Discovery and Validation of Sex-Specific Survival Biomarkers in Early-Stage Cutaneous Melanoma

UNIVERSITÀ
DI TORINO

Eirini Chrysanthou
PhD Thesis

A PhD thesis, part of the 35th cycle of the Complex Systems for Quantitative Biomedicine, submitted to the Department of Life Sciences and Systems Biology at the University of Turin

Supervisor: Dr Giovanna Chiorino
Co-supervisors: Prof. Enzo Medico
Prof. Gian Paolo Dotto

July 2024

DISCOVERY AND VALIDATION OF SEX-SPECIFIC SURVIVAL BIOMARKERS IN EARLY-STAGE CUTANEOUS MELANOMA

Eirini Chrysanthou

Dissertation Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

To the PhD school
Complex Systems for Quantitative Biomedicine (35th cycle)
at the department of
Life Sciences and Systems Biology

Dissertation was externally reviewed by three referees:

1. **Dr Joshua B. Rubin** - Department of Pediatrics, Washington University School of Medicine, USA
2. **Dr Jérémie Nsengimana** - Population Health Sciences Institute, Newcastle University, United Kingdom
3. **Dr Irene Orlow** - Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Centre, New York, NY, USA

University of Turin

July 2024

“The facts are always there, as long as we unsee what we always learned to see.” E.C

Thesis Abstract

Background: There are multiple sex disparities across numerous cancers affecting incidence, survival, and treatment response. In cutaneous melanoma (CM), men have a twofold higher probability to develop the disease and a higher death rate compared to females. This project focuses on identifying sex-specific gene expression survival biomarkers for early-stage (stage I-II) CM.

Methods: We used a subset of the Leeds Melanoma Cohort (LMC) microarray dataset with 311 females and 256 males with stage I and II CM, a subset of the TCGA RNA-seq dataset with 28 females and 46 males, as well as CM tissues from 22 females and 21 males recruited at the local hospital. The latter were used to evaluate the expression of selected genes by RT-qPCR and validate the prognostic models. Sex was first used as a covariate in survival analysis and then as an effect modifier, performing multivariable Cox regression analysis on each sex separately and adjusting for age and stage. In females, genes selected by setting the Benjamini-Hochberg adjusted p-value cut-off less at 0.05 (controlling the false discovery rate at 5%) were evaluated for their predictive performance using 10-fold cross-validation and LASSO penalised Cox regression analysis with 8-year censored data. In males, due to no significant results using the adjusted p-value cut-off, genes with raw p-value less than 0.001 and highest hazard ratio were further tested, using 10-fold cross-validation and bidirectional stepwise Cox regression analysis with 8-year censored data. Two sex-specific prognostic models were built and their survival prediction ability was assessed using UNO's concordance (C) index, UNO's area under the curve (AUC) and observed vs expected ratios. We also tested the gene expression variability (GEV) among females and males and compared gene expression between sexes, as well as the immune cell composition using xCell deconvolution method.

Results: Application of previously proposed single or multigene expression survival analysis using sex as a covariate revealed a strong female dependency of the total significant genes. Contrary to expectations, we observed a lower GEV among multiple gene sets and female subgroups of stage I-II CM, in both sun and not-sun exposed normal skin and in nevi. In stage III CMs, GEV was lower in males compared to females, but generally higher than in stage I-II CMs. The sex-stratified survival analysis revealed for the first time in CM a three gene (*UHRF1*, *UBE7*, *HLA-E*) female-specific prognostic signature and a male-specific prognostic

signature consisting of two genes (*SF3B3*, *BEX3*) combined with age and stage. The female gene signature was validated in the RT-qPCR cohort, yielding an UNO C-index of 0.76 and an UNO AUC of 0.80. The male-specific signature gave an UNO C-index of 0.86 and an UNO-AUC of 0.85. A weaker model performance was observed on the TCGA cohort, especially regarding the male signature. In addition, xCell immune cell type enrichment analysis revealed important sex differences in the enrichments of multiple T-cell populations and in their association with survival.

Conclusions: This study highlighted the importance of sex stratification in CM survival analysis and confirmed the statistically significant superior performance of female features, from genes to immune cell subtypes and correlated profiles. It unveils unexpected differences in GEV between the two sexes in early-stage CM and sheds light into the distinct biological pathways and immune cell populations in female and male CMs.

Dedication

This work is a heartfelt dedication to my late father, Chrysanthos Chysanthou. His sudden diagnosis and subsequent passing, three and a half weeks apart, from Glioblastoma Multiforme on May 16, 2021, deeply impacted my journey through this PhD. His unique kind of support was instrumental in bringing me to where I stand today. I am immensely thankful for the invaluable lessons he taught me during his time with us and for the resilience I've developed while facing every day with his permanent absence. His memory remains a guiding light in my life and work.

Acknowledgements

This PhD journey has been a truly life-changing experience, and the work presented would never have been possible without the help I received from many people.

Firstly, I would like to express my sincerest appreciation to my principal supervisor, Dr. Giovanna Chiorino, for trusting and selecting me for this project. Her enormous help, guidance, and support have been invaluable throughout this PhD journey.

I extend my heartfelt gratitude to all my colleagues at the Foundation in Biella, particularly Dr. Emir Sehovic, for the endless brainstorming sessions, immense project support, and unwavering psychological support during the most challenging times of this journey. I would also like to thank Dr. Paola Ostano for her constant readiness to assist with any practical matters, both personal and related to the PhD project. Additionally, I would like to thank Dr. Maurizia Mellograno for her invaluable guidance on the laboratory work involving RT-qPCR, and for her consistently calm, composed, and empowering presence in the lab. Next, I would like to thank Dr. Caterina Peraldo for introducing me to my favourite little village on planet Earth, Rosazza, which has been my inspiration and peaceful spot during these 3.5 years. Finally, I would like to thank Dr. Ilaria Gregnanin and Dr. Francesca Guana for always being ready to help and answer any of my questions.

As an early-stage researcher at the CancerPrev Consortium, which received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 859860 Marie Skłodowska-Curie, I would like to thank all PhD candidates and PIs for the excellent meetings and the useful scientific interactions. More specifically, I would like to thank my second supervisor, Dr. Paolo Dotto, for welcoming me into his lab in Lausanne and allowing me to interact with and learn from his scientific team. I would also like to thank the PhD candidate, Carlos Venturi Ronchi, for all the times he was readily available and willing to help with my project.

Additionally, I would like to thank all the people at the Biella hospital, especially Dr. Elena Biletta and Kejsi Demo, for welcoming me, providing all requested information, and helping me gather all important samples for my analysis.

I would also like to thank all three external referees, Dr. Irene Orlow, Dr. Joshua Rubin, and Dr. Jeremie Nsengimana, for their invaluable corrections and comments that played a very important role in the quality of this work.

Finally, I would like to thank my family and friends, in Cyprus, Italy and France, who have supported and believed in me throughout this journey.

Table of Contents

Discovery and validation of sex-specific survival biomarkers in early-stage Cutaneous melanoma 1

..... 1

Thesis Abstract..... 1

Dedication..... 3

Acknowledgements..... 4

List of Tables..... 8

List of Figures..... 10

List of papers..... 13

Chapter 1: Background and Introduction..... 15

Discrimination of female sex from biomedical studies..... 15

Sex differences in physiology..... 15

Sex differences in pathology..... 16

Sex differences in cancer..... 17

Cutaneous melanoma (CM)..... 19

Brief history of melanoma..... 19
Melanoma epidemiology and risk factors..... 20
Melanoma diagnosis..... 20
Melanoma classification..... 21
Melanoma diagnostic markers..... 23
Melanoma prognosis..... 23
Melanoma prognostic and serological markers..... 24
Molecular characteristics..... 25
Melanoma treatment..... 26
Melanoma sex differences..... 27

Chapter 2: Aims and Objectives..... 28

General Aim of the Thesis..... 28

Specific aims of the thesis..... 28

Chapter 3: Materials and Methods..... 29

Datasets..... 29

Datasets description and characteristics..... 29
Leeds Melanoma Cohort (LMC)..... 30
The Cancer Genome Atlas (TCGA)..... 31
GSE65904..... 31
GSE53118..... 31
Genotype Tissue Expression (GTEx)..... 32
GSE46517 and GSE3189..... 32

Biella primary melanoma cohort ... 32
Reverse-Transcription quantitative Polymerase Chain reaction (RT- qPCR) ... *32*
Statistical analysis ... *33*
Survival analysis ... 33
Validation of predictive models ... 35
Gene expression variability (GEV) ... 35
Correlation analysis ... 35
Principal component analysis (PCA) ... *36*
Immune cell infiltration analysis ... *36*
Functional enrichment analysis ... *36*
Protein-protein interaction Network ... *36*
Human Protein Atlas ... *37*
Melanoma biomarkers found in literature ... *37*
Study flowchart ... **38**
Results ... **39**
Chapter 4: Sex-specific evaluation of previously published CM biomarkers ... **39**
1.1 Single-gene CM biomarkers ... 39
1.2 The CP-GEP commercial CM signature ... 43
1.3 The 26 gene signature of Gerami et al 2015. ... 44
1.4 The 6 consensus-based classes created by Thakur et al 2019. ... 46
Summary Chapter 4 ... 47
Chapter 5: Explore sex-specific differences in primary CM, normal skin and Nevi samples. ... **49**
2.1 Sex stratified Kaplan-Meier plots ... 49
2.2 Quantifying the effect of sex in survival analysis of CM ... 50
2.3 Principal component analysis (PCA) on gene expression data ... 51
2.4 Gene expression variability ... 53
2.5 Class comparison analysis ... 57
2.6 Sex-stratified immune deconvolution analysis using xCell ... 58
2.7 Multivariable Cox Regression Analysis of Immune Cell Subtypes ... 61
Summary chapter 5 ... 63
Chapter 6: Discovery of sex-specific prognostic gene biomarkers in stage I-II CM ... **64**
3.1 Filtering out genes utilising multivariable Cox regression analysis ... 64
3.2 Investigation of the gene*sex interaction in a multivariable Cox regression analysis ... 68
3.3 Discovery of female-specific prognostic gene signature ... 69
3.4 Discovery of male-specific prognostic gene signature ... 71
3.5 Individual assessment of genes in the discovered sex-specific signatures ... 73
Summary chapter 6 ... 77
Chapter 7: Assessment of the identified sex-specific prognostic models on external cohorts. ... **78**
4.1 Validation of the female-specific gene signature ... 78
4.2 Validation of male-specific gene signature ... 79
Summary chapter 7 ... 80

Chapter 8: Functional analysis of the discovered sex-specific gene biomarkers 81

5.1 Assessment of correlation of the genes of interest with all other genes 81
5.2 Female- specific genes 83
5.3 Male-specific genes 85
5.4 Immune deconvolution using xCell 87
5.5 Network analysis 89
5.6 Protein Atlas analysis 91
Summary chapter 8 96

Chapter 9: Discussion 98

Chapter 10: Strength, limitations, and future directions 111

Chapter 11: Conclusion 112

References 112

Appendix 131

List of Tables

Table 1: Overview of TNM pathologic staging for melanoma, obtained from the Melanoma Research Alliance (1). N = number of tumour-involved regional lymph nodes; M = number of metastases at distant site; T = primary tumour thickness (T1 ≤ 1 mm, T2 > 1.0 – 2.0 mm, T3 > 2.0 – 4-0 mm, T4 > 4 mm).

Table 2: Detailed characteristics of all used gene expression datasets, where N/A no information provided by cohorts. F= females, M = Males

Table 3*:* This table presents sex-specific multivariable Cox regression results in females and males together (adjusted fro age,stage and sex) and separately in females and males (adjusted for age and stage) with stage I-II melanomas.

Table 4: The table presents the results of the C-index, likelihood ratio test, Wald test and score (logrank) test of multivariate Cox regression model of stage I,II (A) and stage III (B) performed on females and males together and on females-only and males-only.

Table 5: The tables present the results of the C-index, likelihood ratio test, Wald test and score (logrank) test of multivariate Cox regression model initially performed on females and males together and on females-only and males-only of stage I,II (A) and stage III (B).

Table 6: Cox regression analysis results for sex in the female-male subgroups; stage I,II,II; stage I; stage II; stage I-II and stage III.

Table 7: Variance explained by the first five principal components in LMC, TCGA, GSE53118 and GSE65904 melanoma cohorts, in GTEx – not sun exposed (NSE) and – sun exposed (SE) normal skin, and in Nevi samples (GSE46517 and GSE3189).

Table 8: Average coefficients of variation obtained from the gene expression data of melanoma (LMC, TCGA, GSE53118) and non-melanoma (Nevi, Normal skin) datasets were calculated for both females and males in the stage stratified groups.

Table 9: Gene Ontology Biological Pathways and Coefficient of Variation Analysis. The table lists 15 gene ontology biological pathways along with the total number of genes in each pathway.

Table 10: Up-regulated autosomal and sex chromosome genes in females and males. The table summarizes the number of up-regulated autosomal and sex chromosome genes identified through class comparison analysis across various datasets.

Table 11: Multivariable Cox Regression Analysis in Female Stage I-II Melanomas. The results shown are obtained from a multivariable Cox regression analysis for female patients with stage I-II melanomas from the LMC dataset.

Table 12: Multivariable Cox Regression Analysis on the LMC Melanoma Cohort. This table presents the results of a multivariable Cox regression analysis on the LMC melanoma cohort.

Table 13: Multivariable Cox Regression Analysis of Age, Stage, BRAF, and NRAS in LMC Stage I-II Melanomas.

Table 14: Average performance metrics of 10-fold cross validation in female stage I-II melanomas.

Table 15: Average performance metrics of 10-fold cross validation in male stage I-II melanomas.

Table 16: Individual gene evaluation in multivariable Cox regression (adjusted for age and stage) for females with stage I-II melanomas, LMC cohort.

