U.PORTO

MSc IN LEGAL MEDICINE

# Loss of chromosome Y (*LOY*) in blood cells and its potential as a molecular biomarker of age-related diseases.

Ana Bela Coelho Barros

M
2019

U.PORTO

Ana Bela Coelho Barros.Loss of Chromosome Y (LOY) in blood cells and its potential as a molecular biomarker of age-related diseases.

Loss of Chromosome Y (*LOY*) in blood cells and its potential as a molecular biomarker of age-related diseases.
Ana Bela Coelho Barros

M.ICBAS 2019

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

Ana Bela Coelho Barros

# Loss of Chromosome Y (LOY) in blood cells and its potential as a molecular biomarker of age-related diseases

Dissertação de Candidatura ao Grau de Mestre em Medicina Legal submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientadora – Doutora Ana Luísa Teixeira
Categoria - Investigadora Júnior
Afiliação - Grupo de Oncologia Molecular e Patologia Viral do Centro de Investigação do Instituto Português de Oncologia do Porto FG
Coorientador – Professor Doutor Rui Manuel de Medeiros Melo Silva
Categoria - Professor Afiliado
Afiliação - Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto e Coordenador, Grupo de Oncologia Molecular e Patologia Viral do Centro de Investigação do Instituto Português de Oncologia do Porto FG
Coorientadora Laboratorial – Mestre Mariana Gomes Morais
Categoria - Aluna de Doutoramento
Afiliação - Grupo de Oncologia Molecular e Patologia Viral do Centro de Investigação do Instituto Português de Oncologia do Porto FG

"*We keep moving forward, opening new doors, and doing new things, because we're curious and curiosity keeps leading us down new paths*" by Walt Disney.

# TABLE OF CONTENTS

## AGRADECIMENTOS

Este trabalho foi realizado, entre Setembro 2018 e Setembro de 2019, no Instituto Português de Oncologia do Porto (IPO-Porto), com a colaboração do Instituto de Ciências Biomédicas Abel Salazar (ICBAS), da Universidade do Porto.

Começo por agradecer ao ICBAS pela oportunidade de realizar este Mestrado em Medicina Legal. Em particular, tenho de realçar a Professora Doutora Maria José Pinto da Costa pela disponibilidade e preocupação que demonstrou durante estes dois anos letivos. De igual forma, gostaria de agradecer ao grupo de investigação Oncologia Molecular e Patologia Viral do IPO-Porto por me receber na sua equipa e me ter integrado como membro da mesma.

Quero destacar a minha equipa de orientação, orientadora Doutora Ana Luísa Teixeira, coorientador, Professor Doutor Rui Medeiros e coorientadora laboratorial Mariana Morais pela orientação e ensinamentos que me passaram acerca das boas praticas de investigação. À Ana Luísa e Mariana tenho que agradecer a paciência que tiveram comigo, pois a minha teimosia dificultou a minha capacidade de saber ouvir. Conseguiram mostrar-me que eu valia a pena e que com trabalho árduo e, sobretudo, com mais atenção, tudo compensaria no final. Sem os vossos conselhos não teria evoluído como evolui este ano, tanto no campo profissional como no pessoal.

Também tenho de agradecer à Mara Fernandes, Bruna Cavalcante e Joana Pinto. Mara tu foste a maior surpresa no laboratório, sem ti não teria sido o mesmo, pois foste a primeira pessoa com quem criei uma ligação. És uma amiga que, com certeza, irei levar para a vida. Bruna, minha “parceira no crime”, sabes que começámos por caminhos tortuosos, mas a nossa amizade prevaleceu e, sem ti, com quem é que eu desabafava sobre a LOY? Ninguém do nosso círculo de amigos percebia a nossa dor! Obrigada por teres estado lá quando precisei. Joaninha sabes que poderás sempre contar comigo para te ajudar no que for preciso, partilhamos o mesmo tipo de personalidade, por esse motivo percebias sempre melhor o que se passava comigo, obrigada por isso!

Agora, sem esquecer o restante grupo de investigação, Francisca Dias - sei que te devo vários “bustos” de várias matérias primas – o meu agradecimento por me teres salvo a pele em vários momentos que precisei. Jani Silva, Augusto Nogueira, Ana Rita Pinto, Márcia Ferreira e Inês Nogueira vocês foram parte integrante da minha estada no IPO, muito obrigada por todas as vezes que tiraram “o peso das coisas” com um sorriso

ou por me fazerem rir. Obrigada por terem aceite o desafio, de me aturarem e por me terem aceitado como eu sou.

Aos meus pais, Fernanda Coelho e Bernardino Barros, o meu maior agradecimento, vocês são o meu grande pilar, sempre me apoiaram em tudo! Obrigada por permanecerem o meu porto seguro. Agradeço também à minha irmã, Felicidade Barros, pelos bons conselhos e por “embirrar” comigo por tudo e por nada. Gosto muito de ti. A próxima és tu, força nisso! Agradeço também o apoio das minhas avós, Maria Rocha e Olimpia Soares embora uma de vocês tenha partido a meio do meu mestrado. Ambas estarão sempre no meu coração.

Amigos são poucos nas horas que mais precisamos. Eu tive a sorte de ter a Juliana Ferreira, Angélica Pinto, Gustavo Sampaio, André Marques, Pedro Moreira e Daniel Silva. A todos vocês obrigada por tudo!

Não posso deixar de agradecer à família do meu namorado que me acolheu durante o segundo ano do mestrado e me apoiaram. Por último, mas dos mais importantes, Pedro Castro, o meu namorado. A ti eu devo o mundo! Foste a pessoa que mais me apoiou, repreendeu e me acordou em momentos maus pelos quais passei. Sempre acreditaste em mim, aceitando que sou uma tagarela, “colas” e que sou capaz de muito mais do que mostro, a ti agradeço com o todo o meu amor.

Finalmente, a mim própria, dedico esta tese que simboliza uma evolução pessoal e profissional, com certezas de que ainda estou no início do caminho do conhecimento e que ainda terei muito que aprender.

# LIST OF ABBREVIATIONS AND ACRONYMS

**A**

| | |
|---|---|
| AFAR | American Federation for Aging Research |
| AIDS | Acquired Immunodeficiency Syndrome |
| Alu | Arthrobacter luteus restriction endonuclease derived short DNA stretch |
| AMELX | Amelogenin, X-linked |
| AMELY | Amelogenin, Y-linked |
| AML | Acute Myeloid Leukaemia |
| AR | Androgen receptor |
| AVC | Acidente Vascular Cerebral |
| AZF | Azoospermia Factor region |

**B**

| | |
|---|---|
| BMI | Body mass index |
| bp | Base pars |
| BPH | Benign Prostate Hyperplasia |
| BPY | Basic charge protein, Y-linked |

**C**

| | |
|---|---|
| CDY | Chromodomain, Y-Linked |

**D**

| | |
|---|---|
| DAZ | Deleted in azoospermia |
| DBY | Aliase to DEAD-Box Helicase 3, Y-Linked (DDX3Y) |
| DDX3Y | DEAD-Box Helicase 3, Y-Linked |
| DFFRY | Deubiquitinating Enzyme FAF-Y |
| DNA | Deoxyribonucleic acid |

**E**

| | |
|---|---|
| EDTA | Ethylenediamine tetraacetic acid |
| EIF1AY | Eukaryotic translation initiation factor 1A, Y-chromosomal |

**F**

| | |
|---|---|
| FISH | Fluorescence in situ hybridization |

**H**

| | |
|---|---|
| HINd III | Polymorphism generated by restriction enzyme produced by *Haemophilus influenzae* |
| HMG-BOX | High mobility group Box |
| HSFY | Heat shock transcription factor, Y-linked |

**I**

| | |
|---|---|
| IR | Inverted repeats |

**K**

| | |
|---|---|
| Kb | Kilo base pairs |
| KDM5D | Lysine-specific demethylase 5D |

**L**

| | |
|---|---|
| LACI | Lacunar Circulation Infarct |
| LINE | Long interspersed nuclear elements |
| LOY | Loss of chromosome Y |

**M**

| | |
|---|---|
| Mb | Mega bases pairs |
| mRNA | Messenger ribonucleic acid |
| MSY | Male-specific region of the Y chromosome |

**N**

| | |
|---|---|
| NLGN4Y | Neuroligin 4, Y-Linked |
| NORF | No long open reading frame |
| NRY | Non-recombinant region Y |

**O**

| | |
|---|---|
| OCSP | Oxfordshire Community Stroke Project |

**P**

| | |
|---|---|
| PACI | Partial Anterior Circulation Infarct |
| PARs | Pseudoautosomal regions |
| PC | Prostate cancer |
| PC-3 | Prostate cancer human cell line – 3 |
| PCDH11Y | Protocadherin-11, Y-Linked |
| $PM_{10}$ | Particulate matter inferior to 10µm of diameter |
| POCI | Posterior Circulation Infarct |
| PRORY | Proline rich, Y-Linked |
| PRY | PTPN13 like protein, Y-Linked |
| PSA | Prostate-specific antigen |

**Q**

| | |
|---|---|
| qPCR | Real time quantitative polymerase chain reaction |

**R**

| | |
|---|---|
| RBMY | RNA binding motif protein, Y chromosome |
| RNA | Ribonucleic acid |
| RPS4Y | (40S) Ribosomal protein S4 |

**S**

| | |
|---|---|
| SHOX | Short stature homeobox |
| SMCY | Selected mouse cDNA on Y, human homolog |
| SNPs | Single nucleotide polymorphisms |
| SPGY | Aliase to deleted in azoospermia 1 (DAZ1) |
| SRY | Sex-determining region Y |
| STRs | Short tandem repeats |

**T**

| | |
|---|---|
| TACI | Total Anterior Circulation Infarct |
| TB4Y | Thymosin beta 4, Y-Linked |
| TBL1Y | Transducin β-like 1, Y-linked |
| TGIF2LY | TGF-Induced Factor 2-like, Y-linked |
| TMSB4Y | Thymosin β-4, Y-linked |
| TTTY | Testis transcript |
| TSPY | Testis-specific Y protein |

**U**

| | |
|---|---|
| UTY | Ubiquitously transcribed tetratricopeptide repeat containing, Y-linked |
| USP9Y | Ubiquitin specific peptidase 9, Y-linked |

**V**

| | |
|---|---|
| VCY | Variable charge, Y-linked |
| VNTRs | Variable number tandem repeats |

**X**

| | |
|---|---|
| XKRY | Testis specific XK-related protein, Y-Linked |
| X-PAR2 | X-Linked Pseudoautosomal region 2 |
| Xq | X chromosome's long arm |

**Y**

| | |
|---|---|
| Yp | Y chromosome's short arm |
| Y-PARs | Y-Linked Pseudoautosomal regions |
| Yq | X chromosome's long arm |
| Y-SNPs | Y-Linked single nucleotide polymorphisms |
| Y-STRs | Y-Linked short tandem repeats |

**Z**

| | |
|---|---|
| ZFY | Zinc Finger Protein, Y-Linked |

# RESUMO

A perda do cromossoma Y (*Loss of chromosome Y*: LOY) é uma aneuploidia que ocorre em mosaico, cuja percentagem aumenta durante a vida útil do indivíduo. O cromossoma Y é um cromossoma sexual e 95% deste não recombina, sendo herdado em bloco de pai para filho, representando assim, a linhagem paterna. O estudo do cromossoma Y tem aplicabilidade na resolução de casos forenses, de que são exemplos a realização de testes de paternidade, identificação genética de restos cadavéricos, investigações de agressão/abuso sexual, entre outras. Para além disso, a presença de LOY no sangue está associada com uma diminuição, em média, de 5,5 anos na esperança média de vida no homem. Assim, a LOY tem sido estudada em associação com o processo de envelhecimento e com doenças associadas à idade, de que são exemplos o cancro e as doenças cardiovasculares.

O objetivo deste estudo foi analisar o potencial da deteção da quantidade relativa de DNA do gene *Sex-determining region Y* (*SRY*) como biomarcador da presença de LOY associada ao envelhecimento de homens na população portuguesa e analisar a sua influência no desenvolvimento de doenças associadas à idade. Deste modo, foi utilizada uma técnica de PCR quantitativo em tempo real para quantificar a quantidade relativa de DNA do *SRY*, em indivíduos saudáveis e indivíduos com diagnóstico de cancro da próstata (*prostate cancer:* PC) e acidente vascular cerebral (AVC).

Neste estudo, na população saudável, não foi encontrada correlação entre a quantidade relativa de DNA do *SRY* com a idade cronológica, nem como associação com hábitos tabágicos e IMC. Contudo, foi observada uma menor quantidade relativa de DNA de *SRY* nos indivíduos com PC comparativamente aos indivíduos saudáveis. Sendo que, foi entre os doentes com PC e com menor quantidade relativa de DNA de *SRY* que se verificou maior ocorrência de formas avançadas da doença. Relativamente aos indivíduos diagnosticados com AVC, não se observaram diferenças estatisticamente significativas na quantidade relativa de DNA de *SRY*, quando comparados aos indivíduos saudáveis.

O presente estudo demonstra o potencial do uso da quantidade relativa de DNA de *SRY* como biomarcador molecular da ocorrência de LOY a partir de amostras de sangue periférico, apresentando potencial de ser analisado como biomarcador de desenvolvimento de PC. Adicionalmente, os resultados sugerem que a diminuição da quantidade relativa de DNA de *SRY* possa refletir o processo de envelhecimento, podendo futuramente ser usado como biomarcador de idade biológica.

## ABSTRACT

Loss of chromosome Y (LOY), a sex chromosome, is a mosaic hypoploidy that increases during the lifespan. 95% of this chromosome does not recombine and, consequently, is inherited from father to son as a haplotype. The study of Y chromosome polymorphisms helps solving forensic cases involving paternity testing, human remains genetic identification, sexual assault investigations, among others. Moreover, men that present occurrence of LOY in their blood cells have, on average, a lifespan decrease of 5.5 years. Thus, LOY has recently started to be studied in association with the aging process and age-related diseases such as cancer and cardiovascular diseases.

The aim of this study was to analyse the potential of *SRY* DNA relative quantity as a biomarker of LOY occurrence during the lifespan of Portuguese men and to explore its potential as a biomarker of both the aging process and age-related diseases. Therefore, this study used a qPCR analysis of the *SRY* DNA relative quantity as a relative measure of the occurrence of LOY in the blood cells of healthy individuals and in individuals diagnosed with prostate cancer (PC) or stroke. This method allowed the exploitation of *SRY* DNA relative quantity associations with biological aging and age-related diseases.

