli, R. & Altarescu, G. X-inactivation in Fabry disease. *Gene* **505**, 266–268 (2012).
 97. Arends, M. *et al.* Characterization of classical and nonclassical Fabry disease: A multicenter study. *J. Am. Soc. Nephrol.* **28**, 1–11 (2016).

98. Wilcox, W. R. *et al.* Females with Fabry disease frequently have major organ involvement: Lessons from the Fabry Registry. *Mol. Genet. Metab.* **93**, 112–128 (2008).
99. Hopkin, R. J. *et al.* Characterization of Fabry disease in 352 pediatric patients in the Fabry registry. *Pediatr. Res.* **64**, 550–555 (2008).
100. Kritzer, A., Siddharth, A., Leestma, K. & Bodamer, O. Early initiation of enzyme replacement therapy in classical Fabry disease normalizes biomarkers in clinically asymptomatic pediatric patients. *Mol. Genet. Metab. Reports* **21**, 100530 (2019).
101. Pitz, S. *et al.* Ocular signs correlate well with disease severity and genotype in Fabry disease. *PLoS One* **10**, 1–13 (2015).
102. Tuttolomondo, A. *et al.* Neurological complications of Anderson-Fabry disease. *Curr. Pharm. Des.* **19**, 6014–6030 (2013).
103. Morrissey, R. P., Philip, K. J. & Schwarz, E. R. Cardiac abnormalities in Anderson-Fabry disease and Fabry's cardiomyopathy. *Cardiovasc. J. Afr.* **22**, 38–44 (2011).
104. Zar-Kessler, C., Karaa, A., Sims, K. B., Clarke, V. & Kuo, B. Understanding the gastrointestinal manifestations of Fabry disease: promoting prompt diagnosis. *Therap. Adv. Gastroenterol.* 1–9 (2016) doi:10.1177/1756283X16642936.
105. Köping, M. *et al.* Characterization of vertigo and hearing loss in patients with Fabry disease. *Orphanet J. Rare Dis.* **13**, 1–9 (2018).
106. Svensson, C. K., Feldt-Rasmussen, U. & Backer, V. Fabry disease, respiratory symptoms, and airway limitation – a systematic review. *Eur. Clin. Respir. J.* **2**, 26721 (2015).
107. Germain, D. P. Fabry disease. *Orphanet J. Rare Dis.* **5**, 151–157 (2010).
108. Patel, V. *et al.* Clinical and genetic predictors of major cardiac events in patients with Anderson-Fabry Disease. *Heart* **0**, 1–6 (2015).
109. Germain, D. P. *et al.* Phenotypic characteristics of the p.Asn215Ser (p.N215S) GLA mutation in male and female patients with Fabry disease: A multicenter Fabry Registry study. *Mol. Genet. Genomic Med.* 1–12 (2018) doi:10.1002/mgg3.389.
110. Oder, D., Nordbeck, P. & Wanner, C. Long term treatment with enzyme replacement therapy in patients with Fabry disease. *Nephron* **134**, 30–36 (2016).
111. Ries, M. *et al.* The early clinical phenotype of Fabry disease: A study on 35

- European children and adolescents. *Eur. J. Pediatr.* **162**, 767–772 (2003).
112. Mehta, A. *et al.* Fabry disease defined: Baseline clinical manifestations of 366 patients in the Fabry Outcome Survey. *Eur. J. Clin. Invest.* **34**, 236–242 (2004).
 113. Linthorst, G. E. *et al.* Misdiagnosis of Fabry disease: importance of biochemical confirmation of clinical or pathological suspicion. *Br. J. Dermatol.* **150**, 575–577 (2004).
 114. Vedder, A. C. *et al.* The Dutch Fabry cohort: Diversity of clinical manifestations and Gb3 levels. *J. Inherit. Metab. Dis.* **30**, 68–78 (2007).
 115. Hoffmann, B., Koch, H. G., Schweitzer-Krantz, S., Wendel, U. & Mayatepek, E. Deficient α -galactosidase A activity in plasma but no Fabry disease - a pitfall in diagnosis. *Clin. Chem. Lab. Med.* **43**, 1276–1277 (2005).
 116. Mayes, J. S., Scheerer, J. B., Sifers, R. N. & Donaldson, M. L. Differential assay for lysosomal alpha-galactosidases in human tissues and its application to Fabry's disease. *Clin. Chim. Acta* **112**, 247–251 (1981).
 117. Mitsias, P. & Levine, S. R. Cerebrovascular complications of Fabry's disease. *Ann. Neurol.* **40**, 8–17 (1996).
 118. Young, E. *et al.* Is globotriaosylceramide a useful biomarker in Fabry disease? *Acta Paediatr. Int. J. Paediatr. Suppl.* **94**, 51–54 (2005).
 119. Nakao, S. *et al.* Fabry disease: Detection of undiagnosed hemodialysis patients and identification of a “renal variant” phenotype. *Kidney Int.* **64**, 801–807 (2003).
 120. Kotanko, P. *et al.* Results of a nationwide screening for Anderson-Fabry disease among dialysis patients. *J. Am. Soc. Nephrol.* **15**, 1323–1329 (2004).
 121. Dobrovolny, R. *et al.* A nationwide blood spot screening study for Fabry disease in the Czech Republic haemodialysis patient population. *Nephrol. Dial. Transplant.* **22**, 179–186 (2007).
 122. Schoenmakere, G. De *et al.* Two-tier approach for the detection of alpha-galactosidase A deficiency in kidney transplant recipients. *Nephrol. Dial. Transplant.* **23**, 4044–4048 (2008).
 123. Terryn, W. *et al.* Two-tier approach for the detection of alpha-galactosidase A deficiency in a predominantly female haemodialysis population. *Nephrol. Dial. Transplant.* **23**, 294–300 (2008).