Table 17: Individual gene evaluation in multivariable Cox regression (adjusted for age and stage) for males with stage I-II melanomas, LMC cohort.

Table 18: Validation metrics of the female-specific survival model applied to Biella and TCGA datasets.

Table 19: Validation metrics of the male-specific survival model applied to Biella and TCGA datasets.

List of Figures

Figure 1: Figure obtained from "*Sex differences in immune responses. Nat Rev Immunol **16**, 626–638 (2016)"* shows immunologic factors varying between females and males throughout their lifespan.

Figure 2: Estimated age-standardised incidence (blue) and mortality (red) of all cancers, across all continents showing an increased incidence and mortality in males compared to females across all continents. Dual bar graph obtained from Globocan 2020.

Figure 3: The 25 literature found melanoma biomarkers plotted according to their beta coefficients and raw p-value obtained from the multivariable Cox regression analysis together in females and males (grey) and separately in females (red) and males (blue) stage I-II melanomas. Horizontal red line at -log10(0.05).

Figure 4: The 25 literature found melanoma biomarkers plotted according to their beta coefficients and raw p-value obtained from the multivariable Cox regression analysis together in females and males (grey) and separately in females (red) and males (blue) stage III melanomas. Horizontal red line at -log10(0.05).

Figure 5: C-index Results from Multivariate Cox Regression Models for Key Nodal Genes of each LMC Class.

Figure 6: Kaplan - Meir overall survival curves with the log-rank test p-value between females and males for the LMC dataset on A) stages I-II and III; B) stage I; C) stage II; and C) stage III.

Figure 7: Principal component 1 (x axis) and 2 (y axis) from PCA analysis applied to the entire LMC melanoma cohort.

Figure 8: Density Plots of Average Bootstrap Coefficients of Variation from Gene Expression Data. Bootstrap analysis was performed on random subgroups of 3000 genes, repeated 10,000 times.

Figure 9: Differential Enrichment of Immune Cell Subtypes in LMC Female and Male Stage I-II (A) stage III (B).

Figure 10: Distribution of genes with significant p-values in adjusted models ($p < 0.05$) in females with stage I-II cancer.

Figure 11: Multivariable Cox regression analysis results for the genes in stage I-II melanomas.

Figure 12 : Interaction Term Analysis of Gene Impact on Survival in Stage I-II CM.

Figure 13: Regression Coefficients vs -log10(raw p-value) in Female and Male separately in stage I, II CM.

Figure 14: Multivariable Cox regression Hazard ratios and related confidence intervals of the three female-specific stage I-II survival covariates on the LMC cohort.

Figure 15: Multivariable Cox regression on LMC Hazard ratios of the four male-specific stage I-II survival covariates.

Figure 16: Kaplan-Meir curves of the five genes of interest in LMC females with stage I-II melanomas.

Figure 17: Kaplan-Meir curves of the five genes of interest in LMC males with stage I-II melanomas.

Figure 18: Spearman correlation coefficients plotted against the -log10 adjusted p-value of the top 200 positively correlated and the top 200 negatively correlated genes, each colour representing one gene of interest.

Figure 19: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for UHRF1.

Figure 20: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for HLA-E.

Figure 21: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for UBA7.

Figure 22: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for BEX3.

Figure 23: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for SF3B3.

Figure 24: Correlation between gene expression of *HLA-E*, *UHRF1* and *UBA7* and xCell enrichment scores in female stage I-II melanomas and between gene expression of *SF3B3* and *BEX3* and xCell enrichment scores in male stage I-II melanomas.

Figure 25: Network created using GeneMANIA for females-identified genes, UHRF1, HLA-E and UBA7.

Figure 26: Network created using GeneMANIA for males-identified genes, BEX3 and SF3B3.

Figure 27: The top plot represents HLA-E normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to HLA-E mRNA expression levels in female primary melanomas from TCGA.

Figure 28: Top plot represents UHRF1 normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to UHRF1 mRNA expression levels in female primary melanomas from TCGA.

Figure 29: Top plot represents UBA7 normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to UBA7 mRNA expression levels in female primary melanomas from TCGA.

Figure 30: Top plot represents BEX3 normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to BEX3 mRNA expression levels in male primary melanomas from TCGA.

Figure 31: Top plot represents *SF3B3* normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to *SF3B3* mRNA expression levels in male primary melanomas from TCGA.

List of papers

1. **Eirini Chrysanthou** , Emir Sehovic , Paola Ostano , Giovanna Chiorino. Comprehensive gene expression analysis to identify differences and similarities between sex and stage stratified melanoma samples. *Cells* 2022, *11*(7), 1099; https://doi.org/10.3390/cells11071099

This study examined gene expression in early-stage melanomas (stages I and II), advanced-stage melanomas (stage III), nevi, and normal skin. Genes were stratified based on autosomal and sex chromosome, emphasising the importance of sex and stage classification in understanding melanoma prognosis.

The analysis revealed significant differences in gene expression between sexes and stages. Females showed better overall survival rates than males in stages I and II, whereas this difference diminished in stage III. Interestingly, based on the class comparison analysis to identify differentially expressed genes, early-stage males displayed under-expression of several tumour suppressor genes located on the X chromosome, potentially influencing the aggressiveness of melanoma at the disease's outset. Moreover, there were more genes showing differential expression between females and males in normal skin and nevi compared to melanoma datasets, indicating pronounced sex-based differences in non-malignant tissues. However, as melanoma progressed through stages, these sex-related differences in gene expression reduced. Interestingly, despite similarities in highly variable genes between sexes and stages, differences in gene expression variability were observed, with higher variability in males, especially in stage III melanomas. This heterogeneity correlates with increased genetic instability and tumour progression. The study also found differences in immune-related cells and pathways in early and late-stage melanomas in both sexes, influencing tumour development and progression.

Overall, the study highlighted the nuanced variations in gene expression between sexes and different stages of melanoma. It emphasises the importance of taking sex and stage into account in future transcriptomic or biomarker melanoma studies.

2. Alexandre Crépin, Audrey Thiroux, Aurélien Alafaci, Amine M. Boukerb, Izelenn Dufour, **Eirini Chrysanthou**, Joanne Bertaux, Ali Tahrioui, Alexis Bazire, Sophie

Rodrigues, Laure Taupin, Marc Feuilloley, Alain Dufour, Jocelyne Caillon, Olivier Lesouhaitier, Sylvie Chevalier, Jean-Marc Berjeaud & Julien Verdon. Sensitivity of *Legionella pneumophila* to phthalates and their substitutes. *Sci Rep* 13, 22145 (2023).

For this analysis I performed the bioinformatic analysis, applied the FASTQC tool for the quality and the raw reads count. Genes were mapped based on the reference genome using Trimmomatic and counted using FeatureCounts. Data were normalised and differential expression was performed using DESeq2 package.

3. **Chrysanthou E**, Sehovic E, Mello-Grand M, Ostano P, Biletta E, Manzoni R, Bolzonaro A, Laurora S, Demo K, Pisacane A, Chiorino G. Discovery and validation of sex-specific survival biomarkers in early-stage melanoma.
   Manuscript in preparation.

This study includes the majority of the results included in this thesis, identifying sex-specific prognostic gene signatures in early-stage melanoma and validating these results in external independent melanoma cohorts.

Chapter 1: Background and Introduction

This thesis refers to sex differences and not gender differences. Sex is defined by the multidimensional biological characteristics that distinguish females and males, referring to anatomical and physiological differences, such as reproductive organs and chromosomes. Notably, intersex individuals may possess variations in sex characteristics, sex chromosomes, hormones, and reproductive anatomy beyond typical definitions of female or male. Gender encompasses the cultural and psychological elements linked to being female, male or non-binary. It relates to the roles, behaviours, identities and expectations that are influenced by norms and cultural environments (2–4).

Discrimination of female sex from biomedical studies

Research results on diagnosis, treatment and prevention mainly originate from multiple studies done on male cells, male mice, and men (5–8). In 2010 a study reported a male bias in 8 out of 10 biological disciplines, male to female ratios of 5.5:1 in neuroscience, 5:1 in pharmacology and 3.7:1 in physiology (9). They also showed that 75% of studies analysed from highly cited journals do not specify the sex of the animals or the cells they used.

This disproportion of the two sexes resulted from the fact that historically, for multiple reasons including safety purposes, women of childbearing age were excluded from clinical trials (10). Another reason for excluding women, is the belief that the estrous cycle and hormone changes cause an increased variability in females (11). Instead of completely excluding 50% of the population, multiple solutions have been proposed over the years for addressing women's oestrous cycle (12) such as the inclusion of oestrous/ hormonal staging (13) or increase and stratification of the sample, but a more robust understanding of sex differences in variability is clearly needed to guide future research and experimental designs .

Sex differences in physiology

Females and males differ in so many aspects under physiological conditions such as hormonal (14) variations, reproductive systems and chromosomal differences with a heterologous XY chromosome pair in males and a homologous XX chromosome pair in females. Females are known for their prolonged lifespan compared to males (15,16). Body compositions are different (17) with men having more muscle mass (18) and increased bone density compared to women

that have higher body fat (19) percentages. Additionally, variances exist in the cardiovascular system, affecting factors like heart rate and blood pressure (20). Females often display a stronger immune response, with distinction being made in both innate and adaptive immunity and differences fluctuating from in utero to old age (Figure 1). This also makes women more susceptible to autoimmune diseases (21–23). Additionally, the human X chromosome contains over 1,100 genes, making about 5% of the total whole genome, and includes a significant number of genes related to the immune system. Metabolic variations show that males lean towards carbohydrate-based cellular metabolism, while females tend towards lipid-based metabolism (24,25). Neurologically, disparities in brain structure and neurotransmitter activity lead to differences in cognition, emotional processing, and susceptibility to specific neurological conditions between the sexes (26).

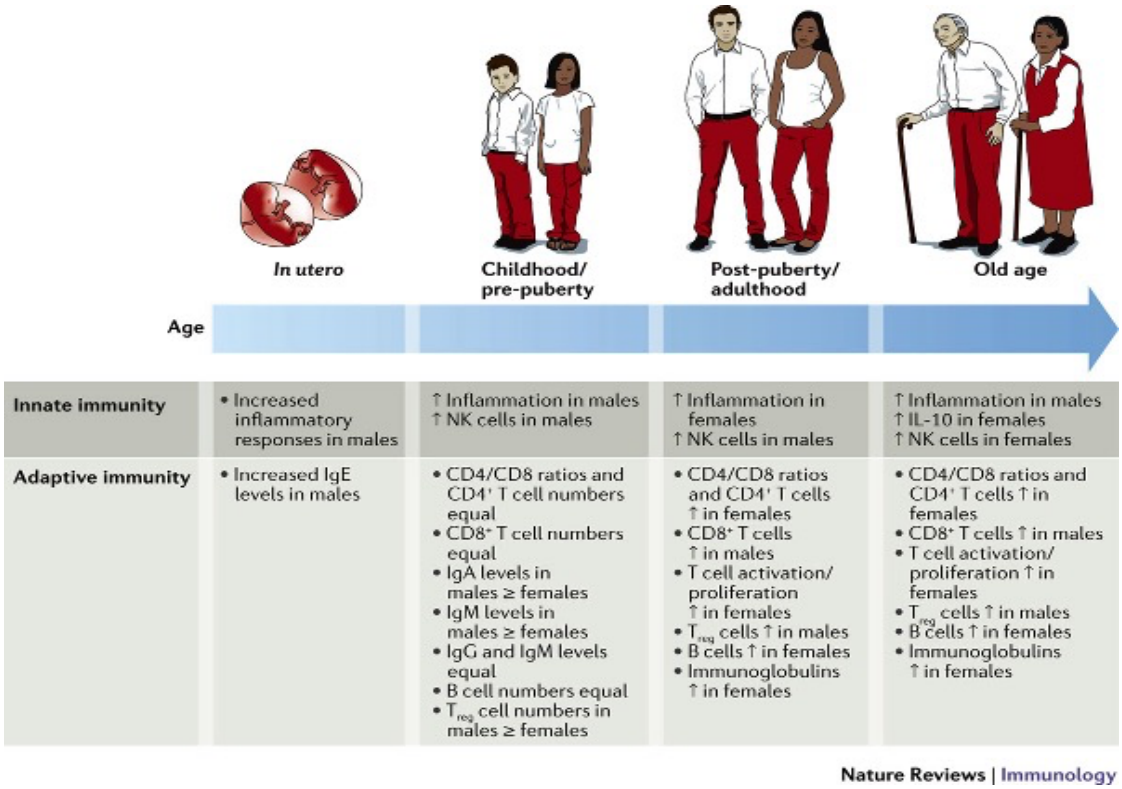


Figure 1: Figure obtained from "*Sex differences in immune responses. Nat Rev Immunol* ***16****, 626–638 (2016)"* shows immunologic factors varying between females and males throughout their lifespan.