In the present study, in healthy individuals, a correlation between *SRY* DNA relative quantity and chronological age was not observed and *SRY* DNA relative quantity associations with smoking habits and BMI classes were not found. However, we observed lower *SRY* DNA relative quantities in individuals diagnosed with PC in comparison to healthy individuals. Moreover, the PC patient's subgroup with the lowest *SRY* DNA relative quantity showed a higher percentage of individuals with advanced disease. Regarding patients diagnosed with stroke, this study did not reveal any statistically significant difference regarding their *SRY* DNA relative quantity when comparing with the *SRY* DNA relative quantity of healthy individuals.

This study demonstrated the potential of *SRY* DNA relative quantity as a molecular biomarker of LOY occurrence in peripheral blood samples, which, in addition, has the potential to be analysed as biomarker of PC development. Moreover, these results suggest that the decrease on *SRY* DNA relative quantity may reflect the aging process, thus, having the potential to be used as a biomarker of biological age in future.

# LIST OF FIGURES

***Figure 9***. Dispersion of individuals from stroke group per age, hypertension, TACI, smoking habits and BMI classes and according to *SRY* DNA relative quantity (-ΔCt). **9 A:** Dispersion of individuals according to age, below 49, between 50 to 59, between 60 to 69 and above 70 years old, respectively (P=0.317); **9 B:** Dispersion of individuals according to hypertension and no-hypertension, respectively (P=0.262); **9 C:** Dispersion of individuals according to TACI and other types, respectively (P=0.695); **9 D:** Dispersion of individuals according to smoking habits, smoker, never smoker and former smoker, respectively (P=0.513); **9 E:** Dispersion of individuals according to BMI classes, normal weight, pre-obesity, obesity class I and II, respectively (P=0.558). (mean ± standard deviation). .....54

# LIST OF TABLES

# INTRODUCTION

# 1 – INTRODUCTION

## 1.1 - Loss of chromosome Y (LOY)

Loss of chromosome Y (LOY) is a hypoploidy event and it is described as a mosaic loss (1). LOY occurs during cell division and its association with age has been reported more than 50 years ago, when Jacobs and collaborators stated that LOY was increasing in healthy men over 65 years old (2). Later, in 1995, this loss was detected in healthy males with ages ranging from 1 week to 93 years old (3). However, a correlation between LOY and age was only observed in males after 16 years old (3). LOY is different from other events that are observed in male infertility, such as microdeletions (1). For instance, Y nullisomy cannot be defined as LOY, because a nullisomic event (2n-2) is a result of non-disjunction during meiosis (4). Thus, Y nullisomy does not act as a mosaic event, rather it affects all post-zygotic cells (4).

Healthy patients can have detectable levels of LOY in their blood cells and its percentage rises throughout their lifespan (5). Men that present LOY in their blood cells have a lifespan decrease of 5.5 years when compared with men without LOY (5). This evidence shows the potential of LOY as a biomarker of biological age (4, 6).

LOY, as well as gene deletions and epigenetic modifications in the Y chromosome may alter this chromosome functions (7). Consequently, these alterations, along with other factors, may contribute to the development of diseases in men (7). In Portugal, data from 2017 shows that, regarding life expectancy, men live on average less 5.6 years than women (8). Apart from the difference in life expectancy, there are gender differences in risk, incidence and progression of human diseases (9). In fact, LOY has been observed in cells of cancer patients and in cancer cell lines of several tumoral models, such as prostate carcinoma, bladder cancer, pancreatic cancer, oesophageal carcinoma, gastric cancer, head and neck carcinoma, renal cell carcinoma and testicular germ cell tumour (10-23). Moreover, there are also reports of LOY being associated with cardiovascular events (24-27).

Nevertheless, the mechanism by which the Y chromosome is lost during lifespan is not yet well understood (28). One possible mechanism of LOY is based on environmental triggers, such as smoking, that can lead to missegregation during mitosis (29). This missegregation causes aneuploidy, leading some daughter cells to become hypoploid and others to become hyperploid (30). Other possibility resides on telomeric shortening, which

increases chromosomal instability, promoting chromosomal degradation and its eventual loss in elderly males (3, 28).

## 1.2 - The Y chromosome and its loss implications

The Y chromosome was firstly discovered in 1905, by Nettie Stevens (31). Later, in 1923, Theophilus S. Painter identified the human Y chromosome during observations of meiotic divisions in spermatogonia (31, 32). Finally, in 1959, a *Sex-Determining Region Y* (*SRY*) was discovered: a gene encoding for a protein that triggers the development of the male gonadal organs during embryonic development (33, 34).

The Y chromosome is one of the shortest chromosomes of the human genome, presenting approximately 60 Mega base pairs (Mb) of length (35). The Male-Specific region of the Y chromosome (MSY), also known as non-recombinant region Y (NRY) represents 95% of its length (35). The MSY can be subdivided into two main parts: the euchromatin (40% of its length) and the heterochromatin (60% of its length) (36). Additionally, the Y heterochromatin is full of satellite Deoxyribonucleic acid (DNA) repeats and is highly abundant in *Alu* and *Line* repetitive elements which are distributed throughout the entire chromosome (37). There is no meiotic recombination in the MSY and the repeats that exist throughout the chromosome are important assets to maintain chromosomal integrity (36).

Additionally, the pseudoautosomal regions (PARs) account for the remaining 5% of the Y chromosome and are also present on the X chromosome (38-40). There are three Y Chromosome PARs (Y-PARs): Y-PAR1, on short chromosomal arm (Yp) which is 2.6 Mb long; Y-PAR2, on long chromosomal arm (Yq) which is 320 kilo base pairs (kb) long; Y-PAR3, which is located near Y-PAR1 and has approximately 2.3 Mb length. Y-PAR1 and Y-PAR2 are located near telomeric regions of the Y chromosome and Y-PARs are the only sections of this chromosome that undergo crossing over during meiosis (41, 42). Y-PAR1 has a recombination rate 17-fold greater than the genome-wide average, whereas Y-PAR2 displays a lower frequency of paring and recombination with X-PAR2 (43, 44). Y-PAR3 was only introduced in 2013 and it shares more than 98% sequence homology with Xq 21.3, but its functional significance is not totally understood, since it was only identified in 2% of the general population (45).

Genes in the human Y chromosome can be organized according to two criteria: their origin or their expression among different tissues (33). Regarding their origin, these genes can be split into three different classes: X-transposed, X-degenerate and ampliconic (33). The X-transposed genes are originated from X-Y transpositions that have happened 3.5 million years ago and they share more than 98% identity to X chromosome (Xq21) (33). The X-degenerate segments represent sequences of ancient autosomal chromosomes that have evolved towards both sex chromosomes (33, 35, 42, 46). This class includes pseudogenes or single-copy genes with 60 to 96% of homology to X-linked genes (33, 35, 42, 46). Finally, the ampliconic class genes are similar in length to the other MSY genes. This class includes testis-specific protein Y-linked (*TSPY*) gene repeats (~35 copies – varies among individuals), three inverted repeats (IR1, IR2 and IR3), two arrays of no long open reading frame (NORF) clusters (Testis transcript Y (*TTTY*) *1*, *TTTY2*, *TTTY6*, *TTTY7*, *TTTY8*, *TTTY18*, *TTTY19*, *TTTY21* and *TTTY22*) and eight palindromes (corresponding to 25% of the euchromatin) (33). Until now, 47 MSY protein-coding genes have been identified and are shown in **Figure 1** and **Table 1**.

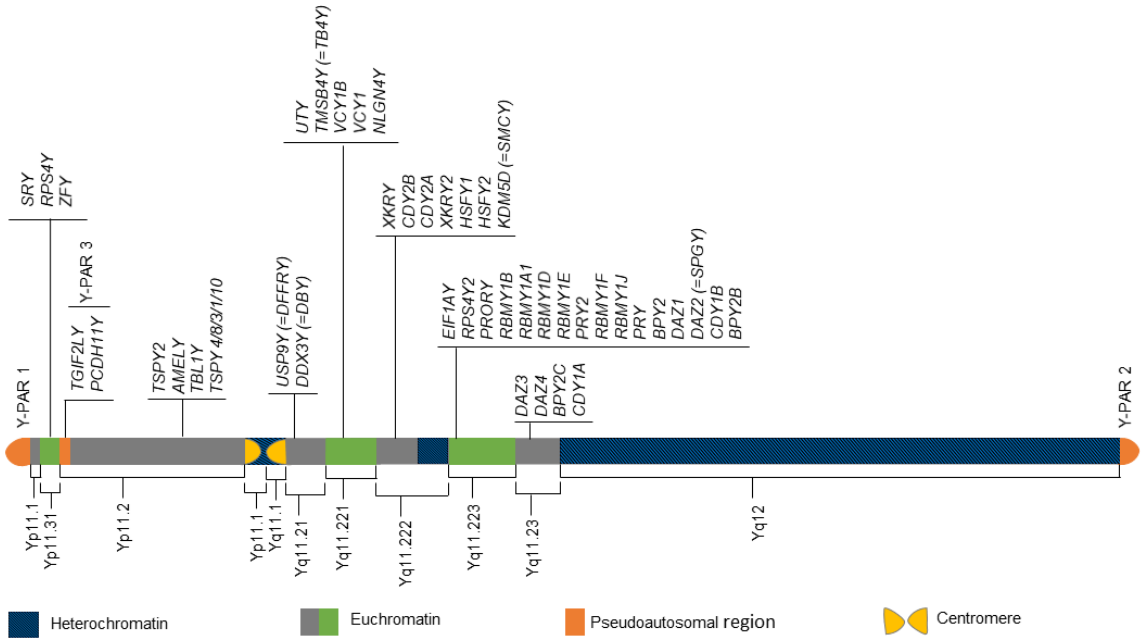


***Figure 1.*** Map of the Y chromosome. The colours used are merely representative to make loci distinguishable. On top of the map, Y chromosome's genes are presented in their respective location order. Below the diagram, loci are shown.

According to gene expression in different tissues, Y chromosome genes can be divided in testis-specific/predominant and ubiquitously expressed. **Table 1** shows the Y chromosome genes and the respective human tissues where they are expressed (9, 33, 34, 41, 47).

***Table 1.*** MSY's genes/families, respective coding proteins and tissue expression in humans.

| Loci [35] | Gene/ Family [33, 47-50] | Coding proteins [35, 47] | Tissue expression[d] [9, 33, 35, 47, 51] |
|---|---|---|---|
| **Yp 11.31** | *SRY*[a] | Sex-determining region Y | Testis, foetal Sertoli cells, adult germinative cells, Adrenal Gland, Oesophagus, Kidney, Brain and Adipose |
| **Yp 11.31** | *RPS4Y*[a] | (40S) ribosomal protein S4, Y-linked 1 | Ubiquitously expressed[b] |
| **Yp 11.31** | *ZFY*[a] | Zinc finger protein, Y-linked | Ubiquitously expressed[b] |
| **Yp 11.2** | *TGIF2LY*[a] | TGF-Induced Factor 2-like, Y-linked | Testis |
| **Yp 11.2** | *PCDH11Y*[a] | Protocadherin-11, Y-linked | [b] except Blood, Liver, Pancreas, Heart (Left ventricle) and skeletal muscle |
| **Yp 11.2** | *TSPY2*[a] | Testis-specific protein, Y-linked 2 | Testis |
| **Yp 11.2** | *AMELY*[a] | Amelogenin, Y-linked | Testis, Pancreas, Thyroid and Teeth |
| **Yp 11.2** | *TBL1Y*[a] | Transducin β Like 1, Y-linked | Prostate, Kidney (cortex), Pancreas, Oesophagus, Thyroid and Adipose |
| **Yp 11.2** | *TSPY4*[a] | Testis-specific protein, Y-linked 4 | Presumed to be expressed on Testis |
| **Yp 11.2** | *TSPY8*[a] | Testis-specific protein, Y-linked 8 | Presumed to be expressed on Testis |
| **Yp 11.2** | *TSPY3*[a] | Testis-specific protein, Y-linked 3 | Presumed to be expressed on Testis |
| **Yp 11.2** | *TSPY1*[a] | Testis-specific protein, Y-linked 1 | Presumed to be expressed on Testis |
| **Yp 11.2** | *TSPY10*[a] | Testis-specific protein, Y-linked 10 | Presumed to be expressed on Testis |
| **Yq 11.21** | *USP9Y*[a] *(=DFFRY)* | Ubiquitin Specific Peptidase 9, Y-linked | Ubiquitously expressed[b] |
| **Yq 11.21** | *DDX3Y (=DBY)* | DEAD-Box Helicase 3, Y-linked | Ubiquitously expressed[b] |
| **Yq 11.221** | *UTY*[a] | Ubiquitously Transcribed Tetratricopeptide Repeat Containing, Y-linked | Ubiquitously expressed[b] |
| **Yq 11.221** | *TMSB4Y*[a] *(=TB4Y)* | Thymosin β-4, Y-linked | Ubiquitously expressed[b] |
| **Yq 11.221** | *VCY1B*[a] | Variable Charge 1B, Y-linked | Testis |
| **Yq 11.221** | *VCY1*[a] | Variable Charge 1, Y-linked | Testis |
| **Yq 11.221** | *NLGN4Y*[a] | Neuroligin 4, Y-linked | Ubiquitously expressed[b] |
| **Yq 11.222** | *XKRY* | Testis-specific XK-related protein, Y-linked | Testis specific |
| **Yq 11.222** | *CDY2B* | Chromodomain, Y-linked 2B | Testis |
| **Yq 11.222** | *CDY2A* | Chromodomain, Y-linked 2A | Testis specific |
| **Yq 11.222** | *XKRY2* | Testis-specific XK-related protein, Y-linked 2 | Testis specific |
| **Yq 11.222** | *HSFY1* | Heat shock transcription factor, Y-linked 1 | Testis, Prostate, Adrenal Gland, Coronary Artery, Thyroid, Brain (Cortex and Cerebellum) |