124. Andrade, J. *et al.* Screening for Fabry disease in patients with Chronic Kidney Disease: Limitations of plasma α -galactosidase assay as a screening test. *Clin. J. Am. Soc. Nephrol.* **13**, 139–145 (2008).
125. Linthorst, G. E. *et al.* Screening for Fabry disease in high-risk populations : a systematic review. *J. Med. Genet.* **47**, 217–222 (2010).
126. Gal, A., Hughes, D. A. & Winchester, B. Toward a consensus in the laboratory diagnostics of Fabry disease - recommendations of a European expert group. *J. Inherit. Metab. Dis.* **34**, 509–514 (2011).
127. Ohashi, T., Iizuka, S., Ida, H. & Eto, Y. Reduced α -Gal A enzyme activity in Fabry fibroblast cells and Fabry mice tissues induced by serum from antibody positive patients with Fabry disease. *Mol. Genet. Metab.* **94**, 313–318 (2008).
128. Schiffmann, R. *et al.* Enzyme replacement therapy in Fabry disease: A randomized controlled trial. *JAMA* **285**, 2743–2749 (2001).
129. Eng, C. M. *et al.* Safety and efficacy of recombinant alpha-galactosidase replacement therapy in Fabry's disease. *N. Engl. J. Med.* **345**, 9–16 (2001).
130. Aerts, J. M. *et al.* Elevated globotriaosylsphingosine is a hallmark of Fabry disease. *Proc. Natl. Acad. Sci.* **105**, 2812–2817 (2008).
131. Beck, M. Therapy for lysosomal storage disorders. *IUBMB Life* **62**, 33–40 (2010).
132. Schiffmann, R., Ries, M., Timmons, M., Flaherty, J. T. & Brady, R. O. Long-term therapy with agalsidase alfa for Fabry disease: Safety and effects on renal function in a home infusion setting. *Nephrol. Dial. Transplant.* **21**, 345–354 (2006).
133. Wraith, J. E. *et al.* Safety and efficacy of enzyme replacement therapy with agalsidase beta: An international, open-label study in pediatric patients with Fabry disease. *J. Pediatr.* **152**, (2008).
134. Mehta, A. *et al.* Therapeutic goals in the treatment of Fabry disease. *Genet. Med.* **12**, 713–720 (2010).
135. Schaefer, R. M., Tylki-Szymanska, A. & Hilz, M. J. Enzyme replacement therapy for Fabry disease a systematic review of available evidence. *Drugs* **69**, 2179–2205 (2009).
136. Desnick, R. J. *et al.* Fabry disease, an under-recognized multisystemic disorder: expert recommendations for diagnosis, management, and enzyme replacement