Sex differences in pathology

As expected, the divergences observed between sexes manifest in varied disease incidence, mortality rates, and treatment responses. Particularly, in prevalent causes of death, such as heart disease (27), sex disparities exist in symptoms, diagnosis, and treatment. Notably, women

frequently encounter distinct forms of heart disease and tend to receive less effective care, resulting in higher mortality rates than in men. Another disease that exhibits important differences is chronic pulmonary disease (28), with women showing higher susceptibility and encountering challenges in diagnosis. Stroke affects more men globally, but its incidence in women is rising, contributing to higher prevalence and mortality among women (29). Alzheimer's disease affects more women, often progressing faster and posing diagnostic challenges. Women bear a heavier burden in Alzheimer's caregiving (30). In diabetes, women and men exhibit varied patterns in type 2 diabetes onset and cardiovascular risks, with women facing severe consequences from the condition (31). Chronic kidney disease affects women due to pregnancy risks, while men progress faster (32). Men face higher chronic liver disease risks from alcohol, with women being more vulnerable post-menopause (33). Depression is more common in women, but men often show different symptoms. Suicide rates are higher in men and attempts are more frequent in women (34). Finally only 39% of participants in clinical trials concerning diseases such as HIV/AIDS, chronic kidney disease, and cardiovascular disease are women (35). Regarding cardiovascular disease, as women's blood pressure (BP) is lower than men's the risk for these diseases starts at a lower BP threshold than the standard "high BP" level (36). Drug pharmacokinetics are also influenced by sex-differences found in both intrinsic factors such as body weight, genetic predisposition, renal or hepatic function and extrinsic factors such as smoking, concomitant medication including herbal products, alcohol use and diet. These factors can result in sex-differences in the pharmacokinetics or the exposure to a drug, therefore mg/kg dosage should be prioritised instead of the one-size fits all (37,38).

Sex differences in cancer

Across various cancers, differences between females and males can be observed in multiple aspects such as incidence, tumour biology, treatment response, and survival rates (39,40). Men have a higher cancer incidence and mortality (Figure 2) than women in most shared anatomic sites including, rectum, kidney, gastric cardia, biliary tract, skin, liver, oropharynx, bladder, larynx, gastric noncardiac, and oesophageal adenocarcinoma (39–41). The only two cancer types that are more common in females than males are thyroid and gallbladder cancer (42). A study investigating carcinogenic exposures in behavioural and environmental factors among females and males identified sex-related biological mechanisms as the primary determinant of differences in cancer risk between sexes (43).

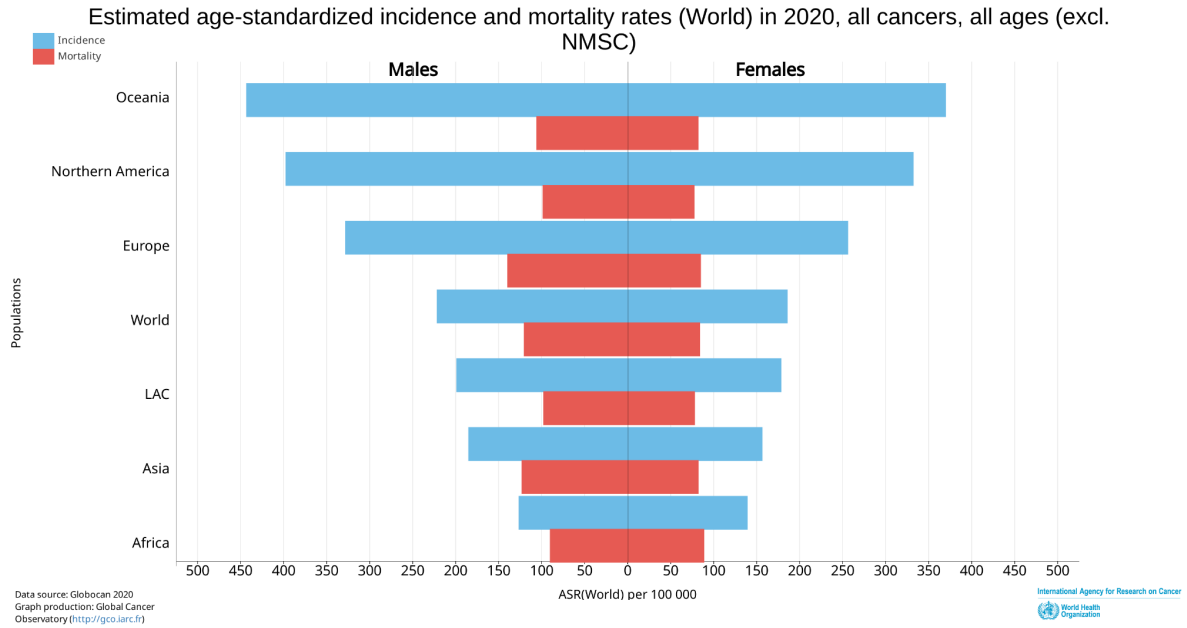


Figure 2: Estimated age-standardised incidence (blue) and mortality (red) of all cancers, across all continents showing an increased incidence and mortality in males compared to females across all continents. Dual bar graph obtained from Globocan 2020.

Differences between the two sexes in several cancers have been observed on epigenetics mechanisms. Dysregulation in genes governing epigenetic processes is frequently observed in human tumours, underscoring their link to tumorigenesis. DNA methylation, a prominent epigenetic mark, displays varying sex-specific patterns (44) in various tissues such as blood (45), liver (46), insulin secretion of the pancreas (47), myoblasts and myotubes (48), heart (49) and the brain (50,51). Another very important difference affecting many biological pathways is the presence of bi-allelic XX chromosomes in females, as opposed to the XY chromosomes in males. In females, early X chromosome inactivation aims to balance gene expression; some genes escape this process, exhibiting sex-biased expression and influencing cancer susceptibility. Among the genes that escape X-inactivation, we find ATRX, CNKSR2, DDX3X, KDM5C, KDM6A and MAGEC3, all known for their tumour suppressor activities (52,53). Due to the bi-allelic expression of these genes, they might provide extra protection against tumours in females only. Another cellular response to DNA damage known for its protective mechanism is cellular senescence, which also shows differences between sexes, with female cells more prone to undergoing senescence than male cells (54,55). Senescent cells are known for their senescence associated secretory phenotype (SASP), which is secreted as senescent cells remain metabolically active after their proliferation arrest. The numerous proinflammatory cytokines produced as part of the SASP increase immune cell infiltration, promoting a more tumour-permissive environment (56). Furthermore, sex hormones are known

to influence the immune system, with estrogens usually boosting the immune response (57,58), progesterone showing an anti-inflammatory (59,60) effect and androgen generally suppressing immunity (61,62). As a result, they alter both the innate and the adaptive immunity, differently across sexes. Moreover, there have been reports of sex-differences also in response to chemotherapy, with women experiencing more adverse effects and being more susceptible to chemotherapy than men. In colorectal cancer, females had increased risk of vomit, alopecia, heartburn, mucositis (63) and in small-cell lung cancer women experienced more hematologic toxicity than men (64). Even if there have been several improvements in recent years, still many researchers, even though they are aware of the number of females and males in their datasets, do not take sex in consideration during their analyses.

Cutaneous melanoma (CM)

The skin, our body's protective shield, is intricately defended against the sun's harmful rays by melanocytes producing melanin. This pigment shields skin cells from mutating under UV radiation. Yet its balance, the ratio of eumelanin to pheomelanin (65), and genetic variations, can dictate skin colour and influence susceptibility to skin cancer. CM, an infrequent yet deadly form of skin cancer, stems from these melanocytes (66). CM is by far the most common subtype of melanoma, other than mucosal and uveal melanoma (67). Despite its rarity, melanoma accounts for 80% of skin cancer-related deaths, even though it only accounts for 4% of skin cancer cases (68). In the US there are 100,640 new cases of CM per year with 59,170 occurring in males and 41,470 females (69). Additionally, there is an estimation of 8,290 deaths per year with 5,430 in males and 2,860 in females (70). If melanoma rates continue to rise according to the 2012 rate, by 2024 there will be a roughly 50% incidence rate increase and a 68% increase in deaths caused by CM (71).

Brief history of melanoma

Melanoma traces back centuries, with Hippocrates of Cos documenting cases as early as the 5th century (72). However, medical understanding remained limited until the 18th century when Scottish surgeon John Hunter performed the first surgical removal of a melanoma tumour, initially mistaking it for a fungal growth (73). This “mistake” led to the first recognition of this disease, which was then classified based on the location of the tumour. Dr. William Norris's observations in the 1820s hinted at hereditary links and advocated wide excision for treating melanoma lesions (74). Throughout the 19th century, surgical resection

and lymph node removal were the primary treatment methods, with little advancement until the advent of chemotherapies in the mid-20th century. Since the 1980s, after realisation of the importance in cancer development of accumulating genetic mutation, there has been an effort to identify individual mutations that commonly occur in melanoma and develop targeted therapies (75). These historical milestones laid the groundwork for understanding and managing this formidable disease, offering glimpses into its complexities well before modern medical advancements.

Melanoma epidemiology and risk factors

Unlike other cancers that have a declining trend, CM incidence and mortality rate continues to rise with a striking 41% in incidence from 2012 to 2020, especially in fair-skinned populations, with low levels of melanin, of European decent. The rise of CM varies among humans is caused by several risk factors such as UV exposure (76), ethnicity , geographic location, age group and sex (77). The highest incidence of CM cases is observed in Australia and New Zealand, followed by Western Europe, North America and North Europe (71). Both incidence and mortality are higher in men than in women. Even though CM has the lowest incidence rate among skin cancers, its ability to rapidly metastasise and affect younger individuals makes CM a significant health and economic burden (78,79). The success of skin cancer prevention efforts varies among different countries, depending on public education about UV protection and the implementation of skin cancer screening programs.

Melanoma diagnosis

Early and accurate diagnosis of melanoma is crucial for effective treatment and improving patient outcomes. Melanoma diagnosis involves a combination of clinical evaluation, dermoscopic examination, and histopathological analysis. Advances in diagnostic techniques have significantly enhanced the ability to detect melanoma at an early stage, thereby increasing the chances of successful treatment (80).

Clinical evaluation starts by a thorough visual examination of the skin, looking for any unusual moles or lesion. The ABCDE criteria involving assessment based on Asymmetry, Border irregularity, Colour variation, Diameter greater than 6 mm, and Evolving characteristics of moles is used (81). Dermoscopic examination is performed using a dermatoscope to magnify

and illuminate skin lesions, allowing for a more detailed examination of pigmented skin lesions, or by digital dermoscopy to monitor changes over time (82). Histopathological analysis consists of removal of a suspect lesion or a part of it for microscopic examination by a pathologist to confirm the diagnosis.
Latest technological advancements regard molecular diagnostics (gene expression profiling to identify specific gene mutations associated with melanoma; Next-Generation Sequencing to comprehensively analyse multiple genes to detect mutations that can guide targeted therapies) and imaging technologies, such as reflectance confocal microscopy and optical coherence tomography (83).

Challenges in distinguishing melanoma from benign skin lesions include atypical presentation of some melanomas, similarities to benign lesions such as dysplastic nevi or seborrheic keratosis, variability in diagnostic expertise, high cost of advanced imaging technologies, which are not universally available, and the fact that some melanomas can evolve rapidly, changing their appearance over short periods, which complicates long-term monitoring and diagnosis.

Melanoma classification

Wallace Clark suggested in the 1960s that melanoma should be classified on histological features instead of tissue origin. He then identified three variants of melanoma that exist until today: superficial spreading melanoma (SSM), lentigo malignant melanoma (LMM) and nodular melanoma (NM) (84). In addition to these, other variants have been identified, such as acral lentiginous melanoma, mucosal melanoma, desmoplastic melanoma and nevoid melanoma (85). Clark also proposed a level-system of evaluating melanoma based on depth of invasion into the dermis and subcutaneous fat. Thus, Clark's level has been proposed. It is used until today (86,87) and is one of the most important tissue biomarkers for predicting invasiveness of the disease, going from level 1 - level 5, 1 corresponding to melanoma cells confined to the epidermis and 5 to invasion of melanoma cells into the subcutaneous fat . In the 1970, Alexander Breslow independently devised a more accurate method for classifying melanoma based on a measured depth of invasion reporting the thickness of the tumour. Breslow thickness is used until today and stratifies melanoma into five levels (82). Breslow showed that first two levels correlated with better survival rates, with reduced metastatic risks. This approach led to smaller resections and aided in assessing lymph node involvement,

guiding decisions on lymph node removal, especially important for patients with a Breslow thickness exceeding 1.5 mm. Over time, Clark's level and Breslow thickness, were refined and remain crucial in predicting patient outcomes (90).

Over the next 40 years, advancements in statistical methods and accumulating data on melanoma treatment led to the creation of refined staging systems, notably the TNM staging system by the American Joint Committee on Cancer (AJCC) (91,92). This system utilises histologic attributes of the primary tumour (T), regional lymph node involvement (N), and distant metastasis (M) to categorise patients, providing crucial guidance for clinicians based on linked survival and prognosis factors. Beyond Breslow depth, attributes like ulceration, mitotic rate, inflammation, and regression are integrated into the AJCC staging for melanoma (93) (Table 1).

Table 1: Overview of TNM pathologic staging for melanoma, obtained from the Melanoma Research Alliance (1). N = number of tumour-involved regional lymph nodes; M = number of metastases at distant site; T = primary tumour thickness (T1 ≤ 1 mm, T2 > 1.0 – 2.0 mm, T3 > 2.0 – 4-0 mm, T4 > 4 mm).

Stage	Tumour	Node	Metastasis
0	Tis	N0	M0
IA	T1a or T1b	N0	M0
IB	T2a	N0	M0
IIA	T2b or T3a	N0	M0
IIB	T3b or T4a	N0	M0
IIC	T4b	N0	M0
IIIA	T1a/b or T2a	N1a or N2a	M0
IIIB	T0	N1b or N1c	M0
	T1a/b or T2a	N1b/c or N2b	M0
	T2b or T3a	N1a/b/c or N2a/b	M0
IIIC	T0	N2b/c or N3b/c	M0
	T1a/b, T2a/b or T3a	N2c or N3a/b/c	M0
	T3b or T4a	Any N > N1	M0
	T4b	N1a/b/c or N2a/b/c	M0
IIID	T4b	N3a/b/c	M0
IV	Any T, Tis	Any N	M1

The latest update in the AJCC staging manual, the 8th edition, brought critical changes that refined measurements, expanded existing categories, and enhanced prognostic stratification in

melanoma staging. This staging system is pivotal for clinicians, aiding in prognosis assessment and treatment planning for patient recovery or prolonged survival (94). However, despite an international staging system, diagnosing melanoma consistently remains challenging due to significant observer variability among pathologists (95,96). The subjectivity in visual observations for diagnosis contributes to this variability, prompting ongoing research for more objective methods in CM diagnosis. Efforts to enhance detection and diagnosis include utilising non-invasive imaging before biopsies and employing advanced quantitative techniques post-biopsy, such as fluorescent in situ hybridization (FISH), comparative genomic hybridization (CGH), DNA/RNA sequencing, mass spectrometry (MS) for proteins and/or metabolites evaluation, and immunohistochemistry (IHC) (97,98).