***Table 1***. MSY's genes/families, respective coding proteins and tissue expression in humans. (cont.)

| Loci [35] | Gene/ Family [33, 47-50] | Coding proteins [35, 47] | Tissue expression[d] [9, 33, 35, 47, 51] |
|---|---|---|---|
| **Yq 11.222** | *HSFY2* | Heat shock transcription factor, Y-linked 2 | Testis specific |
| **Yq 11.222** | *KDM5D (=SMCY)* | Lysine-specific demethylase 5D | Ubiquitously expressed[b] |
| **Yq 11.223** | *EIF1AY*[a] | Eukaryotic translation initiation factor 1A, Y-linked | Ubiquitously expressed[b] |
| **Yq 11.223** | *RPS4Y2*[a] | (40S) Ribosomal protein S4, Y isoform 2 | Ubiquitous expression |
| **Yq 11.223** | *PRORY* | Proline-Rich Protein, Y-linked | Testis |
| **Yq 11.223** | *RBMY1B*[a] | RNA-binding motif protein, Y chromosome, family 1 member B | Testis |
| **Yq 11.223** | *RBMY1A1*[a] | RNA-binding motif protein, Y chromosome, family 1 member A1 | Testis specific |
| **Yq 11.223** | *RBMY1D*[a] | RNA-binding motif protein, Y chromosome, family 1 member D | Testis specific |
| **Yq 11.223** | *RBMY1E*[a] | RNA-binding motif protein, Y chromosome, family 1 member E | Testis specific |
| **Yq 11.223** | *PRY2* | PTPN13-like protein, Y-linked 2 | Testis, Heart, Lung, White blood cells |
| **Yq 11.223** | *RBMY1F*[a] | RNA-binding motif protein, Y chromosome, family 1 member F | Testis specific |
| **Yq 11.223** | *RBMY1J* | RNA-binding motif protein, Y chromosome, family 1 member J | Testis specific |
| **Yq 11.223** | *PRY* | PTPN13-like protein, Y-linked | Testis |
| **Yq 11.223** | *BPY2* | Basic charge protein, Y-linked, 2 | Testis |
| **Yq 11.223** | *DAZ1* | Deleted in azoospermia 1 | Testis, Stomach and Liver |
| **Yq 11.223** | *DAZ2 (=SPGY)* | Deleted in azoospermia 2 | Predominantly in Testis[c] |
| **Yq 11.223** | *CDY1B*[a] | Chromodomain, Y-linked 1B | Testis specific |
| **Yq 11.223** | *BPY2B* | Basic charge protein, Y-linked, 2B | Testis specific |
| **Yq 11.23** | *DAZ3* | Deleted in azoospermia protein 3 | Predominantly in Testis[c] |
| **Yq 11.23** | *DAZ4* | Deleted in azoospermia protein 4 | Predominantly in Testis[c] |
| **Yq 11.23** | *BPY2C* | Basic charge protein, Y-linked, 2C | Testis specific |
| **Yq 11.23** | *CDY1A*[a] | Chromodomain, Y-linked 1A | Testis specific |

In green, orange and purple are represented the X-degenerate, ampliconic and X-transposed genes, respectively. Y-PAR1 and Y-PAR2 are part of the Y chromosome but are not represented. Only the Y-PAR3's genes are represented (in purple). [a] X-homologue genes are in blue and the Y-specific are in black. [b] includes the testis, prostate, kidney (cortex), adrenal gland, blood, spleen, liver, stomach, transverse colon, terminal ileum, pancreas, oesophagus, coronary artery, heart (left ventricle and atrial appendage), lung, thyroid, brain (cerebellum and cortex), skeleton muscle and adipose tissues. [c] Other tissue(s) also express(es) it, but testis expression has higher levels. [d] there is a different intensity of expression between genes and among the same gene; there are different expressions between tissues as well.

The Y chromosome is crucial for the development of the male phenotype during embryonic development (52). S*RY* is a single-copy gene located on the p arm of the Y chromosome and encodes a transcription factor which has the ability to trigger the differentiation of gonadal tissue into male gonads (53). *SRY* gene is only composed of exons and its protein has an high mobility group box (HMG-box) which binds to DNA and gives this protein its transcription factor activity (51).

In general, the Y chromosome contributes to stature, tooth size and spermatogenesis processes (36). For example, Short Stature Homeobox-containing (*SHOX*) gene of the Y chromosome is a transcription factor which participates in stature-related processes (40). The depletion of this gene leads to less transcriptional activation of growth-related genes and, consequently, to shorter individuals (54). Another Y chromosome gene, Amelogenin Y-linked (*AMELY*), has an homologue gene in the X chromosome (*AMELX*) and both are involved in biomineralization during tooth enamel development (55). Regarding spermatogenesis, there are three Azoospermia Factor (*AZF*) regions on the Y chromosome, *AZFa*, *AZFb* and *AZFc*, that are important to assure male fertility, since deletions occurring on these regions, cause spermatogenesis failure (53).

Additionally, the fact that chromosome Y genes are expressed in other tissues apart from the testis further suggests that the Y chromosome may participate in more processes apart from this organ (35). Jangravi and co-workers established that MSY proteins participate in several biological processes including DNA transcription, cell differentiation, metabolic processes, gonad development, nucleosome assembly, tissue development, chromatin modification, messenger Ribonucleic acid (mRNA) translation, sex differentiation, cell adhesion, cell proliferation, RNA metabolism and sex determination (35). For example, *TSPY* is a gene that interacts with Cyclin B protein, a mitotic cell cycle regulatory protein, suggesting that this gene may cooperate in the regulation and promotion of cell proliferation (56). On the other hand, Yuan X and collaborators have reported that the *SRY* gene has the ability to negatively regulate androgen receptor (AR) which represses AR transcriptional activity (57). Since these genes may contribute to processes outside of the testis, this chromosome is also being studied in other diseases besides infertility, such as prostate cancer (11, 58).

Even though in a forensic context, autosomal profiling is preferred and has more validity in court, the study of the Y chromosome is widely applied in forensic science (59). The study of this chromosome is applied to support paternity testing studies, sexual assault investigations and exclusion of males from crime scenes (59). It is frequently used in cases when standard autosomal DNA profiling is not enough to answer forensic

questions (60). This includes cases of DNA mixture (for example, blood-blood and blood-saliva samples), “gang rape” (group of people that participates in the rape of one victim) and mass disaster victim identification (60). Even when the father is missing, if there is a male relative, such as a grandfather, the Y chromosome can be used to test if these two individuals belong to the same family (61).

Forensic applications of this chromosome are based on the analysis of genetic polymorphisms such as short tandem repeats (STRs) and single nucleotide polymorphism (SNPs) (59). STRs are polymorphic individualizers that are disperse all over the genome and consist of repeat units ranging from two to six base pairs (bp) in length (62). On the other hand, SNPs are genetic variations of a single nucleotide, which on average occur in every 300 nucleotides (nt) throughout an individual’s DNA (63). Every individual has a different combination of SNPs and STRs that make their own “fingerprint”, allowing for individual differentiation in forensic cases (59). 27 Y-STRs are already being included in forensic commercial kits used in the routine practice of forensic laboratories, allowing paternal lineage characterization with a high degree of certainty (64). Besides, these commercial kits allow the detection and characterization of low male DNA quantities from mixed stains (59, 64). Even with an excess of female DNA, male DNA can still be detected, which is particularly important in sexual assault sample’s analysis (59, 64).

To study Y chromosome’s diversity among individuals and populations, two groups of polymorphisms can be distinguished: bi-allelic and multi-allelic, which define haplogroups and haplotypes, respectively (61). Haplogroups can be defined using Y-SNPs and *Alu* elements that have lower mutations rates (~$10^{-8}$ to $10^{-9}$ per generation) than Y-STRs (65). This group is currently applied in human migratory studies and bio-geographic ancestry, which, in this case, researches a person’s paternal ancestors (66, 67). On the other hand, haplotypes are characterized based on Y-STRs and minisatellites (Variable Number of tandem Repeats - VNTRs, 10 to 15 bp in length), which have higher mutation rates than Y-SNPs (68). Since the Y chromosome is inherited only from father to son, haplotypes are similar to a “family fingerprint”, so Y-STRs haplotyping can also identify a paternal lineage (59). Haplotypes can make criminal’s trace easier if the offenders are related to each other (59). Cases involving human remains of missing person, where only relatives are available for Y-STRs analysis, allow the identification of the human male remains as one of their family member (69). Standard Y-STRs can also be applied to paternity testing, when autosomal STRs are not available for analysis or technician errors occur, if the Y-haplotypes match, it means that, that man and child, belong to the same family (59).

The Y chromosome was thought to be a relatively stable chromosome since most of it did not recombine during meiosis and its content remained almost unaltered from father to son. However, this idea overshadowed the thought that Y chromosome microdeletions or even its total loss could interfere with Y chromosome analysis. Recently, a study suggested that microdeletions seen in the Y chromosome of infertile men could introduce bias results in Y-STRs and Y-SNPs forensic analysis (70). Although LOY is different from microdeletions, it represents a mosaic loss of the entire chromosome which could also introduce bias to forensic techniques. Moreover, blood and buccal samples are the most frequent sources of information when a new forensic case appears and the occurrence of LOY has already been observed in both samples (71).

## 1.3 - LOY and the aging process

Recently, studies have proposed that LOY can be a potential biomarker of biological age (4, 6). Unlike chronological age, which reflects the time that has passed since birth, biological age is a reflection of the physiological aging process which considers the biological state of cells, tissues and organs (72). There is a need to quantify biological age, but a quantification based on numbers cannot be given since the aging process is diverse and varies among individuals of the same age (72). The aging process is a result of progressive biological deterioration related to age, which leads to the progressive loss of functions in tissues and in organs (73). This loss of function is a consequence of increased oxidative stress, telomeric shortening, DNA damage, mitochondrial disfunction, chromatin disruption, inflammation status, epigenetic deregulation and oncogenic activation (73). The disfunction of biological processes, such as DNA damage repair, increases the susceptibility to age-related diseases development in elderly (73). Intrinsic and extrinsic factors, such as smoking habits and pollution exposure, can promote cell stress and DNA damage which accelerates the aging process (73). It is well accepted that smoking habits are a key aging accelerator, triggering inflammatory responses and favouring the occurrence of several diseases (74). This accelerated aging separates biological age from the chronological age, making biological age a best predictor of lifespan than chronological age (72).

Since biological age distances itself from the chronological age of an individual, there is a need to discover new specific biomarkers of biological age. Biomarkers are molecules

that can be objectively measured in the body fluids as an indicator of normal biological or pathological processes, or a response to a therapeutic intervention (75). The ideal biomarker must be easy to access through minimal invasive procedures (such as taking a blood sample) and be analysed by low cost quantification methods that are specific and sensible to the parameter of interest (75, 76). According to American Federation for Aging Research (AFAR) a biomarker of aging must fulfil some criteria in order to be applied: 1) predict aging rate; 2) monitor basic processes that correlate with aging rather than disease effects; 3) be able to be repeatedly tested with minimal harm to the person; 4) work in both humans and laboratory animals (77). Given the complex nature of the aging process, it is difficult to find a biomarker that fits all four items proposed by AFAR (77).

LOY is a normal aging process that can be measured in peripheral blood and its percentage rises throughout the lifespan (3, 6). This loss is dependent on environmental factors exposure, such as smoking and outdoor pollution, which can also increase the percentage of LOY in male individuals (29). Consequently, higher LOY in blood cells increases biological age and takes years from the lifespan of an individual (5). Using three different cohorts, a strong association between smoking habits and LOY was identified: the highest risk of carrying blood cells without the Y chromosome was found to be 2.4 to 4.3 times greater in current smokers comparing with non-smokers and past smokers (29). However, this association between increased LOY and smoking began to fade years after current smokers stopped smoking (29, 71). Regarding outdoor air pollution, in a study performed in the United States, the outdoor pollution was measured considering particulate matter lower or equal to 10µm ($PM_{10}$), carbon monoxide, nitrogen dioxide, sulphur dioxide and ozone data (26). These agents, particularly $PM_{10}$, can induce a variety of effects at the cellular level, such as DNA damage, inflammation and genomic instability which cause malfunctioning in mechanisms such as chromosomal segregation and cell cycle progression (26). These pollutes were associated with higher frequency of LOY in older American men (over 65 years old) (26).

The influence of internal factors that are intrinsic to the individuals but also of external factors in LOY shows its potential to be a biomarker of biological age, since it has the potential of showing the state of aging of the body of an individual (78). Because of this, researchers have started to study the relation of LOY with age-related diseases, such as cancer and cardiovascular events (1, 4, 6, 24, 28).

## 1.4 - Age-related diseases

Biological systems in older individuals are less able to overcome the effects of the exposure to intrinsic and extrinsic factors (79). The struggle to combat the effect of those agents and their consequences increases the risk for life-threatening conditions (80). To the elderly, these conditions bring a series of psychological, cognitive, social and environment changes and challenges (81). World population is getting older and age-related diseases have considerable implications in public health (81). The increased need for caregiver support, palliative care, assistive technologies and transportation are current increasing problems in our older population (81). Age-related disorders such as cancer, cardiovascular and neurodegenerative diseases are the most prevalent disorders in elderly (80). Each individual has its own genetic background and has different responses to risk factors exposure: two individuals with the same age have different aging rates and disease susceptibilities (80). Usually, therapies target age-related diseases individually which is a current problem since all these diseases have the aging process in common (80). Targeting the aging process itself would improve age-related diseases prevention, early diagnosis and therapy resulting in a need for biomarkers of age-related diseases (79).

The first study to report LOY associated with a disease was published in 1985, when Holmes and co-workers reported LOY in acute myeloid leukaemia (AML) patients (82). Since then, LOY has been reported to be associated with the development of several age-related diseases, such as Alzheimer's disease, cardiovascular diseases and cancer (4, 6, 24, 26-28). Cancer is an heterogenous disease that could present different aetiology, evolution and treatment (83). Moreover, cancer is a major public health problem and is the second leading cause of death worldwide (84). Approximately 1 in every 6 deaths is consequence of cancer, causing more deaths than acquired Immunodeficiency syndrome (AIDS), tuberculosis and malaria combined (84). Prostate cancer (PC) is the second most frequent cancer in men (13.5%) and the fifth (6.7%) cause of cancer death in the world, as shown in **Figure 2** (85). In Portugal, 6609 Portuguese men were diagnosed with PC in 2018 (86). In that same year, this type of cancer was the fourth leading cause of death by cancer in the Portuguese population (86).