- therapy. *Am. Coll. Physician-American Soc. Intern. Med.* **138**, 338–346 (2003).
137. Laney, D. A. *et al.* Fabry disease practice guidelines: Recommendations of the national society of genetic counselors. *J. Genet. Couns.* **22**, 555–564 (2013).
 138. Sessa, A. *et al.* Evolution of renal pathology in Fabry disease. *Acta Paediatr.* **443**, 6–8 (2003).
 139. El Dib, R. P., Nascimento, P. & Pastores, G. M. Enzyme replacement therapy for Anderson-Fabry disease. *Cochrane Database Syst. Rev.* (2016)
doi:10.1002/14651858.CD006663.pub3.
 140. Ishii, S. *et al.* Mutant α -galactosidase A enzymes identified in Fabry disease patients with residual enzyme activity: biochemical characterization and restoration of normal intracellular processing by 1-deoxygalactonojirimycin. *Biochem. Soc. J.* **406**, 285–295 (2007).
 141. Fan, J.-Q. & Ishii, S. Active-site-specific chaperone therapy for Fabry disease Yin and Yang of enzyme inhibitors. *FEBS J.* **274**, 4962–4971 (2007).
 142. Desnick, R. J. & Schuchman, E. H. Enzyme replacement and enhancement therapies: Lessons from lysosomal disorders. *Nat. Rev. Genet.* **3**, 954–966 (2002).
 143. Yam, G. H. F., Bosshard, N., Zuber, C., Steinmann, B. & Roth, J. Pharmacological chaperone corrects lysosomal storage in Fabry disease caused by trafficking-incompetent variants. *Am. J. Physiol. - Cell Physiol.* **290**, 1076–1083 (2006).
 144. Warnock, D. G. *et al.* Oral migalastat HCl leads to greater systemic exposure and tissue levels of active α -galactosidase A in fabry patients when co-administered with infused agalsidase. *PLoS One* **10**, 1–17 (2015).
 145. Abe, A. *et al.* Reduction of globotriaosylceramide in Fabry disease mice by substrate deprivation. *J. Clin. Invest.* **105**, 1563–1571 (2000).
 146. Sitaskas, C. & Medin, J. A. Gene therapy for Fabry disease. *J. Inherit. Metab. Dis.* **24**, 25–41 (2001).
 147. Human Gene Mutation Database. www.hgmd.cf.ac.uk.
 148. Azevedo, O. *et al.* Founder effect of Fabry disease due to p.F113L mutation: Clinical profile of a late-onset phenotype. *Mol. Genet. Metab.* **129**, 150–160 (2020).
 149. Eng, C. M. *et al.* Fabry Disease: Thirty-five mutations in the alpha-galactosidase A gene in patients with classic and variant phenotypes. *Mol. Med.* **3**, 174–182 (1997).

150. Burlina, A. P. *et al.* The pulvinar sign: frequency and clinical correlations in Fabry disease. *J. Neurol.* **255**, 738–744 (2008).
151. Gaspar, P. *et al.* Frequency of Fabry disease in male and female haemodialysis patients in Spain. *BMC Med. Genet.* **11**, (2010).
152. Caetano, F., Botelho, A., Mota, P., Silva, J. & Marques, A. L. Fabry disease presenting as apical left ventricular hypertrophy in a patient carrying the missense mutation R118C. *Rev. Port. Cardiol.* **33**, 183.e1-183.e5 (2014).
153. Ferreira, S. *et al.* The alpha-galactosidase A p.Arg118Cys variant does not cause a Fabry disease phenotype: Data from individual patients and family studies. *Mol. Genet. Metab.* **114**, 248–258 (2015).
154. Baptista, M. V. *et al.* Mutations of the GLA gene in young patients with stroke: The PORTYSTROKE Study-screening genetic conditions in PORTuguese Young STROKE Patients. *Stroke* **41**, 431–436 (2010).
155. Zampetti, A. *et al.* Angiokeratoma: Decision-making aid for the diagnosis of Fabry disease. *Br. J. Dermatol.* **166**, 712–720 (2012).
156. Chaves-Markman, Â. V. *et al.* GLA gene mutation in hypertrophic cardiomyopathy with a new variant description: Is it fabry's disease? *Arq. Bras. Cardiol.* **113**, 77–84 (2019).
157. Cerón-Rodríguez, M., Ramón-García, G., Barajas-Colón, E., Franco-Álvarez, I. & Salgado-Loza, J. L. Renal globotriaosylceramide deposits for Fabry disease linked to uncertain pathogenicity gene variant c.352C>T/p.Arg118Cys: A family study. *Mol. Genet. Genomic Med.* **7**, 1–9 (2019).
158. Arends, M. *et al.* Characterization of Classical and Nonclassical Fabry Disease: A Multicenter Study. *J. Am. Soc. Nephrol.* **28**, 1631–1641 (2017).
159. Koulousios, K. *et al.* Fabry disease due to D313Y and novel GLA mutations. *BMJ Open* **7**, 1–9 (2017).
160. Lukas, J. *et al.* Functional Characterisation of Alpha-Galactosidase A Mutations as a Basis for a New Classification System in Fabry Disease. *PLoS Genet.* **9**, (2013).
161. Lenders, M. *et al.* Multifocal White Matter Lesions Associated with the D313Y Mutation of the α -Galactosidase A Gene. *PLoS One* **8**, (2013).
162. Oder, D. *et al.* Organ manifestations and long-term outcome of Fabry disease in