Melanoma diagnostic markers

Cost effective IHC serves a critical role in diagnosing melanoma, particularly for complex cases challenging to discern through traditional methods. In melanoma diagnosis, two key marker types, melanocytic and proliferative, are commonly employed (99). Melanocytic markers help determine melanocytic origin, while proliferative markers assess cell cycle activity. Various genes are involved in this process, including Melan-A/MART-1 (100), HMB-45 (101), S100 protein (102), microphthalmia transcription factor (MITF) (103), Tyrosinase, SOX10 (104), gp100 (Pmel 17) (105), and BRAF and NRAS (106,107). These markers play vital roles in identifying melanoma, though each has its limitations (108). While sensitive, some markers lack specificity in distinguishing between melanoma and benign lesions. Additionally, specific markers may not express in certain melanoma types, potentially leading to false-negative diagnoses without more sensitive markers (97).

Melanoma prognosis

Prognosis in melanoma is a critical aspect of patient management, providing essential information that influences treatment decisions and survival outcomes. Breslow thickness, ulceration, mitotic rate and the involvement of lymph nodes are amongst the most important adverse prognostic factors. In addition to these established features, improvements in transcriptomic analysis have facilitated the discovery of numerous gene expression signatures to personalise CM prognosis.

The signatures are derived from gene expression obtained from high-throughput RNA sequencing or microarray profiling and have been shown to predict survival, especially the risk of developing metastases. The Cam_121 gene signature, for example, was found to significantly enhance the prognostic prediction for metastasis and survival in primary melanoma (109). Another study evaluated the accuracy of gene expression profiles (GEP) in predicting the metastatic risk in CM. They classified patients in low and high-risk applying both univariate and multivariate analysis of 28 genes, that were previously found to be significantly associated with recurrence-free survival and distant metastasis-free survival (110,111). Even though these studies have adjusted for sex, additional to other variables like age and stage, they have not further investigated findings in a sex-stratified manner. Moreover, high quality evidence for gene expression predictors in early stage CM are lacking. Another study tried addressing the difficulties faced in stage I CM, where tumours are very small in size for sampling. They identified a six-class gene expression signature that predicts outcome in patients treated with immunotherapies. The insufficient number of stage I cases and the lack of comprehensive data on treated patients with immunotherapy and targeted thereby limit these findings (112).

Melanoma prognostic and serological markers

Proliferation markers like Ki-67 and phosphohistone H3 (PHH3) play a crucial role in determining cell cycle activity in melanoma lesions (113–115). Ki-67, associated with cell proliferation, is commonly used in diagnosing aggressive melanomas. PHH3, introduced later, detects cells undergoing mitosis more specifically than Ki-67, which captures cells in the cell cycle (116,117). While both markers offer limited independent prognostic insights in melanoma, they lack specificity for melanocytes (105). In tumours with highly active immune cells, these markers may overestimate proliferation. Although informative, their use is not mandated in melanoma reporting or staging (118). Serum biomarkers like lactate dehydrogenase (LDH) serve as independent prognostic indicators for melanoma. Elevated LDH levels are associated with poorer survival, especially in stage IV melanoma (119). The diversity of melanoma prevents the identification of a single perfect diagnostic or prognostic biomarker. Using panels of multiple biomarkers compensates for individual limitations. Ongoing research aims to discover new biomarkers to improve melanoma diagnosis and prognosis. However, up to our knowledge, neither “in use” nor “under evaluation” prognostic markers have been separately analysed in females and males.

Molecular characteristics

- Inherited (familial) melanoma

Hereditary melanomas, making up to 5-12% of cases, show distinct mutations, with *CDKN2A* mutation being the most commonly found in familial melanomas, although it also occurs in sporadic cases. CDKN2A mutations lead to defects in tumour suppressor proteins and regulators of p53 pathway, specifically p14ARF and p16INK4A (120). Other notable mutations in hereditary melanomas include those in the BAP1 gene(121), associated with melanocytic tumours, and mutations in POT1, which impact telomere maintenance and genome stability(122).

- UV-induced melanoma

UV- induced melanomas are primarily driven by mutations resulting from UV radiation exposure. These melanomas usually harbour mutations in the MAPK pathway genes, which are responsible for initiation early oncogenic events. About 70% of melanoma feature mutations in this pathway (123).

One of the key mutations found in about 50% of melanomas is in the BRAF gene, with the activating mutation V600E being particularly common. These BRAF mutations are mainly observed in melanomas linked to intermittent sun exposure and the superficial spreading phenotype(124,125). Although 80% of benign moles carry BRAF mutations, a single BRAF mutation alone is insufficient for malignancy, requiring additional mutations to transform normal cells into tumours (125,126). Another significant mutation, found in 15-20% of melanomas, is in the NRAS gene, which is associated with chronic sun exposure (127). Additionally, mutations in the C-KIT gene are present in 2% of melanomas and are common in mucosal melanomas (128). About 50% of uveal melanomas exhibit mutations in the GNAQ gene (129).

In addition to these key mutations, other significant mutations include those in the PTEN and KIT genes, which are involved in metabolism dysregulation, and mutations in the TP53 gene, which induce resistance to apoptosis (130,131). Mutations in the promoter of the telomerase reverse transcriptase (TERT) gene can cause up to a twofold increase in transcription (132).

In UV-induced carcinogenesis, MAPK kinases play a major role. When activated, these proteins translocate to the nucleus, where they phosphorylate key transcriptional factors responsible for proliferation, development, and cell death (133,134).

Melanoma treatment

The primary treatment to localised melanoma is **surgical resection**, where the surgeons remove the tumour area and surrounding healthy tissue. They only perform sentinel lymph node biopsy if the tumour is thicker than 0.8mm and to thinner tumours if they have ulceration (135). **Targeted therapy** has revolutionised the approach to treating advanced melanoma, particularly with drugs like BRAF inhibitors such as vemurafenib and dabrafenib. These medications, approved in 2011 and 2013 respectively, directly address specific molecular defects found in melanoma. Although initially promising for BRAF-mutated melanomas, these inhibitors often encounter secondary resistance, prompting ongoing research into novel drug combinations and mechanisms to achieve sustained effectiveness (136). The development of targeted therapies determined an important shift in melanoma treatment, showcasing a more precise and personalised approach that aims to counteract specific genetic mutations driving the cancer's growth (137). **Immunotherapy** in treating advanced melanoma uses the body's immune response against cancer cells, especially in melanoma, known for its high mutational burden (138) that produces unique antigens (139). These tumour-specific antigens, arising from mutated genes beyond BRAF and NRAS, trigger an immune reaction. Immune cells like CTLs recognize these antigens presented on MHC-I, while antigen presenting cells (APCs) interact with CD4+ Th cells via MHC-II, amplifying the immune response (140). Melanoma often evades this response by downregulating antigen presentation or secreting inhibitory molecules. Immunotherapies like checkpoint inhibitors disrupt these evasion tactics by blocking checkpoint proteins from binding with their partner proteins, reactivating the immune system against melanoma cells (141). Ongoing research aims to enhance effectiveness and predict patient responses by exploring biomarkers and alternative treatment targets in this constantly evolving field. Finally, in advanced or metastatic melanoma when other treatments fail, **chemotherapy** might be used alone or combined with other treatments as a last option for treating this disease (142).

Melanoma sex differences

Melanoma is a very aggressive and heterogenous skin cancer that tends to spread to other parts of the body (143). It is the fifth most common cancer among men and the sixth among women. Female overall survival rate advantage has been continuously confirmed for melanoma throughout the years. Males were shown to harbour unfavourable primary tumour characteristics, such as age, histological subtype, ulceration, Breslow thickness, mitotic rate, vascular invasion, and recurrence. These factors were shown to independently predict poor outcome in melanoma patients and lower the survival rate, reinforcing the importance of sex and sexual dimorphism in melanoma (144).

Sex hormones play a significant role in the development and progression of. Melanoma with various mechanisms influencing the disease differently in men and women. More specifically, oestrogens, characterised by their mediation through receptors such as ER-β, predominate in early stages but diminish as tumours progress, potentially affecting metastasis. Oestrogen (E2) and its metabolite (2ME2) inhibit melanoma cell growth and influence tumour microenvironment (145,146). Conversely, testosterone, operating via the androgen receptor, sparks cell proliferation in melanoma, indicating potential clinical implications (147). Notably, anti-androgens display a promising role in tumour progression, highlighting a potential avenue for therapeutic intervention (148,149). Furthermore, sex-based differences in immunotherapy responses stem from variations in genetic factors, including genes linked to immune pathways like Toll-like receptors (TLRs) and cytokines, to genomic instability, impacting the effectiveness of treatments like immune checkpoint inhibitors (ICIs) (150). Melanoma patients treated with ICIs demonstrated a more substantial therapeutic advantage in males compared to females. Genetic variations in key immune-related genes, such as PD-1 and PD-L1 (151), differ between sexes, affecting therapy outcomes. Additionally, distinct immune cell responses exist; females often exhibit more robust T cell reactions, while males may have more prolonged but subdued immune responses (152).

Chapter 2: Aims and Objectives

General Aim of the Thesis

This thesis aims to comprehensively investigate the impact of sex-specific gene expression on survival outcomes in early-stage melanoma. Through detailed analysis of diverse datasets, it seeks to identify and validate sex-specific gene signatures associated with prognosis, shedding light on the distinct gene expression variability, biological pathways and immune responses in female and male melanomas. These sex-specific signatures aim to classify early-stage CM patients as high or low risk helping adjustment of their follow-up care. Ultimately, in challenging current views on sex-based survival differences, this research advocates for a paradigm shift towards more refined personalised and targeted sex-specific approaches in CM management, research and patient care.

Specific aims of the thesis

1. Sex-specific evaluation of previously published prognostic genes/signatures for primary cutaneous melanoma (CM).
2. Exploration of sex-specific differences in primary CM, normal skin and Nevi samples.
3. Discovery of sex-specific prognostic signatures in stage I-II CM.
4. Assessment of the identified sex-specific prognostic models on external cohorts.
5. Functional analysis of the discovered sex-specific biomarkers

Chapter 3: Materials and Methods

Datasets

Datasets description and characteristics

A total of eight different transcriptomic profiling datasets were analysed (see detail in Table 2): Leeds Melanoma Cohort (LMC), TCGA, GSE65904, GSE53118 and the Biella qPCR cohort, all containing primary melanoma samples, except GSE53118 which had frozen lymph node biopsies; GTEx with both sun-exposed (SE) and not sun-exposed (NSE) skin and the two merged Nevi GEO datasets GSE46517 and GSE3189. Three of the datasets (TCGA, GTEx-NSE, GTEx-SE) are based on NGS platforms, while the other five were obtained from microarray platforms, either Illumina or Affymetrix. The probes or the ENSEMBL gene IDs, from now on are referred to as genes.

Table 2: Detailed characteristics of all used gene expression datasets, where N/A no information provided by cohorts. F= females, M = Males

	Sex	Stage	Number of samples	Mean age	Events
LMC	F	I	130	51.4	15
		II	181	53.2	59
		III	60	56.2	30
	M	I	95	56.1	18
		II	161	59.3	60
		III	47	58.9	22
GSE65904	F	Primary.Mel	7	78.9	
	M	Primary.Mel	8	69.5	
TCGA	F	I	1	61	0
		II	26	66.3	8
	M	I	1	81	0
		II	40	66.4	5
GSE53118	F	III	26	57.8	14
	M	III	45	57.2	25

Biella cohort	F	I	8	58.7	0
		II	14	64.9	5
	M	I	11	59.4	1
		II	10	68	3
GTEx - NSE	F	N/A	83	N/A	N/A
	M		170		
GTEx - SE	F	N/A	117	N/A	N/A
	M		220		
Nevi (GSE46517+GSE3189)	F	N/A	16	N/A	N/A
	M		11		

Leeds Melanoma Cohort (LMC)

To access the LMC primary melanoma transcriptome dataset (dataset ID EGAS00001002922) an agreement was signed between the University of Leeds and Fondazione Edo ed Elvo Tempia in Biella, Italy. The Leeds Melanoma Cohort (LMC) Study enrolled participants from a region of the UK between 2000 and 2012. The study obtained approval from the Research Ethics Committee (REC) with the reference number 01/3/057. It was also registered under the NIHR/CPMS ID. 15064 (Central Portfolio Management System). Participants were recruited five months after their diagnosis and their participation was based on informed consent. The study adhered to the guidelines outlined in both the World Medical Association Declaration of Helsinki and the Department of Health and Human Services Belmont Report (153). Initially the dataset had 703 samples, however 8 samples had no stage information and 1 sample had no overall survival information. Out of the remaining 694 samples, 26 samples were not used because they had a non-melanoma related death. Of the 668 remaining patients, 370 were females (311 stages I, II and 59 stage III) and 298 males (255 stages I, II and 43 stage III). The dataset includes quantile normalised gene expression data of 29,355 genes, the overall survival time in years for all the patients, information about their vital status, sex and age, as well as tumour stage. All events accounted for in our survival analysis were melanoma related deaths. Gene expression profiling was carried out on totRNA extracted from bulk fixed melanoma tissues, using the Illumina (Illumina HT 12.4, GPL28098-111924) c**D**NA-mediated **A**nnealing, **S**election, extension and **L**igation (DASL) protocol, designed to generate reproducible RNA profiles from degraded tissue samples such as formalin fixed, paraffin-embedded (FFPE) ones.