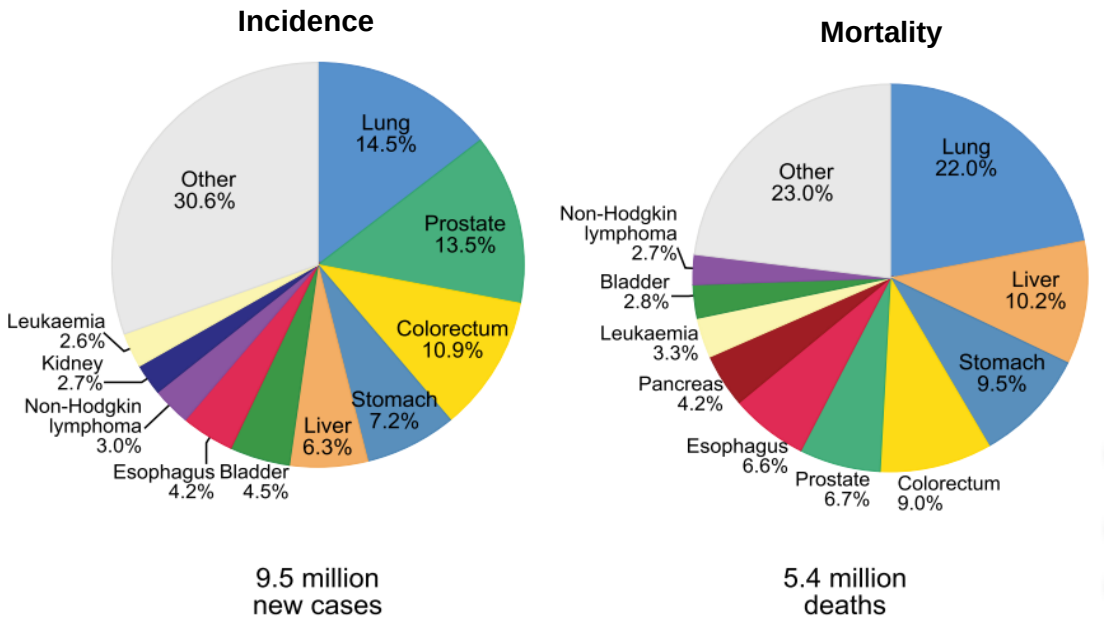


***Figure 2.*** Pie charts representing the distribution of cases and deaths (proportion of the total number of cases of death) for the 10 Most Common Cancers in 2018 for men. Source: GLOBOCAN 2018.

Risk factors associated with PC include, age, ethnicity, disease family history, diet, physical activity, hormone levels and the genetic background of the individual (87). PSA test has revolutionised PC diagnosis since 1980's, but PSA levels are organ-specific and not cancer specific (87, 88). Higher PSA levels may indicate early stages of prostatitis, benign prostate hyperplasia (BPH) and other non-malignant conditions (87, 88). Even though PSA test is good to detect early stages of the disease, this test cannot be used to distinguish between the benign or malign tumour (88). Moreover, PSA discovery was for a long time a problem because its overuse as an indicator of PC's presence lead to overdiagnosis and overtreatment with the performance of unnecessary invasive exams in some patients (88). However, this problem has been attenuated with guideline improvements and by considering individual background (87, 88). The confirmatory exam for PC diagnosis is a prostate biopsy or surgical resection, which can be very aggressive (87, 89). After microscopic observation of prostate tissue, a Gleason score is given, determining prostate cancer severity (90). Since PSA test cannot be used alone, there is a need of more accurate and specific biomarkers that can help the early diagnosis of PC without the misuse of invasive procedures.

Reports suggest that PC is genetically complex. Its development is dependent on somatic copy number variations, structural rearrangements, point mutations and changes in chromosomal number (91). Noveski and co-workers reported that a higher occurrence

of LOY in blood cells conferred a higher risk of developing PC in a case-study performed in individuals from the Republic of Macedonia (28). The uncontrolled proliferation characteristics of cancer mean that every cell division provides an opportunity for chromosome mutations, deletions and even total chromosomal loss (90). Several genes of the Y chromosome have been reported lost in PC cancer cells and they also have been associated with cancer progression (92). In a study done in athymic nude mice, when an exogenous chromosome Y was added, the tumorigenicity of PC-3 cell line was suppressed (93). This could mean that tumour suppressor genes reside on the Y chromosome and that their loss may contribute to cancer progression and aggressiveness (28). However, the Y chromosome may also have oncogenes, like *TSPY* which oncogenic activity has been a focus of study (11, 56, 94). The fact that LOY may provide an advantage to cancer progression allows this loss to have the potential to become a biomarker of cancer aggressiveness and prognosis (16).

Other age-related diseases that have also been studied in association with LOY are cardiovascular diseases (24, 25, 27). This type of disease is associated with the aging process, since several age-related factors, such as atherosclerosis, hypertension, vascular remodelling and stiffness, increase the risk for cerebrovascular events such as strokes (79, 81).

Cerebrovascular diseases are the second cause of death in Europe and one of the main causes of death in Portugal (95). Premature death in Portugal occurs largely in consequence of a stroke and ischemic heart disease (96). Portuguese data shows that six individuals per hour are having a stroke at this moment, in which two to three result in death (97). The main risk factors that increase the probability of having a stroke include, hypertension, high cholesterol levels, smoking habits, low physical activity and a poor diet (98).

A stroke is characterized as a neurological deficit, which occurs due to an acute focal injury in the central nervous system (99). The causes of this injury can be subdivided in haemorrhagic or ischemic (99). The first one could be due to a cerebral infarction, intracerebral haemorrhage or subarachnoid haemorrhage (99). On the other hand, ischemic stroke results from an atherosclerotic plaque detachment or a gas embolism (98). These embolisms obstruct the artery channel and block the oxygen from reaching the brain causing a stroke (98). Oxfordshire Community Stroke Project (OCSP) classify stroke accordingly with the affected area in the brain (100). Accordingly with OCSP, strokes are classified in Total anterior circulation infarct (TACI), Partial anterior circulation infarct (PACI), Lacunar infarct (LACI) and Posterior circulation infarct (POCI) (100).

Although TACI are the less frequent type of stroke diagnosis, they have been associated with a worse prognosis since these patients have a higher mortality and morbidity rate (101).

There is an increased prevalence of strokes in men as well as hypertension, one of its main risk factors (97). These prevalence differences suggest that Y chromosome may be a possible contributor to this disease in men (102). The study of the Y chromosome genes and polymorphisms is currently being used to explore this chromosome associations with strokes. For example, the presence of a polymorphic restriction site on the MSY region of the Y chromosome, *HINd III*, is associated with higher blood pressure (103). These authors associated this polymorphism with increased risk of higher diastolic blood pressure in men (103). Besides hypertension, Y chromosome polymorphisms have been studied in association with the risk of developing atherosclerotic plaques (104). While studying different Y-haplogroups, the risk of having an atherosclerotic plaque present was 2.5-fold higher in haplogroup K (ancient Southwest of Asia) than the other haplogroups (104). Additionally, when blood samples from healthy patients and after-stroke patients were compared, two MSY genes, along with some Y-PAR1 genes were found to be upregulated in after-stroke patient's samples (105). However, the significance of this upregulation and its consequences are yet to be elucidated.

Haitjema and co-workers reported that LOY was present in both, blood cells and atherosclerotic lesions of male patients that had been submitted to a Carotid Endarterectomy (24). Age and smoking habits are coincident risk factors for increased LOY and risk of having a stroke (1, 29). LOY in peripheral blood increases with these risk factors and it is associated with a 2-fold risk for secondary major cardiovascular events (24). Therefore, LOY detection in blood cells has the possible application as risk predictor and has the potential to help the implementation of preventive measures in patients at risk (1).

The association of LOY occurrence with the aging process and age-related diseases, such as PC and strokes, reveals a need to evaluate its potential as a biomarker of biological age, useful in a legal medicine context.

# OBJECTIVES

# 2 – OBJECTIVES

## 2.1 – Main objective

The main objective of this study is to analyse the potential of *SRY* DNA relative quantity, as a biomarker of LOY occurrence, during the lifespan of men from the North of Portugal and to explore LOY's potential as aging biomarker.

## 2.2 – Specific objectives

- Analyse the *SRY* DNA relative quantity in peripheral blood samples from healthy individuals and in individuals diagnosed with PC and stroke and its influence in the development of these age-related diseases;
- Analyse the association of *SRY* DNA relative quantity with chronological age, smoking habits and body mass index.

# MATERIALS AND METHODS

# 3 – MATERIALS AND METHODS

## 3.1 – Population selection

The study of LOY in Portuguese men involved a total of 1011 male individuals from the North of Portugal. According to the Declaration of Helsinki, all individuals gave their written consent after previously understanding of the study.

In this study, 259 healthy male individuals (control group), without evidence of neoplastic disease or stroke clinical history, were selected from the blood donor's biobank of Portuguese Oncology Institute of Oporto. Their mean age was 39.95 ± 12.48. In this control group, 50% individuals were never smokers, 28% were smokers and 22% were former smokers. Regarding their body mass index (BMI), 30% had a normal BMI, 52% had a pre-obesity BMI, 16% had obesity class I and 2% were obese class II.

Additionally, 520 men with histopathological diagnosis of Prostate cancer (PC group) were recruited in Portuguese Oncology Institute of Oporto. This group had an mean age of 68.42 ± 7.65 years old. 42% of these individuals had T1+T2 stage of tumour development, while 58% had advanced tumour development (T3+T4). Concerning smoking habits, from the selected individuals, 55% were never smokers, 12% were smokers and 33% were former smokers. Regarding their BMI, 31% had a normal BMI, 50% had pre-obesity BMI and 19% had obesity class I, II and III.

Lastly, 232 male individuals with stroke diagnosis (stroke group) were recruited from São João Hospital Center. These individuals showed an mean age of 67.27±12.34 years. Within this group, 18% had been diagnosed with TACI and the remaining 82% had other types of clinical stroke diagnosis. From the selected individuals 30% were never smokers, 33% were smokers and 37% were former smokers. Regarding their BMI, 30% had a normal BMI, 50% had pre-obesity BMI and 20% had obesity class I and II.

To access LOY in the blood cells of the three groups, an 8 mL sample of peripheric blood was collected from the subjects, through a standard method of intravenous collecting with Ethylenediamine tetraacetic acid (EDTA) tubes.

## 3.2 – DNA extraction from peripheral blood samples

DNA from the selected blood samples was extracted with *GRS - Genomic DNA Kit – Blood & Cultured Cells*, accordingly with *GRiSP Research Solutions*® recommendations. The quantity and purity of the isolated DNA was accessed based on light absorbances measured at 260 and 280nm by *NanoDrop Lite Spectrophotometer* (*Termo Fisher Scientific*®). After accessing DNA quantity and quality, the obtained DNA was stored at -20ºC until its molecular analysis.

## 3.3 – *SRY* DNA relative quantification

The relative quantification of *SRY* DNA as a LOY occurrence biomarker was performed in each sample via real time quantitative polymerase chain reaction (qPCR), using *StepOne™ qPCR Real-Time PCR* (*Applied Biosystems*®) and *StepOnePlus™ qPCR Real-Time PCR* (*Applied Biosystems*®). All the collected results were analysed by *StepOne version 2.3* (*Applied Biosystems*®) software.

The qPCR reaction was performed using *Xpert Fast SYBR (uni) 2X mastermix – GriSP Research Solutions*® and specific primers for the amplification of an endogenous control and a gene that would represent chromosome Y. The selected endogenous control was β-globin (HBB). The selected primers for HBB were *HBB1* [5'GCTTCTGACACAACTGTGTTCACTAGC3'] and *HBB2* [5'CACCAACTTCATCCACGTTCACC3'], both in a concentration of 500nM per reaction. The qPCR conditions included an initial 95ºC denaturation for 3 minutes, followed by 32 cycles of denaturation at 95ºC for 5 seconds and amplification at 58ºC for 30 seconds. After the 32 cycles, a melting curve with a melting temperature of 84.11ºC was obtained.

The *SRY* gene was selected as the target gene. To amplify a fragment of *SRY* gene the following primers were selected: *SRY-150F* [5'AGCAGTCAGGGAGGCAGATCA3'] and *SRY-245R* [5'CCCCCTAGTACCCTGACAATGTAT3'], both in a concentration of 100nM per reaction. The qPCR conditions included an initial denaturation at 95ºC for 3 minutes, followed by 35 cycles of denaturation at 95ºC for 5 seconds and amplification at 64ºC for 30 seconds. After the 35 cycles, a melting curve with a melting temperature of 80.09ºC was obtained.

In addition, an efficiency curve was performed for each qPCR reaction. All qPCR reactions obtained a *SYBR* efficiency superior to 90% and the respective negative controls and quantifications were performed in duplicates.

## 3.4 – Statistical analysis

The statistical analysis was conducted using *IBM®SPSS®Statistics* software for Windows (*version 25.0*). The Livak method ($2^{-\Delta\Delta Ct}$) was used to quantify differences in *SRY* DNA relative quantity among the groups. Moreover, the Mann-Whitney U test and t student test were used in order to evaluate the statistical differences observed between the groups.

# RESULTS

# 4 – RESULTS

## 4.1 - *SRY* DNA relative quantification analysis in the Control Group

In order to study the occurrence of LOY, the DNA relative quantity of the *SRY* gene, one of the chromosome Y most studied genes, was assessed. Therefore, less *SRY* DNA relative quantity will suggest a higher occurrence of LOY on the analysed blood samples.

We observed a high variability of *SRY* DNA relative quantity in the control group. Based on the dispersion of the -ΔCt values, it was possible to observe that our sample population can be subdivided into two different subgroups as shown in Figure 3. In fact, group 1 presented significant lower *SRY* DNA relative quantity than group 2, with a fold-decrease of 0.15 ($P<0.001$).

According to the results, we did not observe in control group 1 any association among the *SRY* DNA relative quantity and age, smoking habits and BMI (Figure 3 A-B-C). In the same control group, we also verified a heterogeneous dispersion of the *SRY* DNA relative quantity. On the other hand, even though there was no association among the *SRY* DNA relative quantity and age, smoking habits and BMI in the control group 2, its *SRY* DNA relative quantity was more homogeneous among the individuals (Figure 3 D-E-F). Due to the homogeneity observed in *SRY* DNA relative quantity of the control group 2, this group was used in the subsequent analysis.