- patients with the GLA haplotype D313Y. *BMJ Open* **6**, 1–14 (2016).
163. Niemann, M. *et al.* Lyso-Gb3 indicates that the Alpha-Galactosidase A mutation D313Y is not clinically relevant for Fabry disease. *JIMD Rep.* **7**, 99–102 (2013).
 164. Froissart, R., Guffon, N., Vanier, M. T., Desnick, R. J. & Maire, I. Fabry disease: D313Y is an α -galactosidase A sequence variant that causes pseudodeficient activity in plasma. *Mol. Genet. Metab.* **80**, 307–314 (2003).
 165. Hasholt, L. *et al.* The D313Y variant in the GLA gene—no evidence of a pathogenic role in Fabry disease. *Scand. J. Clin. Lab. Invest.* **77**, 617–621 (2017).
 166. Monserrat, L. *et al.* Prevalence of Fabry Disease in a Cohort of 508 Unrelated Patients With Hypertrophic Cardiomyopathy. **50**, (2007).
 167. Sachdev, B., Takenaka, T., Teraguchi, H. & Tei, C. Prevalence of Anderson-Fabry Disease in Male Patients With Late Onset Hypertrophic Cardiomyopathy. *Circulation* (2002) doi:10.1161/01.CIR.0000012626.81324.38.
 168. Morita, H. *et al.* Single-Gene Mutations and Increased Left Ventricular Wall The Framingham Heart Study. *Circulation* (2014) doi:10.1161/CIRCULATIONAHA.105.593558.
 169. Yorifuji, T. *et al.* PCR-based detection of mosaicism in Turner syndrome patients. *Hum. Genet.* **99**, 62–65 (1996).
 170. Edwards, A., Civitello, A., Hammond, H. A. & Caskey, C. T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**, 746–756 (1991).
 171. Callewaert, L. *et al.* Implications of a polyglutamine tract in the function of the human androgen receptor. *Biochem. Biophys. Res. Commun.* **306**, 46–52 (2003).
 172. Zhou, Z. X., Sar, M., Simental, J. A., Lane, M. V. & Wilson, E. M. A ligand-dependent bipartite nuclear targeting signal in the human androgen receptor. Requirement for the DNA-binding domain and modulation by NH₂-terminal and carboxyl-terminal sequences. *J. Biol. Chem.* **269**, 13115–13123 (1994).
 173. Araújo, A. & Ramos, E. S. Cryptic mosaicism involving a second chromosome X in patients with turner syndrome. *Brazilian J. Med. Biol. Res.* **41**, 368–372 (2008).
 174. Talbot, A. & Nicholls, K. Elevated Lyso-Gb3 suggests the R118C GLA mutation is a pathological Fabry variant. *JIMD Rep.* **45**, 95–98 (2019).

175. Hauth, L. *et al.* Galactosidase Alpha p. A143T Variant Fabry disease may result in a phenotype with multifocal microvascular cerebral involvement at a Young Age. *Front. Neurol.* **9**, 1–5 (2018).
176. De Brabander, I. *et al.* Phenotypical characterization of α -galactosidase A gene mutations identified in a large Fabry disease screening program in stroke in the young. *Clin. Neurol. Neurosurg.* **115**, 1088–1093 (2013).
177. Van Der Tol, L. *et al.* A systematic review on screening for fabry disease: Prevalence of individuals with genetic variants of unknown significance. *J. Med. Genet.* **51**, 1–9 (2014).
178. Lenders, M. *et al.* Alpha-Galactosidase A p.A143T, a non-Fabry disease-causing variant. *Orphanet J. Rare Dis.* **11**, 1–9 (2016).
179. Cecchi, F. *et al.* Intraoperative diagnosis of Anderson-Fabry disease in patients with obstructive hypertrophic cardiomyopathy undergoing surgical myectomy. *JAMA Cardiol.* **2**, 1147–1151 (2017).
180. Lavalle, L. *et al.* Phenotype and biochemical heterogeneity in late onset Fabry disease defined by N215S mutation. *PLoS One* **13**, 1–20 (2018).
181. Ebrahim, H. Y., Baker, R. J., Mehta, A. B. & Hughes, D. A. Functional analysis of variant lysosomal acid glycosidases of Anderson-Fabry and Pompe disease in a human embryonic kidney epithelial cell line (HEK 293 T). *J. Inherit. Metab. Dis.* **35**, 325–334 (2012).
182. Azevedo, O. *et al.* Mild left ventricular hypertrophy unravels a novel nonsense mutation of the GLA gene associated with the classical phenotype of fabry disease. *Cardiol.* **137**, 67–73 (2017).