The Cancer Genome Atlas (TCGA)

Gene expression and stage of TCGA skin cutaneous melanoma (TCGA-SKCM) samples, together with overall survival time, vital status, age and sex of the corresponding patients, were downloaded from the cBioPortal (www.cbioportal.org accessed on 25 September 2021). The dataset used only frozen melanoma tissue samples, and the gene expression profiles of 20,501 genes were generated using the Illumina HiSeq 2000 RNA-seq platform. Genes with more than 70% of zero read counts across samples were removed, ending with a total of 17,550 used in all analyses. The TCGA cohort included tumours with sufficient mass and quality for their downstream molecular analyses, therefore advanced primary and/or metastatic tumours were overrepresented (154). Any stage IV and metastatic melanomas were excluded from the analysis giving a total of 94 samples (69 stage I, II: 28 females and 41 males; 27 stage III: 13 females and 14 males).

GSE65904

The GSE65904 (155), PMID: 31942071,dataset was downloaded from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/ accessed on 28 February 2023). Total RNA was extracted from fresh-frozen melanomas and genome-wide expression profiling was performed using Illumina Human HT-12 V4.0 expression beadchip. The study cohort included 31,810 quantile normalised genes from a total of 15 patients with primary melanoma, of which 7 are females and 8 are males.

GSE53118

The GSE53118 (156), PMID: 25521200, dataset, which was downloaded from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/ accessed on 15 September 2021), refers to 79 patients with stage III cutaneous melanoma, of which 29 are females and 50 males. Gene expression profiles were obtained from lymph node specimens in which macroscopic tumour was observed, obtained from patients without distant metastasis. The quantile normalised cohort had a total of 25,004 genes. This dataset was used for exploratory analyses not directly related to the other datasets as it did not consist of early-stage CM samples.

Genotype Tissue Expression (GTEx)

The normal skin gene expression profiles (gene read counts from RNA-Seq) were obtained from the GTEx portal (https://gtexportal.org/home/ accessed on 25 September 2021). A total of 590 samples were analysed: 253 (83 from females and 170 from males) for non-sun-exposed skin with a total of 25,238 genes and 337 (117 from females and 220 from males) for sun-exposed skin with a total of 25,139 genes. Gene expression profiles were obtained by RNA-sequencing and log2 normalised gene expression values were used for all analyses.

GSE46517 and GSE3189

Two nevi datasets, GSE46517, PMID: 20520718, (4 female and 5 male samples), and GSE3189, PMID: 16243793, (12 female and 6 male samples), were also downloaded from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/ accessed on 17 September 2021). Both datasets contain gene expression profiles obtained from frozen nevi and were analysed using microarrays and the same Affymetrix Human Genome U133A array. The two datasets were merged and batch effects removed, using the "removeBatchEffect" function in the "limma" (157) R package, to obtain the final nevi dataset. This yielded a total of 16 female and 11 male samples profiled for 21,430 genes.

Biella primary melanoma cohort

This dataset refers to a retrospective cohort of patients diagnosed with stage I or II CM between 2008 and 2015 at the hospital of Biella. All alive patients signed informed consent and the study was approved by the Novara ethical committee on 28/07/2021 (Protocol n. 802/CE). This cohort has a total of 43 patients of which 22 are females and 21 males. We collected information on age and vital status for each patient. RNA was extracted from FFPE tissue sections after macro-dissection of tumour areas highlighted by a pathologist. Reverse-Transcription quantitative Polymerase Chain Reaction (RT-qPCR) analysis was carried out in the lab to evaluate the expression of five genes of interest and a housekeeping gene (ACTB).

Reverse-Transcription quantitative Polymerase Chain reaction (RT-qPCR)

RT-qPCR was performed in order to robustly quantify the gene expression of the selected genes from the LMC and validate their prognostic power on an external cohort of patients (Biella

cohort). Both the paperwork procedure and the experimental part of the RT-qPCR was performed in collaboration with my colleagues from the Cancer Genomics lab in Biella. Seventy-one patients were selected by a dermatologist and tissue blocks were collected from the archive of the histopathology department at the Biella hospital. Consent was acquired for 64 out of 71 patients. FFPE blocks were cut and the slices were put on glass slides. Sections underwent haematoxylin and eosin (H&E) labelling and a histopathologist assessed the cancerous area which resulted in adequate only for 51 samples. FFPE blocks were then sliced using a microtome (~ 8 micron thickness), after macrodissection of the tumour area with a scalpel. Micro-dissected sections were then put into tubes and stored at -80°C until RNA isolation. Total RNA was extracted using the miRNeasy FFPE kit (Qiagen) following the indicated volume of 1-2 sections per tube. We evaluated the quantity and the purity of the samples using the Nanodrop and 8 samples were eliminated, because totRNA quantity was lower than 500 ng, leaving only 43 usable samples. For the RT we used the SuperScript™ VILO™ cDNA Synthesis Kit (Invitrogen, Thermo Fisher). For the qPCR the kits used were the TaqMan Fast Advance master Mix (Thermo Fisher) and TaqMan Gene Expression Assay (Thermo Fisher) with a total reaction volume of 20µl. Analyses were performed using the CFX96 BioRad instrument. The threshold cycle (Ct) of each gene was evaluated in triplicates. We calculated the standard deviation (std.dev) of the triplicates and if std.dev > 0.05 the Ct with the more distant reading was excluded. We then calculated the mean of the readings and subtracted it from the mean of our housekeeping gene, ACTB, obtaining the final estimate for the log2 relative expression of each tested gene. These values were used in the validation of the selected sex-specific gene biomarkers.

Statistical analysis

All statistical and computational analyses were performed within the R statistical environment (158). Unless specified otherwise, the significance cut-off for the Benjamini-Hochberg (BH) adjusted *p*-value for all relevant analyses was 0.05.

Survival analysis

Kaplan–Meier survival curves were generated and compared between sexes and stages, using the “survfit” function in the “survival” package and plotted using the “ggplot” package.

To filter the genes not associated with survival, for every gene, a multivariable Cox regression analysis was performed on gene expression data using the "coxph" function from the "survival" package on 8-year censored data. The survival time was calculated from time of diagnosis to the time of last follow up or time of death from CM. Patients with non-CM related deaths were removed from the analysis. This analysis was conducted on females and males together adjusting for age, stage, sex and on each sex separately correcting for age and stage every time. To create a reduced female stage I-II sample size, equal to male stage I-II, we used the "CreateDataPartition" function in R.

For stage I-II female CM samples the significant genes (with BH adjusted p-value < 0.05 from Cox regression) were further analysed using penalised Cox regression. Specifically, the "glmnet" R package was used, with alpha = 1 that corresponds to LASSO penalty. By using the "cv.glmnet" function, cross validation was done to find the optimal parameter lambda, which was set to "lambda.1se" that is within 1 standard error of cross-validated errors from "lambda.min", the one giving the smallest cross-validated error. To assess the performance of the predictors, cv.glmnet was applied 10 times on 9 of 10 random sample partitions each with the same proportion of events (10-fold cross-validation) and the 10 fitted models were applied to the remaining partitions. The genes with non-zero coefficients in all 10 models were selected and used to build the female-specific "coxph" model, which was further validated on two external cohorts (RT-qPCR and TCGA).

For stage I-II male melanoma samples, as there were zero genes with an adjusted p-value lower than 0.05 from the age and sex adjusted Cox regression, we selected the genes with top 5 hazard ratios within those with a raw p-value < 0.001. These selected male-specific genes were further analysed using bidirectional stepwise regression with the "step" function from the "stats" R package. As for females, to assess the performance of the predictors, a 10-fold cross-validation approach was applied. The predictors selected in all 10 models were then used to build the male-specific "coxph" model, which was further validated on two external cohorts (RT-qPCR and TCGA).

In both female and male specific analyses, we built a multivariable regression model using the cross-validation results to avoid using the average of the beta coefficients obtained from the cross-validation and to include all samples in each dataset and not a subgroup only.

Validation of predictive models

In order to apply the two sex-specific prognostic models to external datasets, we extracted the Cox regression beta coefficients, linear predictors from the "coxph" models and calculated the baseline survival S0(t). We measured the predictive ability of our models using several metrics such as the C-index which is defined as the ratio of correctly ordered (concordant) pairs to comparable pairs. Pairs with individuals that experienced an event are expected to rank higher than pairs that did not experience an event, the closer the value is to 1 the better the prediction. We used the UNO C-index (159) as it assesses the survival model focusing on a fixed time. The UNO AUC, or Time-Dependent Area Under the Receiver Operating Characteristic Curve (ROC AUC), quantifies a survival model's ability to predict event risk over time. It assesses the model's effectiveness in distinguishing between individuals experiencing an event at a fixed time point, in our case 8-years. The closer the value is to 1 the better predictive performance. The UNO AUC (160) was calculated using the "AUC.uno" function in R which implements the estimator of cumulative/dynamic AUC that is based on the true positive rate (TPR) and false positive rate (FPR). The ratio of the observed/expected predicted risk (OE ratio) was used as a calibration metric comparing observed number of events to the expected number of events predicted by the model: the closer the value is to 1 the better the calibration of the model.

Gene expression variability (GEV)

GEV was calculated in all datasets using the coefficient of variation (CV) statistic which is defined by standard deviation divided by the mean of expression across samples, for each gene. Comparison of CVs between the sexes (stratified by stages) was performed using the Wilcoxon two-sample paired test using the "wilcoxon.test" R function. In addition, we performed random sampling using bootstrapping, where we created splits of 3,000 genes and set the iterations to be at 10,000 and plotted the distribution of average CVs using "geom_density" from the "ggplot" R package.

Correlation analysis

The correlation between genes of interest and the remaining genes was assessed using Spearman correlation using the "cor.test" function in R. We performed five different correlation analyses, three in females analysing the genes positively and negatively correlated with female-specific genes of interest and two in males. The top 200 positively and top 200

negatively correlated with each gene were further run into functional enrichment and network analyses (described below).

Additionally, we evaluated the correlation between gene expression of genes of interest and immune xCell cell proportion estimates (described below).

Principal component analysis (PCA)

PCA was performed in both LMC and TCGA early-stage melanoma subgroups using the "prcomp" function from the "stats" package in R.

Immune cell infiltration analysis

xCell (161) is a gene signature-driven technique that generates cell type enrichment scores based on gene expression data for immune cell types. It was validated using extensive in-silico simulations and cytometry immunophenotyping. The "immunedeconv" package was used including the xCell function in R. This analysis was repeated in LMC stage I-II female, stage III female, stage I-II male and stage III male melanomas, as well as in GTEx, for both sun-exposed and not-sun exposed skin of both sexes. We conducted a comparison of the immune infiltration cell estimates of our sex-stratified datasets comparing the results using the "t.test" function. The significant results (adjusted p-value < 0.05) were visualised using the "ggplot2" package.
In addition, the xCell immune cell estimates were correlated with the LMC quantile normalised gene expression data of our selected sex-specific biomarkers.

Functional enrichment analysis

To analyse the biological processes (BP) overrepresented within the top 200 genes positively and negatively correlated with our sex-specific survival biomarkers, we used DAVID tools (https://david.ncifcrf.gov/). To retrieve enriched BPs in a statistically significant manner, the cut-off for the Bonferroni adjusted p-value was set to 0.05.

Protein-protein interaction Network

We used Genemania (162) for the group of three female-specific genes of interest and for the group of two male-specific genes. Genemania finds genes that are related to the set of input

genes using a large set of functional association data. We extracted the two network images for the group of three genes in females and for the group of two genes in males. In addition, we extracted for each group the functional pathways and reported the ones with an FDR 5%.

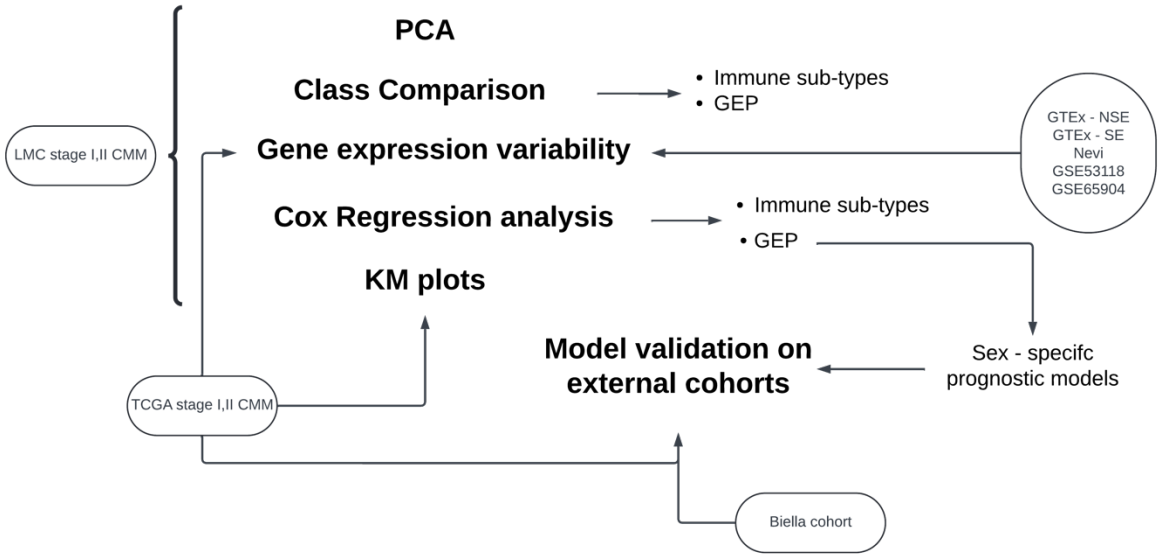
Human Protein Atlas

The Human Protein Atlas (163,164) maps human proteins in cells, tissues and organs using integration of various technologies. We searched individually our genes of interest and extracted information about their protein expression in different organs and their RNA tissue specificity, both found in the "tissue" tab. Using the "single cell" tab the summary of normalised single cell RNA (nTPM) from all single cell types is given. In addition, under the "pathology" tab dataset we selected our cancer of interest: melanoma. We extracted the Kaplan-Meier plot corresponding to the sex in which we detected our gene of interest. All extracted figures from the Human Protein Atlas, except from the survival plots, are not sex specific. The databases utilised by the Human Protein Atlas, for transcriptomic data under physiological conditions, is mainly GTEx and for pathological conditions TCGA.