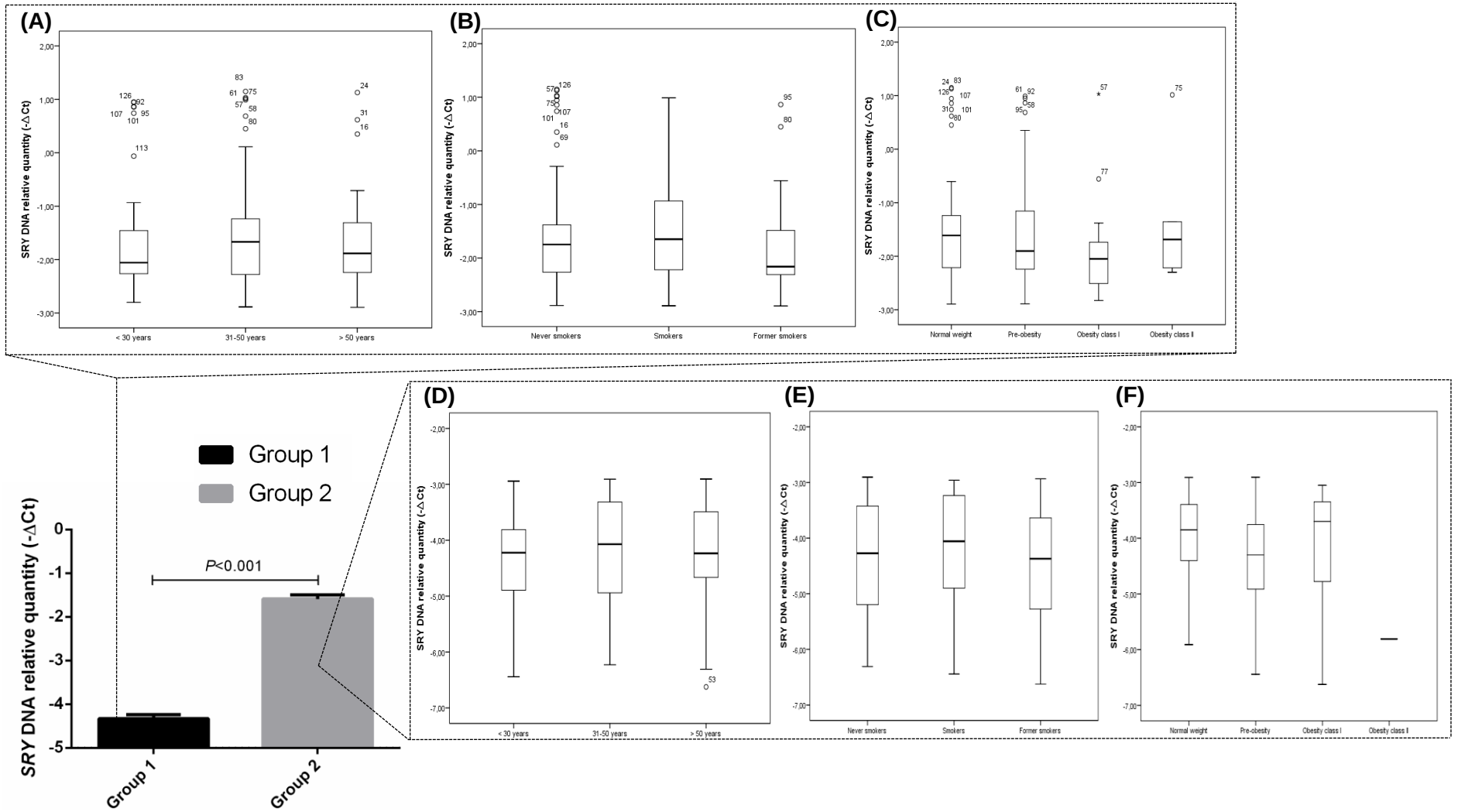


***Figure 3***. *SRY* DNA relative quantity (-ΔCt) detected in blood cells in the group 1 and 2 from the healthy individuals group (mean ± standard error of mean). **3 A-B-C**: Dispersion of individuals from group 1, considering *SRY* DNA relative quantity and according to age (A), smoking habits (B) and BMI (C); **3 D-E-F**: Dispersion of individuals from group 2, considering *SRY* DNA relative quantity and according to age (D), smoking habits (E) and BMI (F). (mean ± standard deviation).

In figure 4, we can observe the correlation between *SRY* DNA relative quantity and the chronological age. In our healthy population, we did not find any correlation between the age and the blood *SRY* DNA relative quantity.

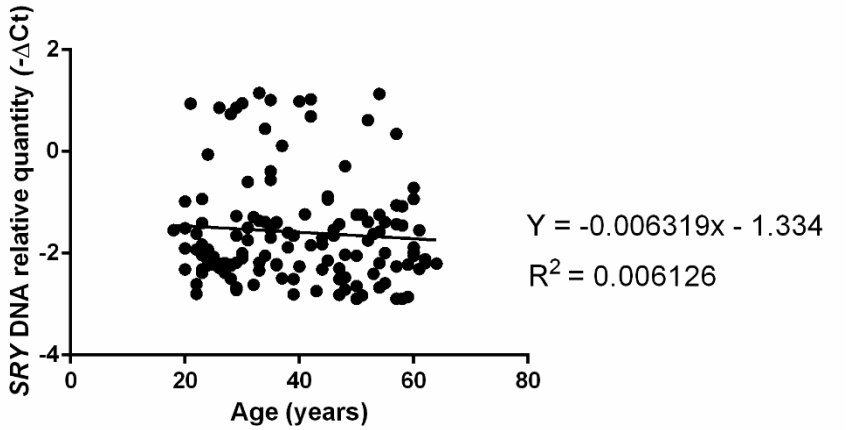


***Figure 4.*** Correlation between age (years) and *SRY* DNA relative quantity (-ΔCt) detected in blood cells of individuals from the control group (Group 2).

## 4.2 - *SRY* DNA relative quantification analysis in the PC Group

In the PC Group, we observed that individuals with this diagnosis had significantly less *SRY* DNA relative quantity in their blood cells than the individuals of control group (fold-decrease=0.48, *P*<0.001) (Figure 5 A). Additionally, we observed that PC group could be stratified into three subgroups based on *SRY* DNA relative quantity: lower quantity, intermediate quantity and higher quantity (Figure 5 B).

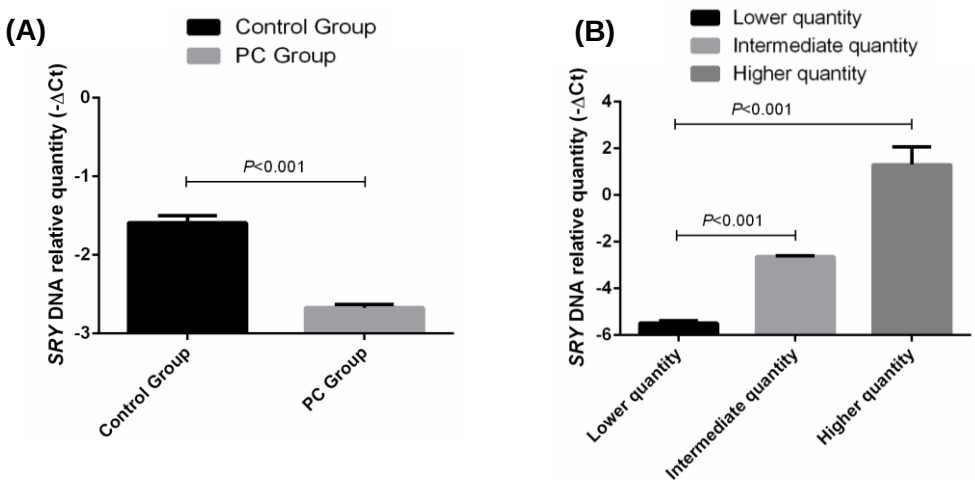


***Figure 5.*** *SRY* DNA relative quantity (-ΔCt) detected in blood cells in the control group and in PC group. **5 A**: *SRY* DNA relative quantity comparation between individuals from the control group and PC group; **5 B**: *SRY* DNA relative quantity comparation between individuals from the three groups defined in the PC group: lower quantity, intermediate quantity and higher quantity. (mean ± standard error of mean).

Considering PC aggressiveness, in the lower quantity group, there was a higher percentage of individuals with advanced stage (85%), Gleason score of eight or higher (54%) and PSA levels above 10 ng/mL (62%) (Figure 6 A). Whereas within each of the other two groups, intermediate quantity and higher quantity, the percentages of individuals according to advanced stage were not that discrepant. Regarding Gleason score, there were higher percentages of individuals with Gleason score under eight, 71% and 70%, for Intermediate quantity and higher quantity groups, respectively (Figure 6 B-C). Additionally, we could observe a higher percentage of individuals with PSA levels above 10ng/mL in both groups, 59% and 67%, respectively (Figure 6 B-C).

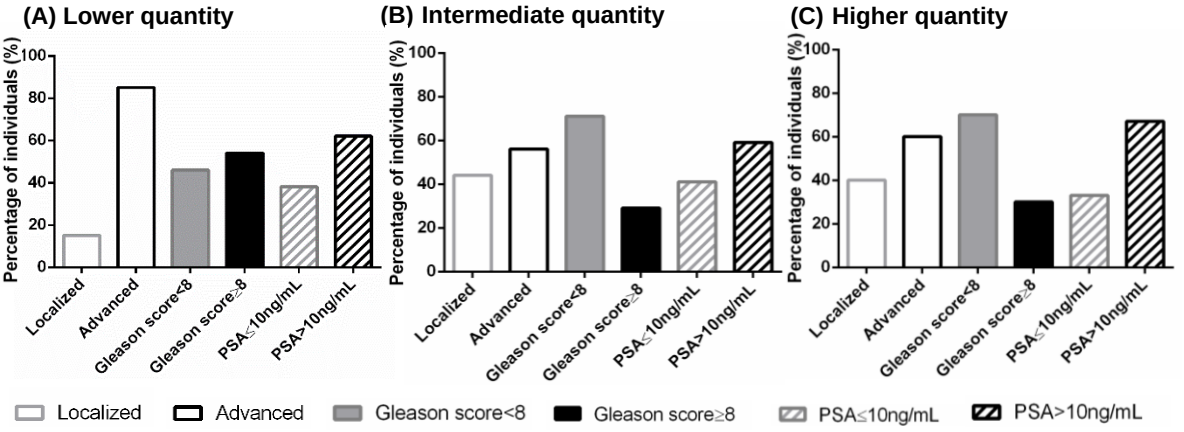


***Figure 6.*** Percentages of individuals from the three PC groups divided according to stage, Gleason score and PSA levels, respectively. **6 A**: Percentages within the lower quantity group; **6 B**: Percentages within the intermediate quantity group; **6 C**: Percentages within the higher quantity group.

Regarding the distribution of factors such as age, smoking habits and BMI within the three PC subgroups, there was a higher percentage of individuals older than 61 years old (85%), never smokers (72%) and pre-obese individuals (90%) within the lower quantity group (Figure 7 A-B-C). The intermediate quantity group was the only group with individuals under 50 years old (Figure 7 D). Additionally, in the intermediate quantity group, there was a higher percentage of individuals between 61 to 70 years old (46%), of never smokers (54%) and pre-obese (49%) (Figure 7 D-E-F). Regarding the higher quantity group, there was a higher percentage of individuals with more than 70 years old (45%), never smokers (57%) and normal weight (50%) (Figure 7 G-H-I).

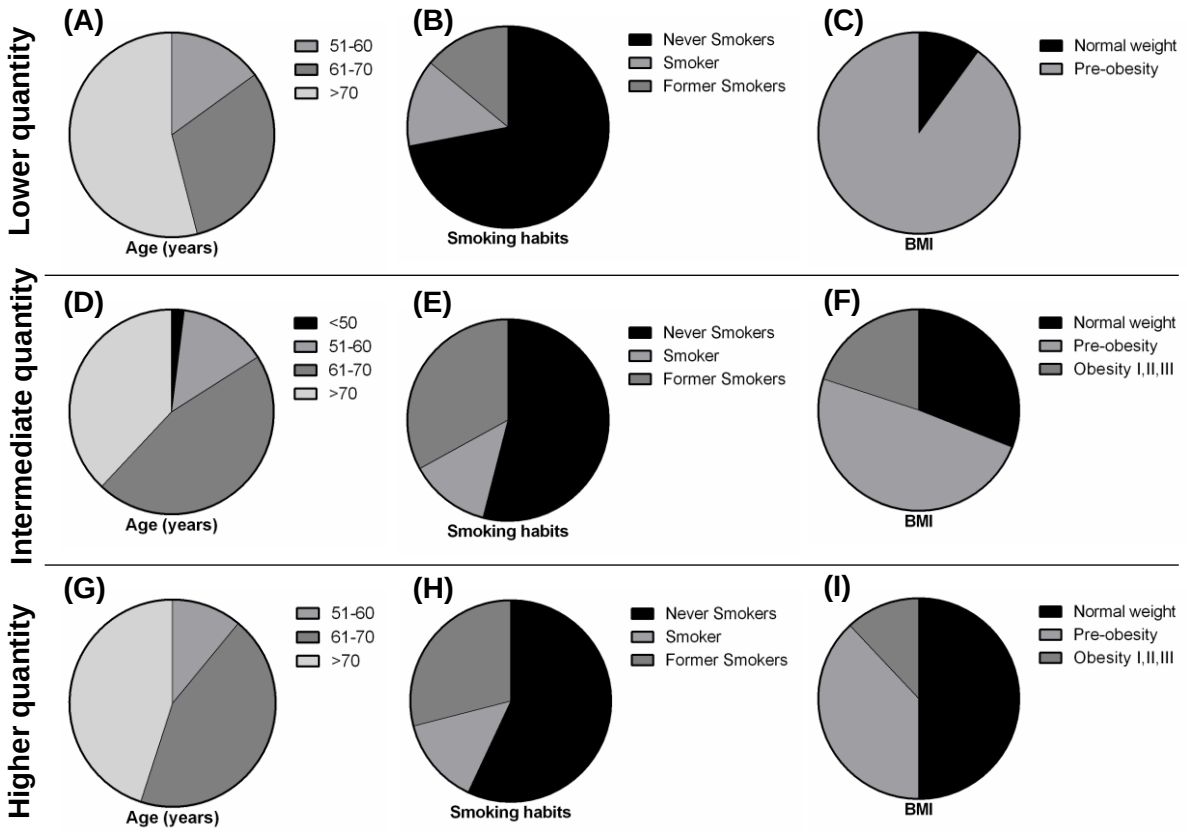


***Figure 7***. Percentage of individuals from the three PC groups distributed by age (**7 A-D-G**), smoking habits (**7 B-E-H**) and BMI classes (**7 C-F-I**), respectively. **7 A-B-C**: Lower quantity group; **7 D-E-F**: Intermediate quantity group; **7 G-H-I**: Higher quantity group.

## 4.3 - *SRY* DNA relative quantification analysis in the Stroke Group

In the stroke group, regarding the *SRY* DNA relative quantity detected in blood cells, we could not find any statistically significant difference when comparing the control group with the stroke group (*P*=0.057) (Figure 8).

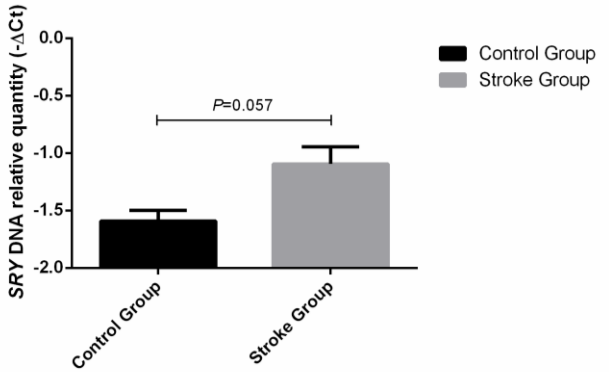


***Figure 8.*** *SRY* DNA relative quantity (-ΔCt) detected in blood cells, compared between control group and stroke group (mean ± standard error of mean).

Moreover, we also could not observe any statistically significant differences in the *SRY* DNA relative quantity according to age (*P*=0.317) (Figure 9 A), hypertension (*P*=0.262) (Figure 9 B), TACI (*P*=0.695) (Figure 9 C), smoking habits (*P*= 0.513) (Figure 9 D) and BMI (*P*=0.558) (Figure 9 E) within stroke group.

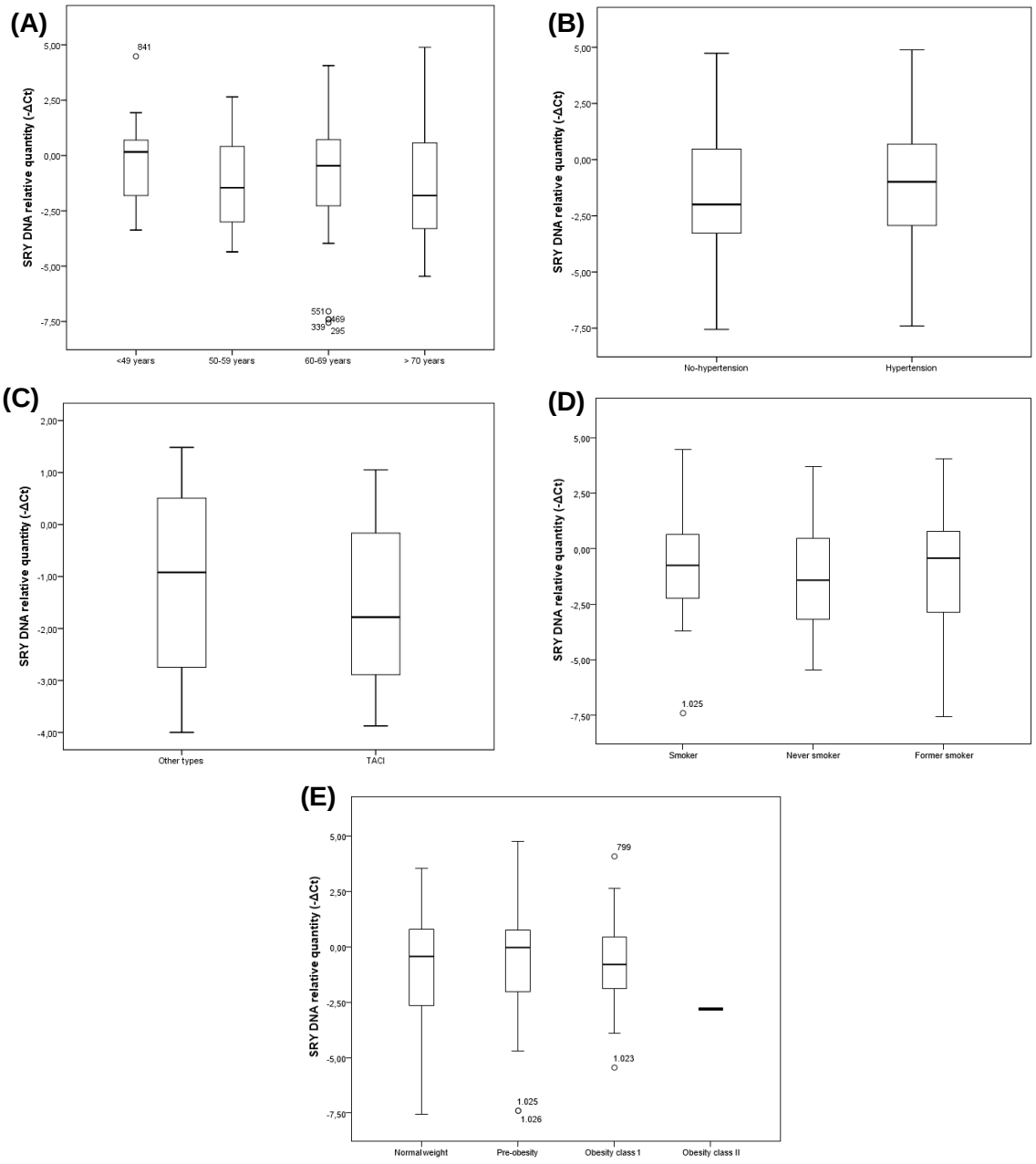


***Figure 9***. Dispersion of individuals from stroke group per age, hypertension, TACI, smoking habits and BMI classes and according to *SRY* DNA relative quantity (-ΔCt). **9 A:** Dispersion of individuals according to age, below 49, between 50 to 59, between 60 to 69 and above 70 years old, respectively (*P*=0.317); **9 B:** Dispersion of individuals according to hypertension and no-hypertension, respectively (*P*=0.262); **9 C:** Dispersion of individuals according to TACI and other types, respectively (*P*=0.695); **9 D:** Dispersion of individuals according to smoking habits, smoker, never smoker and former smoker, respectively (*P*=0.513); **9 E:** Dispersion of individuals according to BMI classes, normal weight, pre-obesity, obesity class I and II, respectively (*P*=0.558). (mean ± standard deviation).

# DISCUSSION

# 5 – DISCUSSION

The aging process is characterized by an irreversible and gradual deterioration of tissue and organ functions due to the action of intrinsic and extrinsic factors (73). These factors have an influence on the aging rate, thus, increasing the susceptibility to age-related diseases (81). LOY in blood increases through the lifespan of the individual and it is a normal aging process (3, 5, 6). This age-dependent loss increased the need to study this molecular alteration in age-related diseases development, such as cancer and cardiovascular diseases (5, 24, 25, 28). The association of LOY with the aging process and age-related diseases suggest its potential to become a biomarker of biological age (4, 6). This study consisted on a qPCR analysis of *SRY* DNA relative quantity in peripheral blood samples as a relative measure of the occurrence of LOY. Consequently, less *SRY* DNA relative quantity detection in blood cells suggests a higher degree of LOY on those cells.

Regarding our healthy population, correlation between *SRY* DNA relative quantity and the chronological age of the individuals was not observed. Several studies have reported the increase of LOY with the subject's age (2, 6, 71, 106-108). However, recent findings suggest that LOY is not equally increasing throughout chronological age (6, 25). In fact, Erikka Loftfield and colleagues have demonstrated that LOY's percentage remained negligible until 50 years old and it only started to be detected after this age in individuals from the UK (25). Similarly, Dumanski and co-workers have demonstrated an accumulation of LOY's percentage in European men older than 65 years old (6). However, until this age, LOY remained stable affecting around 7% of the blood cells (6). LOY's percentage started to increase after 65 years old, affecting 14%, 18% and 20% of blood cells in individuals with ages between 66 to 75, 76 to 85 and older than 86 years old, respectively (6). Additionally, Silva Veiga and colleagues reported a lack of correlation of LOY with age while studying healthy Brazilian individuals ranging from 28 to 68 years old (109). In fact, our study included individuals between 18 and 64 years old but the proportion of individuals below 50 years was higher. This could explain the lack of correlation between *SRY* DNA relative quantity detected and chronological age. Future studies with a higher number of healthy individuals above 50 years old would help to better explore the correlation of *SRY* DNA quantity levels with chronological age. Moreover, in this study, we could observe that men with the same chronological age presented different *SRY* DNA relative quantity. In fact, individuals with the same chronological age can present different aging rates, culminating in a different biological

age (72). Dumanski and co-workers have already suggested that the occurrence of LOY is a better predictor of biological age than chronological age (6).

Biological age reflects the health state of cells, tissues and organs in an individual which can be different among individuals with the same chronological age (72). Intrinsic and extrinsic factors can deregulate processes and increase oxidative stress, telomeric shortening, DNA damage, mitochondrial disfunction, chromatin disruption, inflammation status, epigenetic deregulation and oncogenic activation (73). Among the extrinsic factors that can influence biological age are the smoking habits and the BMI. The healthy population in this study did not show differences in *SRY* DNA relative quantity according to smoking habits and BMI classes. Although current smoking has been associated with higher LOY occurrence, the smoking effect on LOY tends to disappear years after the individual stops smoking, weakening the association between smoking and increased LOY after smoking cessation (29). This transient effect of smoking could explain the similar distribution between former smokers and never smokers. Moreover, regarding individuals who smoke, in the future, information about cigarettes per day or duration of smoking may help to understand the similar distribution observed in this group. Regarding BMI, Dumanski and colleagues have reported no statistical differences between LOY levels and BMI among individuals from three different cohort studies (29). On the other hand, Erikka Loftfield and co-workers have found that individuals with a normal BMI range had significantly higher proportion of LOY in blood cells when comparing with obese class I and obese class II and III (25). However, since the authors found a higher percentage of current smokers in individuals with normal BMI range, they suggested that the smoking habit increased LOY in normal BMI range (25). The knowledge about the impact of higher BMI on DNA damage, chromosomal instability and consequent aneuploidy is still poorly understood and no specific mechanism has yet been stablished (110). Therefore, further research should consider increasing the number of analysed individuals in the healthy population and adding information about the percentage of body fat as well as cigarettes per day since they would help to further elucidate LOY's associations with each BMI class and smoking habit, respectively.

PC and strokes are age-related diseases which are among the main causes of death in the world (85, 95). In the present study, we observed a higher loss of *SRY* DNA relative quantity in the individuals diagnosed with PC comparing with healthy individuals. Remarkably, LOY has been acknowledged as one of the most frequent cytogenic abnormalities in human PC (17, 58, 111-113). Perinchery and co-workers have studied radical prostatectomy specimens from American men and have demonstrated a loss of MSY genes associated with PC development, in which *SRY* and *BPY2* genes were the

most frequently deleted (58). Noveski and colleagues have found an association between increased LOY in blood cells and a higher risk of PC in male individuals from Republic of Macedonia (28). Since LOY is related to the aging process, an increased LOY beyond the age-dependent loss would reflect an acceleration of aging (4). Therefore, this accelerated aging would increase the susceptibility for age-related diseases, such as PC, and explain the lower *SRY* DNA relative quantity found in individuals with PC diagnosis that we studied.

Moreover, in the present study, we also observed that, in the PC subgroup with patients who presented the lowest *SRY* DNA relative quantity, there was a higher percentage of individuals with advanced disease, characterized by a higher percentage of individuals with advanced stage, Gleason score above eight and PSA levels above 10 ng/mL. According with European Association of Urology, these three factors combined confer a high risk of advanced PC which indicates a higher PC aggressiveness and a worse prognosis (87). In fact, LOY has already been associated with PC progression (17, 58, 111). Satoru Takahashi and co-workers obtained PC tissue from radical prostatectomies and verified that PC tissue with higher occurrence of LOY had higher percentages of cells with T3a and greater stages (111). Similarly, Perinchery and colleagues have also studied radical prostatectomy specimens and found that cells with higher Gleason scores (greater than eight) and advanced stages (greater than T3) had a higher loss frequency of most analysed genes (58). According with Perinchery and co-workers, they could not observe deletions of MSY genes in tumours in T1 stages with Gleason scores between 2 and 4 (58). In fact, Jordan and collaborators have found *SRY* deletions in BPH tissue in the presence of malignant prostate tissue (17). However, the authors reported that without malignant prostate tissue, BPH tissue did not present deletions (17). Therefore, suggesting that malignant transformation of BPH was due to the loss of Y chromosome content in the benign tumour tissue (17). Moreover, the introduction of an exogenous human Y chromosome in PC-3 cells (PC cell line lacking the Y chromosome) resulted in hybrid PC-3 cells that were not able to develop PC *in vivo* (93). Vijayakuma and co-workers attributed this tumour suppression activity mostly to the short arm of the Y chromosome, including *SRY* and *ZFY* as candidate genes, which are among the most frequently reported deleted in this type of cancer (93). More recently, the loss of another Y chromosome gene, *KDM5D*, has also been linked to a more aggressive PC (113). KDM5D interacts with histone methylation marks, epigenetically altering expression of cell cycle genes, its loss accelerated cell entry in mitosis (113). The higher loss of *KDM5D* and the observed decrease in its expression was attributed to LOY in PC cells that restrictedly presented higher Gleason scores (113). The fact that Y chromosome

losses are more frequently seen in advanced stages of PC development could explain why we found a higher percentage of individuals with advanced disease in the PC subgroup with the lowest *SRY* DNA relative quantity.

In addition to higher percentages of advanced disease, the lower quantity group, presented a higher percentage of individuals older than 61 years old (85%), never smokers (72%) and of pre-obese individuals (90%). LOY significantly increases with age after 50 years old and its occurrence is higher when cancer is present (5, 28). Therefore, a higher percentage of individuals older than 61 years old is in accordance with the higher occurrence of LOY and the presence of PC in the lower quantity group. Regarding BMI, we observed a higher percentage of pre-obese individuals in the lower quantity group. In fact, obesity contributes to increased oxidative stress, and Wright and collaborators have shown that higher BMI and adult weight gain increased the risk of dying from PC (87, 114).