Melanoma biomarkers found in literature

Using the following terms in pubmed: "melanoma survival biomarkers" and "melanoma tissue biomarkers" we found literature of existing tissue biomarkers, both protective and non-protective in terms of melanoma survival, in any stage of melanoma (165–167). The literature identified melanoma biomarkers were further assessed in the LMC dataset using a Cox regression analysis both on stage I-II and stage III. On stage I-II female and male sets we adjusted for age and stage while, in stage III we only adjusted for age. Their beta coefficients, hazard ratio and raw p-value were extracted and plotted using "ggplot" R package. Additionally, we retrieved articles that identified multidimensional CM signatures and assessed them collectively in our LMC cohort using multivariable Cox regression analysis. The evaluation metrics reported are their C-index, likelihood ratio test, Wald test and score (logrank) test obtained from the "coxph" function in R.

Study flowchart



Flowchart summary

The flowchart gives a summary of the study and the analysis performed for each dataset. On the LMC cohort we performed several analyses, including Principal Component Analysis (PCA), xCell, class comparison, gene expression variability, Cox regression analysis, and Kaplan-Meier (KM) plots. Class comparison and Cox regression variability were both used in immune sub-type and gene expression profile (GEP) analysis. Additional datasets such as GTEx, Nevi, GSE53118, and GSE65904 were utilized specifically for gene expression variability analysis. The model's validity was tested using external cohorts, including the TCGA and the Biella dataset.

Results

Chapter 4: Sex-specific evaluation of previously published CM biomarkers

Considering the persistent issue of incomplete sex stratification in biomedical research and the sex-difference in incidence and mortality in CM, this initial part seeks to address the sex-specific implications of already established CM survival biomarkers. The selected gene biomarkers are either single-gene or signature of genes and are re-assessed in our larger cohort, LMC, consisting of gene expression values obtained from CM biopsies, with 371 females and 302 males. The selected biomarkers, obtained from not sex-stratified analyses, are evaluated in stage I,II and stage III in three groups, 1; females and males together, 2; females-only, 3; males-only. In this chapter, we chose 25 individual prognostic gene biomarkers, 2 gene-signatures of 8-genes and 27 genes, and lastly, we also analysed a group of 6-classes of genes discovered upon unsupervised-clustering in early-stage CM.

1.1 Single-gene CM biomarkers

We found 25 individual gene biomarkers that play crucial roles early and late stage CM survival and participate in functions related to cell cycle, mitosis, inhibition and stimulation of apoptosis, DNA replication, stress response or to intracellular signalling. In addition to the individual biomarkers, we evaluated multiple-gene signatures. We assessed a commercial 8-gene signature called clinicopathological and gene expression (CP-GEP) model(168), developed by Mayo clinic and SkylineDx BV (169) and able to identify early stage CM patients at high or low risk of sentinel lymph node metastasis. A second signature made of 28 genes, created by Gerami et al (111) and predicting the metastatic probability of early-stage highly heterogeneous melanoma, was also tested. Finally, we evaluated the key nodal genes of six-class transcriptomic signatures presented by Thakur et al 2019 (112), with comparable prognostic value to sentinel node biopsies.

The 25 individually selected biomarkers were previously reported to positively or negatively influence survival outcomes (column pro-tumour (+), anti-tumour (-) on Table 3. In Table 3 we reported the beta coefficient, Hazard ratio and raw p-value of the multivariable Cox regression for each gene where we adjusted for sex, age and stage in the LMC combined subgroup of females and males and adjusted only for age and stage on the sex-specific subgroups.

When examining these 25 individual biomarkers across a combined population of females and males, 19 out of 25 exhibited significant p-values, with all 25 demonstrating beta coefficients consistent with existing literature findings. The analysis involving female-specific data showcased a noteworthy consistency with 25 beta coefficients aligned with existing literature findings, and 22 out of these 25 beta coefficients having significant p-values, further confirming their statistical power. In contrast, the same analysis on male stage I-II melanomas revealed that no gene out of the 25 biomarkers had raw p-value lower than 0.1.

Also, in males the Hazard ratio (HR) of these 25 genes consistently hovered around 1, indicating a mostly neutral effect on survival. On the contrary, mainly for stage I-II female melanomas and slightly for the combined female and male group, the HR deviated further away from 1, suggesting a stronger impact on survival.

Table 3: This table presents sex-specific multivariable Cox regression results in females and males together (adjusted fro age,stage and sex) and separately in females and males (adjusted for age and stage) with stage I-II melanomas, including beta coefficients (Coef), hazard ratios (HR), and raw p-values. The third column, labelled pro-tumor (+) and anti-tumor (-), denotes the reported role of each biomarker in existing literature. The 'OGS' column corresponds to the official gene symbol, while the 'Function' column specifies the identified functional involvement of each gene.

			Female/Males stage I,II			Females stage I,II			Males stage I,II		
OGS	Function	pro-tumor (+), anti-tumor(-)	Coef	HR	raw_p-val	Coef	HR	raw_p-val	Coef	HR	raw_pval
CKS2	cell cycle	+	0.2650	1.3034	0.0042	0.4759	1.6095	0.0000	-0.0302	0.9703	0.8311
CDC2	cell cycle	+	0.2029	1.2250	0.0388	0.4639	1.5902	0.0008	-0.0458	0.9553	0.7392
CCNB1	cell cycle	+	0.3879	1.4738	0.0004	0.6559	1.9268	0.0000	0.0994	1.1045	0.5133
CENPF	cell cycle	+	0.3080	1.3607	0.0000	0.5741	1.7755	0.0000	0.0600	1.0619	0.5703
DHFR	cell cycle	+	0.3900	1.4770	0.0065	0.5098	1.6650	0.0163	0.3035	1.3546	0.1195
HCAP-G	mitosis	+	0.2015	1.2233	0.0398	0.4777	1.6124	0.0010	-0.0585	0.9432	0.6352
STK6	mitosis	+	0.1982	1.2192	0.0067	0.3275	1.3875	0.0006	0.0447	1.0457	0.6828
BUB1	mitotic spindle ckeckpoint	+	0.1599	1.1734	0.0196	0.3853	1.4701	0.0000	-0.0685	0.9338	0.4790
BIRC5	inhibition apoptosis	+	0.1891	1.2082	0.0418	0.5391	1.7145	0.0000	-0.1305	0.8776	0.3664
P2RY14	stimulation of apoptosis	-	-0.2929	0.7461	0.0001	-0.5292	0.5891	0.0000	-0.1029	0.9022	0.3326
TOP2A	DNA replication	+	0.2255	1.2529	0.0061	0.3890	1.4754	0.0005	0.0731	1.0759	0.5059
RRM2	DNA replication	+	0.1920	1.2116	0.0234	0.3961	1.4860	0.0011	0.0147	1.0149	0.8988
TYMS	DNA replication	+	0.2455	1.2782	0.0177	0.5612	1.7528	0.0001	-0.0607	0.9411	0.6729
PCNA	DNA replication	+	0.1334	1.1427	0.2016	0.4745	1.6072	0.0006	-0.2144	0.8071	0.1339
MCM4	DNA replication	+	0.4520	1.5715	0.0000	0.7307	2.0765	0.0000	0.1848	1.2029	0.1780
MCM6	DNA replication	+	0.6835	1.9808	0.0007	1.0092	2.7433	0.0001	0.3025	1.3533	0.2842
GLRX2	stress response	+	0.0485	1.0497	0.7086	0.3829	1.4666	0.0590	-0.1128	0.8933	0.4292
DNAJA1	stress response	+	0.3984	1.4894	0.0563	0.6050	1.8312	0.0492	0.2104	1.2341	0.4573
HSPA4	stress response	+	0.2570	1.2931	0.0963	0.4274	1.5332	0.0557	0.1209	1.1285	0.5646
HSPA5	stress response	+	0.2692	1.3089	0.1325	0.4737	1.6060	0.0699	0.1341	1.1435	0.5244
HSPD1	stress response	+	0.8742	2.3970	0.0006	1.4945	4.4573	0.0000	0.1102	1.1165	0.7593
TXNIP	stress response	-	-0.1702	0.8435	0.2758	-0.5559	0.5735	0.0219	0.0893	1.0934	0.6662
CACYBP	ubiquitin cycle	+	0.2465	1.2796	0.0444	0.5909	1.8057	0.0003	-0.0650	0.9371	0.7050
CNN3	Aactin and calmodulin binding	+	0.2963	1.3449	0.0154	0.4642	1.5908	0.0250	0.2281	1.2562	0.1336
STMN2	intracellular signalling	-	-0.0954	0.9090	0.0156	-0.1388	0.8704	0.0099	-0.0517	0.9496	0.3678

This analysis was also repeated together and separately in LMC female and male stage III melanomas adjusted for age, stage and sex for the combined sex population and only for age and stage in females and males only (Appendix Table 6). The 25 genes, both in stage I-II and stage III were plotted using a volcano plot with their female or male beta coefficients and the corresponding p-values (Figure 3, 4). In stages I-II, as expected, only the analysis specific to females and females plus males combined, yielded results surpassing the p-value threshold, displaying beta coefficient values markedly different from those observed in males.

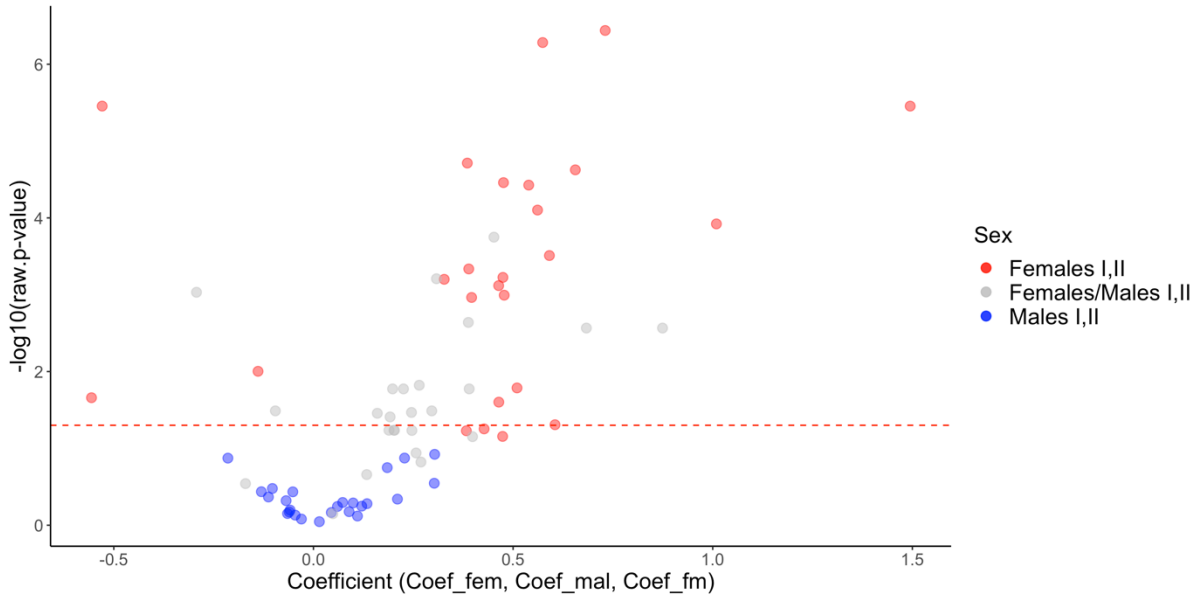


Figure 3: The 25 literature found melanoma biomarkers plotted according to their beta coefficients and raw p-value obtained from the multivariable Cox regression analysis together in females and males (grey) and separately in females (red) and males (blue) stage I-II melanomas. Horizontal red line at -log10(0.05).

Interestingly, in stage III analysis (Figure 4), results are more homogenous. There are 4 genes in females and 7 genes in males with a p-value lower than 0.05 and overall, the beta coefficients of both sexes do not differ to the same extent as for stage I-II.