Considering our results, we also found no statistically significant differences regarding *SRY* DNA relative quantity between individuals diagnosed with stroke and healthy individuals. Haitjema and co-workers reported that occurrence of LOY conferred a higher risk of cardiovascular endpoints (including ischemic and haemorrhagic strokes) in individuals after carotid endarterectomy during a 3-year follow-up (24). On the other hand, Erikka Loftfield and colleagues have reported higher LOY levels in individuals from the UK who self-reported previous events of stroke or ischemic heart attack (25). However, they have studied 11 628 men who self-reported these two types of diagnosis and, remarkably, they only found detectable LOY in 4% of them (25). The fact that previous studies have reported a higher occurrence of LOY only in a small percentage of the analysed individuals could explain the lack of an association between less *SRY* DNA relative quantity and stroke diagnosis in our population.

Furthermore, in individuals diagnosed with stroke, there were no statistically significant differences in the *SRY* DNA relative quantity according to age, hypertension, smoking habits, BMI class and TACI. Haitjema and co-workers studied the association between LOY and cardiovascular endpoints in patients after carotid endarterectomy during a 3 year follow-up and they have found statistical significant differences in LOY according to age (24). However, the same authors did not find association between LOY and hypertension, smoking habits and BMI classes (24). The lack of differences in *SRY* DNA relative quantity also according to TACI type of diagnosis could be due to the fact that only 18% of our group of individuals diagnosed with stroke were diagnosed with this type. Since TACI

is less frequently diagnosed, futures studies should take this in consideration and an effort in increasing the number of subjects with TACI stroke type diagnosis shall be made.

The present study was important to demonstrate the efficiency and usefulness of *SRY* DNA relative quantity as a biomarker to study LOY occurrence in a biological fluid that is easier to access than other more invasive procedures, such as a biopsy. Even though we did not find an association between *SRY* DNA relative quantity and age, smoking habits and BMI in the healthy population, we found an association with PC development, which, reinforces the potential of *SRY* DNA relative quantity detected in peripheral blood cells as a biomarker of age-related diseases, such as PC.

# CONCLUSIONS AND FUTURE PERSPECTIVES

# 6 – CONCLUSIONS AND FUTURE PERSPECTIVES

LOY has been a current object of investigation with applications in different health fields. In this study, LOY occurrence was quantified based on a qPCR method in which *SRY* DNA relative quantity detected in blood cells was relatively quantified. Based on this *SRY* DNA relative quantity, it was possible to compare differences in LOY occurrence in blood cells of healthy men and in comparison with men diagnosed with PC or stroke. There was no correlation between *SRY* DNA relative quantity and chronological age in healthy individuals. However, we found an association of *SRY* DNA relative quantity with PC, an age-related disease and the second most frequent cancer in men. Moreover, among the patients diagnosed with PC, the group who had less amount of *SRY* DNA relative quantity (higher occurrence of LOY) had a higher percentage of individuals with advanced PC, older than 61 years old, never smokers and pre-obese. However, we could not find any differences in *SRY* DNA relative quantity between healthy individuals and individuals with stroke diagnosis.

The present study is the first to use a relative quantification of *SRY* DNA quantity levels as an indirect measure of LOY occurrence in blood cells. The qPCR method presents several advantages to the quantification of LOY occurrence in blood cells, since it is a fast, sensible, specific and less expensive method when compared to other techniques such as karyotyping and Fluorescence in situ hybridization (FISH). Moreover, since *SRY* DNA relative quantity reflects an acceleration of the aging process, it has the potential to become a biomarker of age-related diseases. Therefore, this relationship with aging acceleration mirrors an increase on biological age, which also reinforces the potential of *SRY* DNA relative quantity to become a biomarker of biological age.

Biomarkers of biological age should predict the aging rate and similarly with the aging process must be influenced by intrinsic and extrinsic factors that accelerate aging (72). Biological age estimation has main applications in Legal Medicine, such as in work damage evaluations and indemnity processes, providing several advantages. In fact, work stress and environmental hazards at the workplace are linked to psychological distress and a variety of chronic diseases. Increased LOY occurrence in blood cells of male workers beyond the normal age-dependent loss can be helpful, providing evidence that their work environment is increasing their biological age and, consequently, their risk of developing age-related diseases. Another advantage would be the application of these biomarkers in the study of life stressors impact on one's health status. Childhood abusive experiences, domestic violence and social/racial discrimination (e.g. bullying) are cases

that cause high levels of stress that can have repercussions on these people's health (115). Moreover, patients with depression, generalized anxiety disorder and post-traumatic stress disorder have higher mortality rates and they also usually die of age-related diseases, such as cerebrovascular diseases and cancer (116). Therefore, the use of *SRY* DNA quantity levels in blood cells as a biological age biomarker would provide useful information to access individual aging rate in this type of situations. However, the use of this information can also have disadvantages to these people. In fact, life insurance companies are already trying to use biological age information to raise their prices based on a person's health status (117). Because of this, the use of biological age must have legal controls to assure that citizen's human rights remain protected. Besides these applications, LOY occurrence may have implications to the study of the Y chromosome. Higher LOY occurrence would decrease the detection of Y chromosome content, implying, for example, a lower detection of Y-STRs and Y-SNPs alleles during forensic analysis. However, there are no studies about LOY's association with failure in Y-STRs and Y-SNPs analysis. Future studies should consider exploring the implications that LOY occurrence could have in the analysis of Y chromosome content.

The present study has complemented other studies regarding LOY occurrence in blood associated with age-related diseases. We demonstrated the potential of *SRY* DNA relative quantity detection in blood cells to become a biomarker of PC development. Future studies in other age-related diseases in association with LOY occurrence would support and further explore the associations of this biomarker with aging with applications on both health and legal medicine.

# REFERENCES

# 7 – REFERENCES

1. Dumanski JP, Sundström J, Forsberg LA. Loss of Chromosome Y in Leukocytes and Major Cardiovascular Events. Medical Genetics Cardiac and Cardiovascular Systems. 2017;10(4).

2. Jacobs PA, Brunton M, Court Brown WM, Doll R, Goldstein H. Change of human chromosome count distribution with age: evidence for a sex differences. Nature. 1963;197:1080-1.

3. Guttenbach M, Koschorz B, Bernthaler U, Grimm T, Schmid M. Sex chromosome loss and aging: in situ hybridization studies on human interphase nuclei. American journal of human genetics. 1995;57(5):1143-50.

4. Forsberg LA. Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. Human Genetics. 2017;136(5):657-63.

5. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, et al. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. Nature Genetics. 2014;46(6):624-8.

6. Dumanski JP, Lambert JC, Rasi C, Giedraitis V, Davies H, Grenier-Boley B, et al. Mosaic Loss of Chromosome Y in Blood Is Associated with Alzheimer Disease. American journal of human genetics. 2016;98(6):1208-19.

7. Richard SM, Knuutila S, Peltomaki P, Bianchi MS, Bianchi NO. Y chromosome instability in lymphoproliferative disorders. Mutation research. 2003;525(1-2):103-7.

8. INE, PORDATA. Esperança média de vida à nascença por sexo [Internet]. 2019 [cited 07-08-2019]. Available from: https://www.pordata.pt/Portugal/Esperan%C3%A7a+de+vida+%C3%A0+nascen%C3%A7a+total+e+por+sexo+(base+tri%C3%A9nio+a+partir+de+2001)-418.

9. Kido T, Lau YF. Roles of the Y chromosome genes in human cancers. Asian journal of andrology. 2015;17(3):373-80.

10. Brunelli M, Eble JN, Zhang S, Martignoni G, Cheng L. Gains of chromosomes 7, 17, 12, 16, and 20 and loss of Y occur early in the evolution of papillary renal cell neoplasia: a fluorescent in situ hybridization study. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2003;16(10):1053-9.

11. Lau Y-FC. Gonadoblastoma, Testicular and Prostate Cancers, and the TSPY Gene. American journal of human genetics. 1999;64(4):921-7.

12. Hunter S, Gramlich T, Abbott K, Varma V. Y chromosome loss in esophageal carcinoma: an in situ hybridization study. Genes Chromosomes Cancer. 1993;8(3):172-7.

13. Castedo S, Correia C, Gomes P, Seruca R, Soares P, Carneiro F, et al. Loss of Y chromosome in gastric carcinoma. Fact or artifact? Cancer genetics and cytogenetics. 1992;61(1):39-41.

14. Sauter G, Moch H, Wagner U, Novotna H, Gasser TC, Mattarelli G, et al. Y chromosome loss detected by FISH in bladder cancer. Cancer genetics and cytogenetics. 1995;82(2):163-9.

15. Al-Saleem T, Balsara BR, Liu Z, Feder M, Testa JR, Wu H, et al. Renal oncocytoma with loss of chromosomes Y and 1 evolving to papillary carcinoma in connection with gain of chromosome 7. Coincidence or progression? Cancer genetics and cytogenetics. 2005;163(1):81-5.

16. Klatte T, Rao PN, de Martino M, LaRochelle J, Shuch B, Zomorodian N, et al. Cytogenetic profile predicts prognosis of patients with clear cell renal cell carcinoma. Journal of Clinical Oncology. 2009;27(5):746-53.

17. Jordan JJ, Hanlon AL, Al-Saleem TI, Greenberg RE, Tricoli JV. Loss of the short arm of the Y chromosome in human prostate carcinoma. Cancer genetics and cytogenetics. 2001;124(2):122-6.

18. Alers JC, Rochat J, Krijtenburg PJ, Hop WC, Kranse R, Rosenberg C, et al. Identification of genetic markers for prostatic cancer progression. Laboratory Investigation. 2000;80(6):931-42.

19. Minner S, Kilgue A, Stahl P, Weikert S, Rink M, Dahlem R, et al. Y chromosome loss is a frequent early event in urothelial bladder cancer. Pathology. 2010;42(4):356-9.

20. Wallrapp C, Hahnel S, Boeck W, Soder A, Mincheva A, Lichter P, et al. Loss of the Y chromosome is a frequent chromosomal imbalance in pancreatic cancer and allows differentiation to chronic pancreatitis. International journal of cancer. 2001;91(3):340-4.

21. Veiga LC, Bergamo NA, Reis PP, Kowalski LP, Rogatto SR. Loss of Y-chromosome does not correlate with age at onset of head and neck carcinoma: a case-control study. Brazilian Journal of Medical and Biological Research. 2012;45(2):172-8.

22. Park SJ, Jeong SY, Kim HJ. Y chromosome loss and other genomic alterations in hepatocellular carcinoma cell lines analyzed by CGH and CGH array. Cancer genetics and cytogenetics. 2006;166(1):56-64.

23. Stahl PR, Kilgue A, Tennstedt P, Minner S, Krohn A, Simon R, et al. Y chromosome losses are exceedingly rare in prostate cancer and unrelated to patient age. The Prostate. 2012;72(8):898-903.

24. Haitjema S, Kofink D, van Setten J, van der Laan SW, Schoneveld AH, Eales J, et al. Loss of Y Chromosome in Blood Is Associated With Major Cardiovascular Events During Follow-Up in Men After Carotid Endarterectomy. Circulation: Cardiovascular Genetics. 2017;10(4):e001544.

25. Loftfield E, Zhou W, Graubard BI, Yeager M, Chanock SJ, Freedman ND, et al. Predictors of mosaic chromosome Y loss and associations with mortality in the UK Biobank. Scientfic Reports. 2018;8(1):12316.

26. Wong JYY, Margolis HG, Machiela M, Zhou W, Odden MC, Psaty BM, et al. Outdoor air pollution and mosaic loss of chromosome Y in older men from the Cardiovascular Health Study. Environment international. 2018;116:239-47.

27. Tang D, Han Y, Lun Y, Jiang H, Xin S, Duan Z, et al. Y chromosome loss is associated with age-related male patients with abdominal aortic aneurysms. Clinical interventions in aging. 2019;14:1227-41.

28. Noveski P, Madjunkova S, Sukarova Stefanovska E, Matevska Geshkovska N, Kuzmanovska M, Dimovski A, et al. Loss of Y Chromosome in Peripheral Blood of Colorectal and Prostate Cancer Patients. PLoS One. 2016;11(1):e0146264.

29. Dumanski JP, Rasi C, Lonn M, Davies H, Ingelsson M, Giedraitis V, et al. Mutagenesis. Smoking is associated with mosaic loss of chromosome Y. Science. 2015;347(6217):81-3.

30. Holland AJ, Cleveland DW. Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nature reviews Molecular cell biology. 2009;10(7):478-87.

31. Gilgenkrantz S. [Nettie Maria Stevens (1861-1912)]. Medical Sciences (Paris). 2008;24(10):874-8.

32. Painter TS. Studies in mammalian spermatogenesis. II. The spermatogenesis of man. Journal of Experimental Zoology. 1923;37(3):291–336.

33. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature. 2003;423(6942):825-37.

34. Bruce T. Lahn DCP. Functional Coherence of the Human Y Chromosome. Science. 1997;278.

35. Jangravi Z, Alikhani M, Arefnezhad B, Sharifi Tabar M, Taleahmad S, Karamzadeh R, et al. A fresh look at the male-specific region of the human Y chromosome. Journal of proteome research. 2013;12(1):6-22.

36. Singh NP, Madabhushi SR, Srivastava S, Senthilkumar R, Neeraja C, Khosla S, et al. Epigenetic profile of the euchromatic region of human Y chromosome. Nucleic Acids Research. 2011;39(9):3594-606.

37. Bachtrog D, Charlesworth B. Towards a complete sequence of the human Y chromosome. Genome biology. 2001;2(5):Reviews1016.

38. Cooke HJ, Brown WR, Rappold GA. Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. Nature. 1985;317(6039):687-92.

39. Freije D, Helms C, Watson MS, Donis-Keller H. Identification of a second pseudoautosomal region near the Xq and Yq telomeres. Science. 1992;258(5089):1784-7.

40. Helena Mangs A, Morris BJ. The Human Pseudoautosomal Region (PAR): Origin, Function and Future. Current genomics. 2007;8(2):129-36.

41. Jobling MA, Tyler-Smith C. New uses for new haplotypes the human Y chromosome, disease and selection. Trends in Genetics. 2000;16(8):356-62.

42. Kumar A, Sharma M. Basics of Human Andrology: A Textbook: Springer Singapore; 2017.

43. Charchar FJ, Svartman M, El-Mogharbel N, Ventura M, Kirby P, Matarazzo MR, et al. Complex events in the evolution of the human pseudoautosomal region 2 (PAR2). Genome research. 2003;13(2):281-6.

44. Hinch AG, Altemose N, Noor N, Donnelly P, Myers SR. Recombination in the human Pseudoautosomal region PAR1. PLoS genetics. 2014;10(7):e1004503.

45. Veerappa AM, Padakannaya P, Ramachandra NB. Copy number variation-based polymorphism in a new pseudoautosomal region 3 (PAR3) of a human X-chromosome-transposed region (XTR) in the Y chromosome. Functional & integrative genomics. 2013;13(3):285-93.