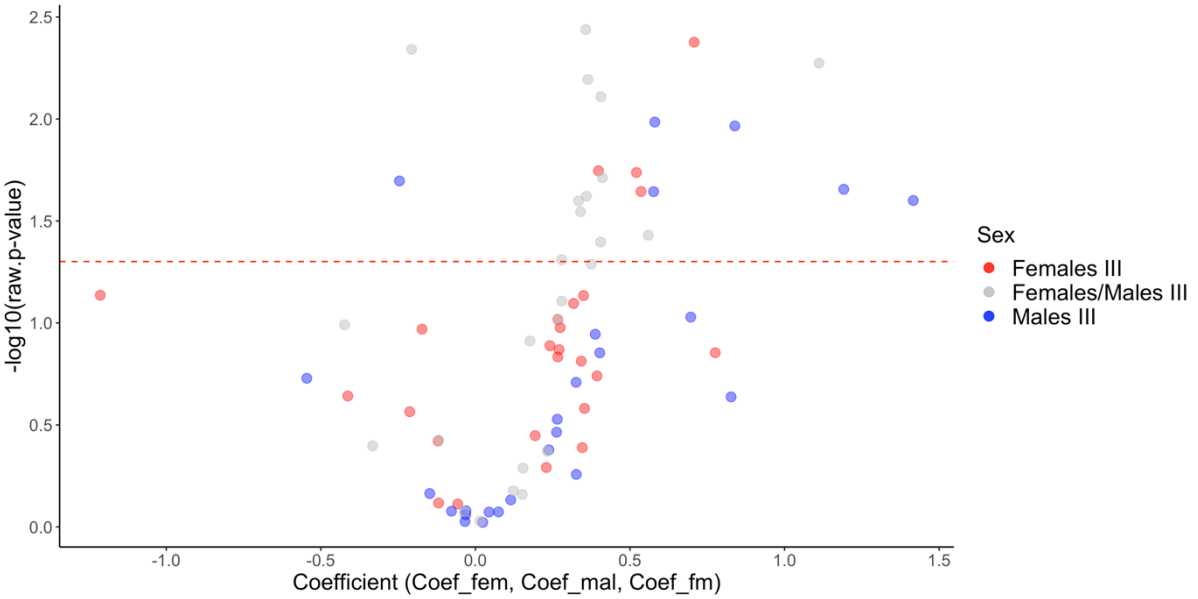


Figure 4: The 25 literature found melanoma biomarkers plotted according to their beta coefficients and raw p-value obtained from the multivariable Cox regression analysis together in females and males (grey) and separately in females (red) and males (blue) stage III melanomas. Horizontal red line at -log10(0.05).

1.2 The CP-GEP commercial CM signature

The 8-gene commercial signature CP-GEP consisting of ITGB3, PLAT, SERPINE2, CDF15, TGFBR1, LOXL4, CXCL8 and MLANA was evaluated in a multivariate manner where all genes were taken together. The analysis was performed on stage I,II and stage III females and males, both together and separately. The model on the combined population of stage I,II was adjusted for age, stage and sex while for the single-sex population it was adjusted on age and stage, the same adjustments, except stage, were also applied for stage III.

Table 4: The table presents the results of the C-index, likelihood ratio test, Wald test and score (logrank) test of multivariate Cox regression model of stage I,II (A) and stage III (B) performed on females and males together and on females-only and males-only.

A)	CP-GEP 8 gene signature LMC stage 1,2		
	Female/Males	**Females**	**Males**
C-index	0.706	0.725	0.699
Likelihood ratio test	69.28, p=2e-10	36.39, p=7e-05	32.66, p=3e-04
Wald test	61.98, p=4e-09	33.34, p=2e-04	29.27, p=0.001
score (logrank) test	64.63, p=1e-09	35.37, p=1e-04	30.52, p=7e-04

B)	CP-GEP 8 gene signature LMC stage 3		
	Female/Males	**Females**	**Males**
C-index	0.704	0.726	0.741
Likelihood ratio test	25.87, p=0.004	18.53, p=0.03	6.65, p=0.05
Wald test	25.97, p=0.004	16.3, p=0.06	12.81, p=0.2
score (logrank) test	27.37, p=0.002	18.41, p=0.03	14.81, p=0.1

Overall, the evaluation of the CP-GEP 8 gene signature across stages 1, 2, and 3 of LMC allowed a mild sex-specific performance difference. In stages 1 and 2, females exhibited superior predictive ability, as indicated by a higher C-index (0.725) compared to males (0.699) (Table 4A). In stage 3, when comparing the C-index of all the three groups (females and males, females-only and males-only) the higher C-index was observed in the male-only population (0.741) (Table 4B). The predictive ability of the statistical tests was higher in the combined population of females and males. All tests, likelihood ratio, Wald and logrank test, had a higher χ^2 across all tests when compared to the single-sex populations in both stage 1,2 and stage 3. These results indicate that sex-specific predictive ability of the 8-gene CP-GEP signature varies according to stage.

1.3 The 26 gene signature of Gerami et al 2015.

The next signature that was assessed in a multivariable Cox regression model was the 27-gene signature published by Gerami et al. We only assessed 26 out of the 27 genes as the gene MGP was not present in the LMC cohort. This signature was accessed in both stage 1,2 adjusted for age and stage and stage 3 adjusted for age. Additionally for the combined population of females and males, the model was adjusted on sex, age and stage.

Table 5: The tables present the results of the C-index, likelihood ratio test, Wald test and score (logrank) test of multivariate Cox regression model initially performed on females and males together and on females-only and males-only of stage I,II (A) and stage III (B).

A)	Gerami et al 26 gene signature tested on LMC stage 1,2		
	Female/Males	**Females**	**Males**
C-index	0.716	0.759	0.711
Likelihood ratio test	82.72, p=5e-07	54.62, p=0.002	45.48, p=0.02
Wald test	81.06, p=8e-07	52.1, p=0.004	42.79, p=0.04
score (logrank) test	86.61, p=1e-07	59.55, p=5e-04	45.08, p=0.02

B)	Gerami et al 26 gene signature tested on LMC stage 3		
	Female/Males	**Females**	**Males**
C-index	0.738	0.856	n/a
Likelihood ratio test	41.77, p=0.05	53.48, p=0.002	n/a
Wald test	36.16, p=0.1	29.58, p=0.3	n/a
score (logrank) test	41.88, p=0.04	48.84, p=0.006	n/a

The results observed when comparing the multivariable Cox regression models in females and males together and on sex-stratified groups, are very similar to the CP-GEP results. The C-index values for the combined Female/Males group, Females-only group, and Males-only group were 0.716, 0.759, and 0.711, respectively for stage I,II (Table 5A). Likewise, the statistical tests, including the Likelihood Ratio Test, Wald Test, and Score (Logrank) Test, demonstrated similar trends with the higher χ^2 being observed in the combined population of female and males and lower in the male-only population. For instance, the Likelihood Ratio Test yielded a $\chi^2 = 82.72$, p=5e-07 for Female/Males, $\chi^2 = 54.62$, p=0.002 for Females, and $\chi^2 = 45.48$, p=0.02 for Males, reaffirming the model's predictive strength across mainly the combined and the female-only groups. Similarly, both the Wald Test and Score (Logrank) Test exhibited consistent patterns of significance across sex-based cohorts, with p-values aligning closely with previous analyses. Additionally, the same signature was also assessed on LMC stage III, with the female-only group showing a stronger predictive ability with a C-index of 0.856 compared to the 0.738 obtained from the combined population (Table 5B). There are no results for the male-only group as the multivariable Cox regression model did not converge because of the high number of variables. To further assess these 26 genes we also evaluated

them individually using a univariable Cox regression analysis in the three groups (females/males, females-only, males-only), showing that in the combined group there were 4 significant genes with a raw p-value < 0.05, in females-only there were 9 significant genes and 0 significant genes were found in males (Appendix Table Gerami et al).

1.4 The 6 consensus-based classes created by Thakur et al 2019.

The last signatures that we evaluated are the 6 consensus-based gene classes created by Thakur et al 2019 predicting survival outcome in the whole LMC dataset and in stage I disease. We only assessed them in our group of stage1,2 and not stage 3 and adjusted for age and stage. We evaluated the key nodal genes up- regulated in each of the 6-classes. The following genes were identified for each class:

- Class 1: Up-regulated NFKB1, FYN, ITGB2, STAT1, RAC2
- Class 2: Up-regulated RAD21, RPL10A, SMAD2
- Class 3: Up-regulated MYC, RNF2, RANBP2, PLCG1
- Class 4: Up-regulated HDAC2, RNF2, TCF7L2, MAPK11
- Class 5: Up-regulated EGFR, HRAS, MAPK13
- Class 6: Up-regulated JUN, NFKB1, HDAC1, PLK1, HCK

These class gene-signatures were evaluated in multivariable Cox regression model is adjusted for age and stage. In all 4 out of 6 classes, the C-index is higher when model is run on females-only compared to the combined population and males-only. Classes 1,2,3 and 6 display a higher difference in the C-index between the groups compared to class 4 and 5, where the models display a higher predictive ability in males-only (Figure 5).

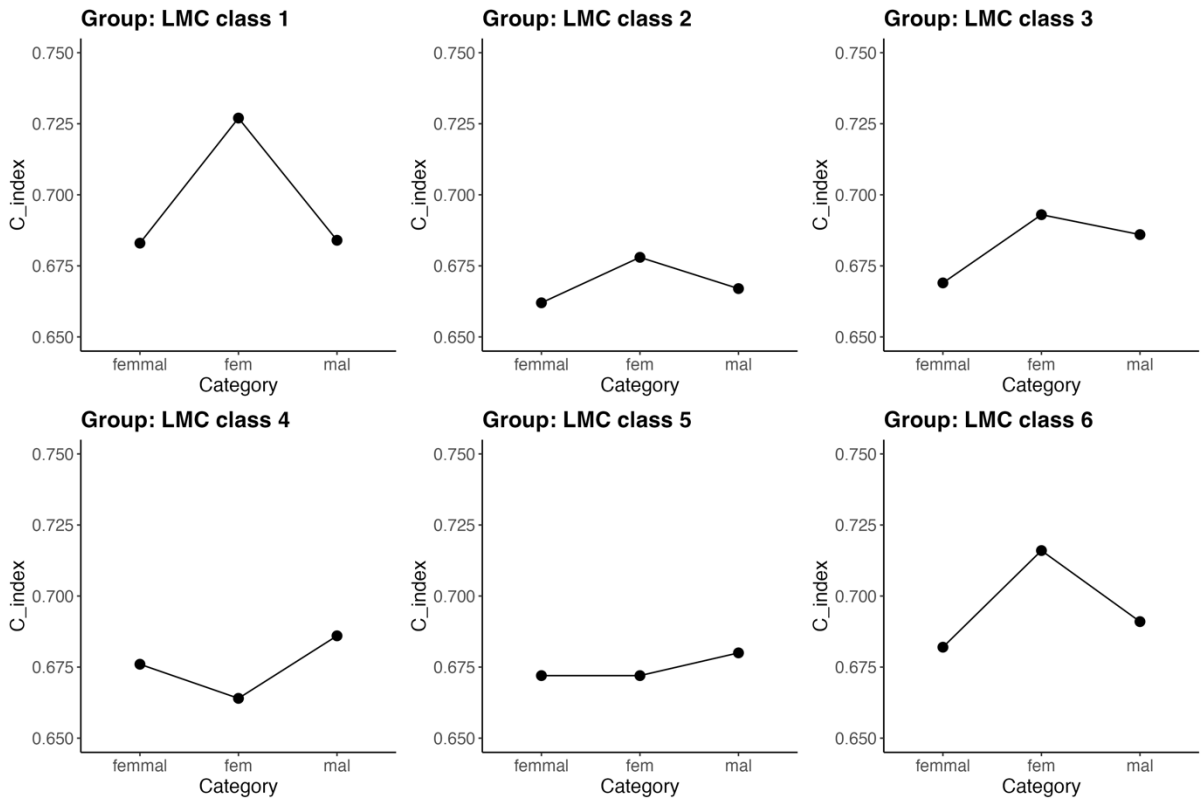


Figure 5: C-index Results from Multivariate Cox Regression Models for Key Nodal Genes of Each LMC Class. Plots represent C-index results obtained from the multivariable Cox regression model for each class of key nodal genes. The C-index values are shown for three categories: the combined population of females and males ("femmal"), females-only ("fem"), and males-only ("mal"). Each plot corresponds to one of the six LMC classes (Class 1 to Class 6).

Sex-stratification analysis on existing survival biomarkers not only reaffirmed their prognostic value in the combined, female and male, cohort but also highlighted significant difference in their impact when patients are stratified by sex. Notably, the analysis revealed a stronger association of these biomarkers with survival outcome in females compared to males. Furthermore, these results highlight the need of accounting for sex as a biological variable and correctly accounting for it in survival biomarker analysis.

Summary Chapter 4

The sex-stratified analysis of previously published CM biomarkers revealed distinct sex-specific patterns in their prognostic value. Significant associations between certain biomarkers and survival were found in the combined sample of females and males, but notable differences emerged when analysed separately by sex. In stage I-II CM, 22 out of 25 individual biomarkers

showed significant p-values in females, with hazard ratios deviating from 1, suggesting a stronger impact on survival. In contrast, none of the biomarkers were significant in males, and hazard ratios were mostly neutral. In stage III CM, differences between sexes were less pronounced, with 4 genes significant in females and 7 in males.

The evaluation of the multidimensional signatures, CP-GEP 8-gene signature and the 26-gene signature by Gerami et al., also showed sex-specific performance differences. Females exhibited superior predictive ability in stages I-II, while males showed higher predictive ability in stage III. Four of the sex consensus-based classes by Thakur et al. revealed higher C-index values in females once more, underscoring sex-specific predictive abilities.

These findings highlight the importance of considering sex as a biological variable or stratifying by sex in the discovery of survival biomarkers. Incorporating sex-specific analyses can improve our understanding of biomarker impact on CM female and male survival and enhance early detection of recurrence and hence personalized treatment strategies.

Chapter 5: Explore sex-specific differences in primary CM, normal skin and Nevi samples.

Chapter five focuses on investigating further sex-specific gene expression differences in CM, normal skin (both sun not-sun exposed) and nevi samples. Initially, we employed a multifaceted approach that includes Kaplan-Meier (KM) analysis, principal component analysis (PCA), analysis of gene expression variability (GEV) and identify the differentially expressed genes (DEG) for datasets, LMC, TCGA, GSE53118, GTEx and Nevi, comparing female vs males. In addition, this part of the thesis explored the immune cell differences between females and males within LMC using an immune deconvolution technique, xCell. Chapter 5 aims to shed light on sex-related disparities in normal skin, nevi and CM.