46. Ginalski K, Rychlewski L, Baker D, Grishin NV. Protein structure prediction for the male-specific region of the human Y chromosome. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(8):2305-10.

47. Maan AA, Eales J, Akbarov A, Rowland J, Xu X, Jobling MA, et al. The Y chromosome: a blueprint for men's health? European Journal of Human Genetics. 2017;25(11):1181-8.

48. Tilford CA, Kuroda-Kawaguchi T, Skaletsky H, Rozen S, Brown LG, Rosenberg M, et al. A physical map of the human Y chromosome. Nature. 2001;409(6822):943-5.

49. Page DC. Hypothesis: a Y-chromosomal gene causes gonadoblastoma in dysgenetic gonads. Development. 1987;101 Suppl:151-5.

50. Sun SL, Horino S, Itoh-Nakadai A, Kawabe T, Asao A, Takahashi T, et al. Y chromosome-linked B and NK cell deficiency in mice. The Journal of Immunology. 2013;190(12):6209-20.

51. Ely D, Underwood A, Dunphy G, Boehme S, Turner M, Milsted A. Review of the Y chromosome, Sry and hypertension. Steroids. 2010;75(11):747-53.

52. Blanche Capel RL-B. The Sry Gene and Sex Determination in Mammals. Advances in Developmental Biology (1992). 1993;2:1 - 35.

53. Quintana-Murci L, Fellous M. The Human Y Chromosome: The Biological Role of a "Functional Wasteland". Journal of Biomedicine and Biotechnology. 2001;1(1):18-24.

54. Binder G. Short stature due to SHOX deficiency: genotype, phenotype, and therapy. Hormone research in paediatrics. 2011;75(2):81-9.

55. Gibson CW. The Amelogenin Proteins and Enamel Development in Humans and Mice. Journal of oral biosciences. 2011;53(3):248-56.

56. Oram SW, Liu XX, Lee TL, Chan WY, Lau YF. TSPY potentiates cell proliferation and tumorigenesis by promoting cell cycle progression in HeLa and NIH3T3 cells. BMC cancer. 2006;6:154.

57. Yuan X, Lu ML, Li T, Balk SP. SRY interacts with and negatively regulates androgen receptor transcriptional activity. The Journal of biological chemistry. 2001;276(49):46647-54.

58. Perinchery G, Sasaki M, Angan A, Kumar V, Carroll P, Dahiya R. Deletion of Y-chromosome specific genes in human prostate cancer. The Journal of urology. 2000;163(4):1339-42.

59. Kayser M. Forensic use of Y-chromosome DNA: a general overview. Human Genetics. 2017;136(5):621-35.

60. Prinz M, Sansone M. Y chromosome-specific short tandem repeats in forensic casework. Croatian medical journal. 2001;42(3):288-91.

61. Butler JM. Chapter 13 - Y-Chromosome DNA Testing. Advanced Topics in Forensic DNA Typing: Methodology. National Institute of Standards and Technology, Gaithersburg, Maryland, USA: Elsevier Inc.; 2012. p. 371 - 403.

62. El-Hafez SHE-AFA. Paternity testing and forensic DNA typing by multiplex STR analysis using ABI PRISM 310 Genetic Analyzer. Journal of Genetic Engineering and Biotechnology. 2012;10(1):101 - 12.

63. Genetics Home Reference. What are single nucleotide polymorphisms (SNPs)? [Internet]. 2018 [cited 25-05-2019]. Available from: https://ghr.nlm.nih.gov/primer/genomicresearch/snp.

64. Gopinath S, Zhong C, Nguyen V, Ge J, Lagace RE, Short ML, et al. Developmental validation of the Yfiler((R)) Plus PCR Amplification Kit: An enhanced Y-STR multiplex for casework and database applications. Forensic science international Genetics. 2016;24:164-75.

65. Hammer MF. A recent insertion of an alu element on the Y chromosome is a useful marker for human population studies. Molecular biology and evolution. 1994;11(5):749-61.

66. Phillips C. Forensic genetic analysis of bio-geographical ancestry. Forensic science international Genetics. 2015;18:49-65.

67. Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, et al. Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications. The American Journal of Human Genetics. 2010;87(3):341-53.

68. Knijff P. Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. American journal of human genetics. 2000;67(5):1055-61.

69. NanoLabs P. Snapshot Genetic Genealogy Snapshot DNA analysis: Parabon Computation, Inc. [Internet]. 2018 [Cited 27-05-2019]. Available from: https://snapshot.parabon-nanolabs.com/intro.

70. Wang YC, Ma XY, Sun XF, Xian JJ, Li SY, He WZ, et al. [Alleles dropout patterns of Y-short tandem repeats in infertile males with Y chromosome microdeletions]. Yi chuan = Hereditas. 2019;41(3):243-53.

71. Zhou W, Machiela MJ, Freedman ND, Rothman N, Malats N, Dagnall C, et al. Mosaic loss of chromosome Y is associated with common variation near TCL1A. Nature Genetics. 2016;48(5):563-8.

72. Jylhava J, Pedersen NL, Hagg S. Biological Age Predictors. EBioMedicine. 2017;21:29-36.

73. Dodig S, Cepelak I, Pavic I. Hallmarks of senescence and aging. Biochemia medica (Zagreb). 2019;29(3):030501.

74. Nicita-Mauro V, Basile G, Maltese G, Nicita-Mauro C, Gangemi S, Caruso C. Smoking, health and ageing. Immunity & ageing : I & A. 2008;5:10-.

75. Strimbu K, Tavel JA. What are biomarkers? Current opinion in HIV and AIDS. 2010;5(6):463-6.

76. Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, et al. The microRNA spectrum in 12 body fluids. Clinical chemistry. 2010;56(11):1733-41.

77. Xia X, Chen W, McDermott J, Han J-DJ. Molecular and phenotypic biomarkers of aging. F1000Research. 2017;6:860-.

78. Kang YG, Suh E, Lee JW, Kim DW, Cho KH, Bae CY. Biological age as a health index for mortality and major age-related disease incidence in Koreans: National Health Insurance Service - Health screening 11-year follow-up study. Clinical interventions in aging. 2018;13:429-36.

79. Franceschi C, Garagnani P, Morsiani C, Conte M, Santoro A, Grignolio A, et al. The Continuum of Aging and Age-Related Diseases: Common Mechanisms but Different Rates. Frontiers in medicine. 2018;5:61-.

80. Niccoli T, Partridge L. Ageing as a Risk Factor for Disease. 2012;22(17):R741-R52.

81. Jaul E, Barron J. Age-Related Diseases and Clinical and Public Health Implications for the 85 Years Old and Over Population. Frontiers in public health. 2017;5:335.

82. Holmes RI, Keating MJ, Cork A, Trujillo JM, McCredie KB, Freireich EJ. Loss of the Y chromosome in acute myelogenous leukemia: a report of 13 patients. Cancer genetics and cytogenetics. 1985;17(3):269-78.

83. Hesketh R. Introduction to Cancer Biology: Cambridge University Press; 2013.

84. Society AC. Global Cancer, Facts & Figures 4th Edition. Atlanta; 2018.

85. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.

86. International Agency for Research on Cancer. The Global Cancer Observatory - Portugal. GLOBOCAN 2018: WORLD HEALTH ORGANIZATION [Internet]. 2018 [cited 10-08-2019]. Available from: https://gco.iarc.fr/today/data/factsheets/populations/620-portugal-fact-sheets.pdf.

87. Grupo Português Génito-Urinário Sociedade Portuguesa de Oncologia. Tumores Malignos - Tracto génito-urinário. edição, editor: Cor Comum - Serviços Gráficos, Lda; 2016.

88. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. European urology. 2017;71(4):618-29.

89. Teixeira AL. Vias de Sinalização EGF/EGGFR-HER2 e TGFB1/TGFBRII no Desenvolvimento de Neoplasias Malignas Urológicas: Epidemiologia Molecular e Farmacogenómica Porto: Universidade do Porto; 2014.

90. Patel R, Khalifa AO, Isali I, Shukla S. Prostate cancer susceptibility and growth linked to Y chromosome genes. Frontiers in bioscience (Elite edition). 2018;10:423-36.

91. Wallis CJD, Nam RK. Prostate Cancer Genetics: A Review. EJIFCC. 2015;26(2):79-91.

92. Yadav SK, Kumari A, Ali S. Fate of the human Y chromosome linked genes and loci in prostate cancer cell lines DU145 and LNCaP. BMC genomics. 2013;14(1):323.

93. Vijayakumar S, Garcia D, Hensel CH, Banerjee M, Bracht T, Xiang R, et al. The human Y chromosome suppresses the tumorigenicity of PC-3, a human prostate cancer cell line, in athymic nude mice. Genes Chromosomes Cancer. 2005;44(4):365-72.

94. Lau YC, Li Y, Kido T. Battle of the sexes: contrasting roles of testis-specific protein Y-encoded (TSPY) and TSPX in human oncogenesis. Asian journal of andrology. 2018.

95. Eurostat Statistics Explained. Causes of death statistics [Internet]. 2018 [cited 11-08-2019]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Causes_of_death_statistics.

96. Direção-Geral da Saúde, Institute for Health Metrics and Evaluation, Portugal. The Nation's Health 1990–2016: An overview of the Global Burden of Disease Study [Internet]. 2016 [cited 15-08-2019]. Available from: https://www.dgs.pt/documentos-e-publicacoes/portugal-the-nations-health-19902016-an-overview-of-the-global-burden-of-disease-study-2016-results.aspx.

97. Mafalda Sousa-Uva CMD. Prevalência de Acidente Vascular Cerebral na população portuguesa: dados da amostra ECOS 2013. In: Departamento de Epidemiologia I, editor. Instituto Nacional de Saúde Doutor Ricardo Jorge www.insa.pt; 2014. p. 3.

98. Parahoo K, Thompson K, Cooper M, Stringer M, Ennis E, McCollam P. Stroke: awareness of the signs, symptoms and risk factors--a population-based survey. Cerebrovascular diseases (Basel, Switzerland). 2003;16(2):134-40.

99. Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ, Culebras A, et al. An Updated Definition of Stroke for the 21st Century. Stroke. 2013;44(7):2064-89.

100. Parmar, Paresh. Stroke: Classification and diagnosis. Pharmaceutical Journal. 2018; 10(1):1-15.

101. Yang Y, Wang A, Zhao X, Wang C, Liu L, Zheng H, et al. The Oxfordshire Community Stroke Project classification system predicts clinical outcomes following intravenous thrombolysis: a prospective cohort study. Therapeutics and clinical risk management. 2016;12:1049-56.

102. Case LK, Teuscher C. Y genetic variation and phenotypic diversity in health and disease. Biology of sex differences. 2015;6:6.

103. Ellis JA, Stebbing M, Harrap SB. Association of the human Y chromosome with high blood pressure in the general population. Hypertension (Dallas, Tex : 1979). 2000;36(5):731-3.

104. Voskarides K, Hadjipanagi D, Papazachariou L, Griffin M, Panayiotou AG. Evidence for contribution of the y chromosome in atherosclerotic plaque occurrence in men. Genetic testing and molecular biomarkers. 2014;18(8):552-6.

105. Tian Y, Stamova B, Jickling GC, Xu H, Liu D, Ander BP, et al. Y chromosome gene expression in the blood of male patients with ischemic stroke compared with male controls. Gender medicine. 2012;9(2):68-75.e3.

106. Persani L, Bonomi M, Lleo A, Pasini S, Civardi F, Bianchi I, et al. Increased loss of the Y chromosome in peripheral blood cells in male patients with autoimmune thyroiditis. Journal of autoimmunity. 2012;38(2-3):J193-6.

107. Forsberg LA, Gisselsson D, Dumanski JP. Mosaicism in health and disease - clones picking up speed. Nature Reviews Genetics. 2017;18(2):128-42.

108. Lleo A, Oertelt-Prigione S, Bianchi I, Caliari L, Finelli P, Miozzo M, et al. Y chromosome loss in male patients with primary biliary cirrhosis. Journal of autoimmunity. 2013;41:87-91.

109. Silva Veiga LC, Bérgamo NA, Reis PP, Kowalski LP, Rogatto SR. Loss of Y-chromosome does not correlate with age at onset of head and neck carcinoma: a case-control study. Brazilian Journal of Medical and Biological Research. 2012;45(2):172-8.

110. Setayesh T, Nersesyan A, Misik M, Ferk F, Langie S, Andrade VM, et al. Impact of obesity and overweight on DNA stability: Few facts and many hypotheses. Mutation research. 2018;777:64-91.

111. Takahashi S, Alcaraz A, Brown JA, Borell TJ, Herath JF, Bergstralh EJ, et al. Aneusomies of chromosomes 8 and Y detected by fluorescence in situ hybridization are prognostic markers for pathological stage C (pt3N0M0) prostate carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research. 1996;2(1):137-45.

112. Tricoli JV. Y chromosome enumeration in touch preparations from 42 prostate tumors by interphase fluorescence in situ hybridization analysis. Cancer genetics and cytogenetics. 1999;111(1):1-6.

113. Komura K, Yoshikawa Y, Shimamura T, Chakraborty G, Gerke TA, Hinohara K, et al. ATR inhibition controls aggressive prostate tumors deficient in Y-linked histone demethylase KDM5D. The Journal of clinical investigation. 2018;128(7):2979-95.

114. Wright ME, Chang SC, Schatzkin A, Albanes D, Kipnis V, Mouw T, et al. Prospective study of adiposity and weight change in relation to prostate cancer incidence and mortality. Cancer. 2007;109(4):675-84.

115. Thoits PA. Stress and health: major findings and policy implications. Journal of health and social behavior. 2010;51 Suppl:S41-53.

116. Shalev I, Moffitt TE, Braithwaite AW, Danese A, Fleming NI, Goldman-Mellor S, et al. Internalizing disorders and leukocyte telomere erosion: a prospective study of depression, generalized anxiety disorder and post-traumatic stress disorder. Molecular psychiatry. 2014;19(11):1163-70.

117. Chang C-C. Adjustable Biological-Age Pricing for the Global Market University in Taichung, Taiwan: Society of Actuaries; 2009.

# APPENDIX

# 8 – APPENDIX

A scientific poster entitled “Loss of chromosome Y (*LOY*): potential implications on health and forensic fields” was presented at the University of Aveiro, in an annual meeting of Biomedical Sciences bachelors.