2.1 Sex stratified Kaplan-Meier plots

Due to the substantially larger sample size compared to the other datasets and to the availability of complete follow-up with overall survival information, the LMC dataset was the most suitable to run survival analysis and look for prognostic biomarkers. Furthermore, since melanoma diagnosis in the LMC dataset occurred between 2000 and 2012 and no metastatic case was analysed, we can be confident that survival information is not affected by the effects of targeted or immunotherapy. Death for melanoma was the only death-cause considered. Only 26 patients died for other reasons and were therefore excluded from survival analysis.

For stages I, II, and III combined, a significantly better overall survival was seen in females (Figure 6), with a log-rank test *p*-value of 0.01. To further investigate the effect of sex on overall survival, samples were stratified based on stage. Kaplan-Meier curves of stage I and stage II (log-rank test *p*-values of 0.043 and 0.06) revealed a stronger correlation of sex in overall survival compared to stage III (log-rank test *p*-values of 0.51).

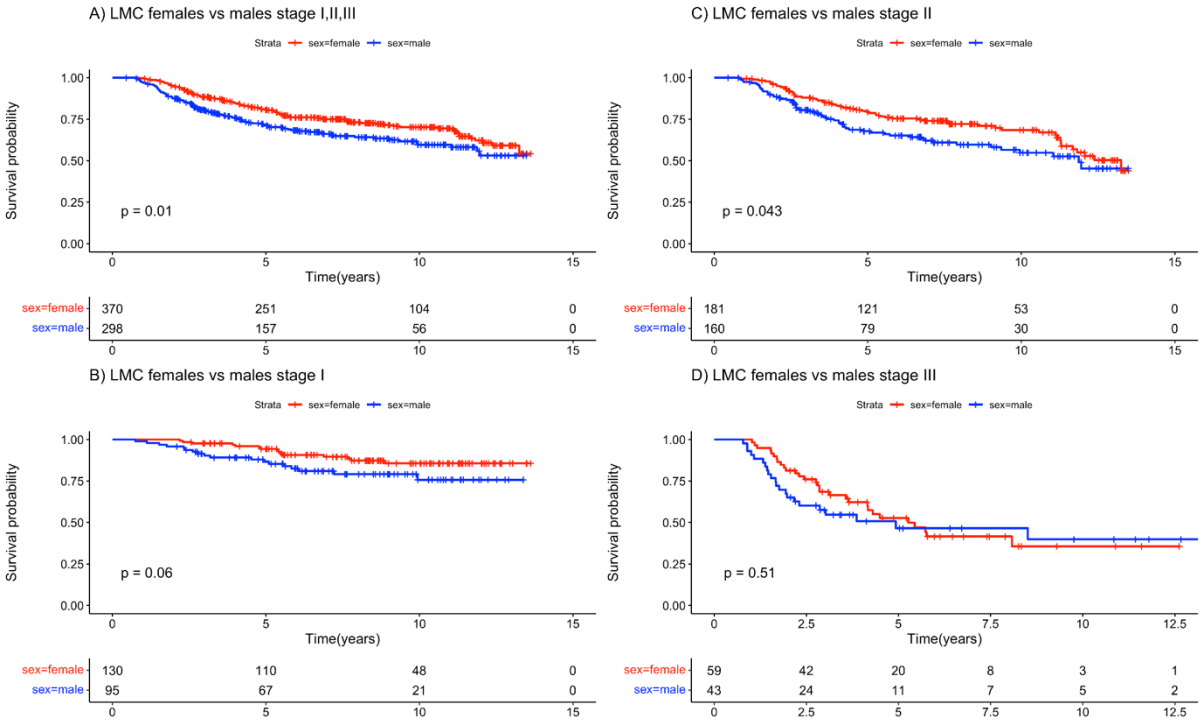


Figure 6: Kaplan - Meir overall survival curves with the log-rank test p-value between females and males for the LMC dataset on A) stages I-II and III; B) stage I; C) stage II; and C) stage III.

2.2 Quantifying the effect of sex in survival analysis of CM

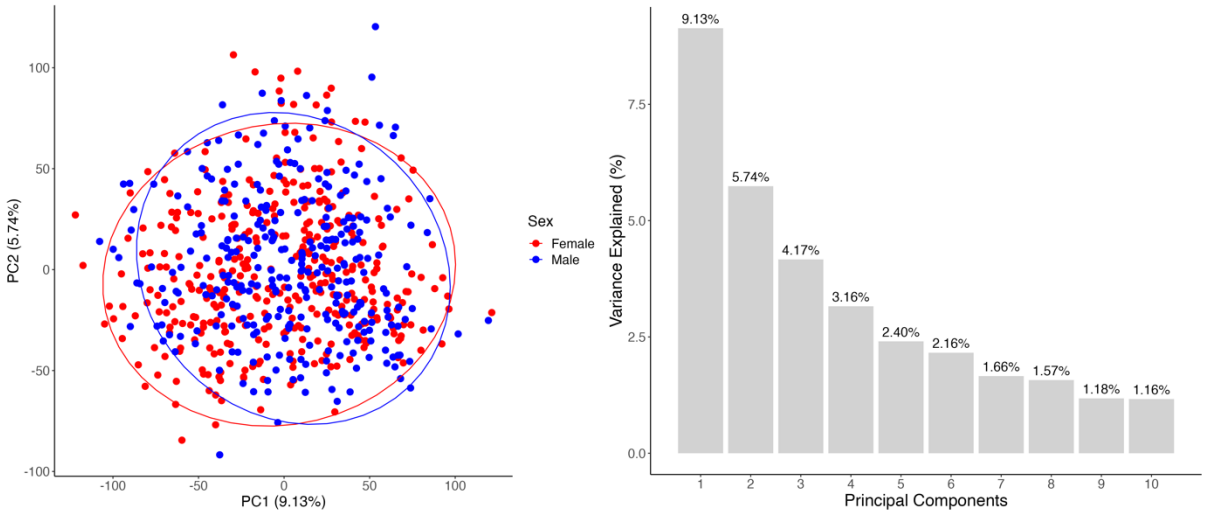
Univariable Cox regression analysis for sex confirmed that males have higher risk of death for melanoma than females. Hazard ratios (HR), confidence intervals (CI) and p-values of all Cox regression analyses are reported in Table 4. Hazard ratio is higher and p-values are lower for stage I and stage II compared to stage III and all the stages combined (Table 6). When stage I-II combined HR, for males is 1.59 (CI 1.16 - 2.19), with p-value = 0.004.

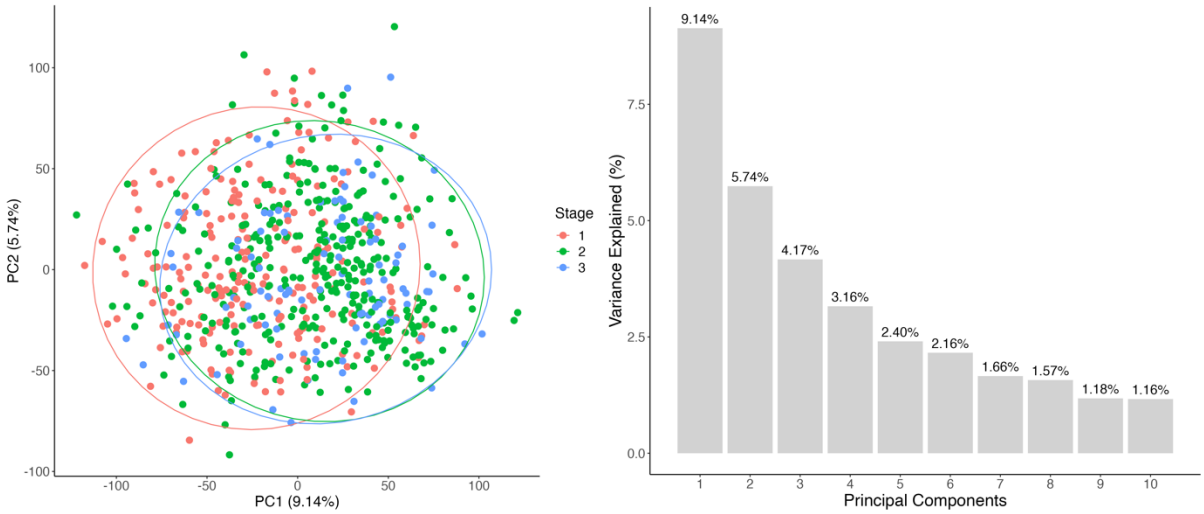
Table 6: Cox regression analysis results for sex in the female-male subgroups; stage I,II,II; stage I; stage II; stage I-II and stage III. This table presents the hazard ratios (HR) with 95% confidence intervals (CI) and p-values from the Cox regression analysis for sex across various stages of melanoma.

	LMC		
Characteristic	HR	95% CI	p-value
stage I,II,II	1.43	1.09-1.88	0.011
stage I	1.91	0.96-3.80	0.064
stage II	1.45	1.01-2.08	0.044
stage I,II	**1.59**	**1.61-2.19**	**0.004**
stage III	1.20	0.69-2.09	0.500

2.3 Principal component analysis (PCA) on gene expression data

PCA is a tool used in the initial steps of building a prediction model, since it helps in the exploratory data analysis for variable dimensionality reduction. In this project PCA was initially considered a data exploration technique to identify potential patterns or clusters within the data, which would facilitate the identification of sex specific survival biomarkers. We first applied it to the entire LMC melanoma cohort (Figure 7) and then to all the other analysed cohorts (Table 7).





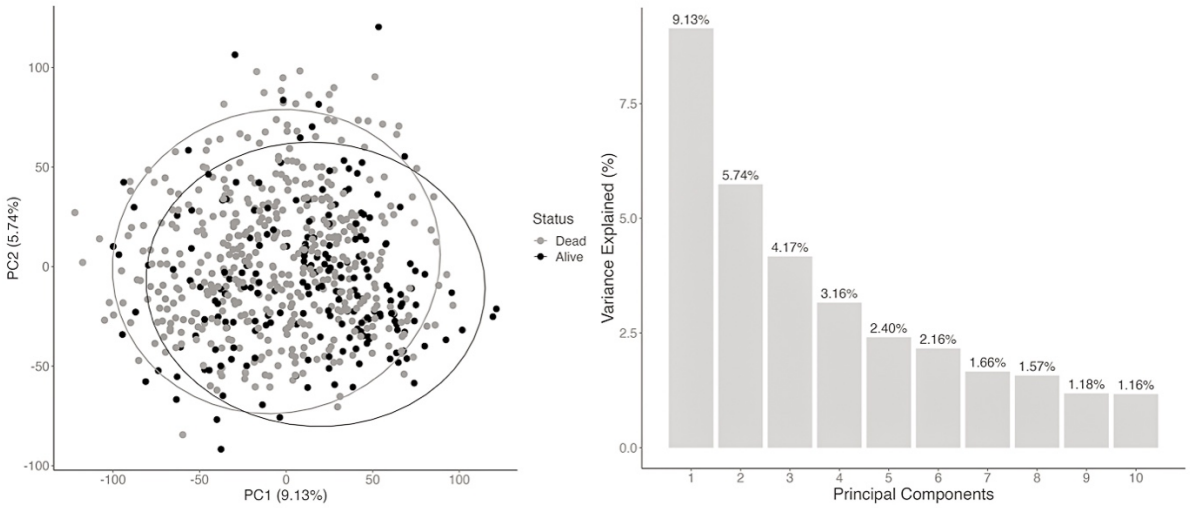


Figure 7: Principal component 1 (x axis) and 2 (y axis) from PCA analysis applied to the entire LMC melanoma cohort. Samples are coloured according to sex (top), stage (middle) or status (alive/dead) (bottom).

Table 7: Variance explained by the first five principal components in LMC, TCGA, GSE53118 and GSE65904 melanoma cohorts, in GTEx – not sun exposed (NSE) and – sun exposed (SE) normal skin, and in Nevi samples (GSE46517 and GSE3189).

	PC1	**PC2**	**PC3**	**PC4**	**PC5**
LMC stage I,II,III	9.13%	5.74%	4.17%	3.16%	2.40%
TCGA stage I,II,III	17.52%	10.73%	5.50%	3.71%	3.43%
GSE53118 stage III	23.43%	10.01%	6.46%	5.15%	4.18%
GSE65904 primary mel.	18.47%	13.57%	11.29%	9.37%	9.07%
GTEx - NSE	11.43%	7.17%	4.29%	3.59%	2.90%
GTEx - SE	8.64%	4.82%	3.66%	3.46%	2.79%
Nevi (GSE46517, GSE3189)	20.97%	10.96%	6.81%	6.05%	5.40%

PCA analysis on the LMC cohort does not show differences between groups when looking at the components both numerically and visually. For example, when plotting PC1 against PC2 and colour coding females and males, different stages, or status (dead from melanoma vs alive), very slight differences are observed. This is evident from the overlapping of positioned ellipses representing these groups (Figure 7).

The scree plot (Figure 7, top), which shows the amount of variance explained by each principal component, provides insight into the results. Specifically, PC1 accounts for 9.13% of the

variance, PC2 for 5.74%, PC3 for 4.17% and so on. This pattern is consistent across datasets, with PC1 consistently explaining less than 24% of the variance, suggesting a general trend of low variance explained by the primary principal components across datasets (Table 7).

Moreover, upon conducting separate PCAs for females-only and males-only within stage I,II,III, a slight difference emerged in PC1 and PC2 between sexes. In females, PC1 accounted for 10.08% and PC2 for 5.77%, while in males, PC1 explained 8.48% and PC2 5.87% of the variance. This sex separation highlighted only a moderate difference in variance captured among females compared to males (Appendix Figure 1).

Taken together this information indicates that the principal components capture a modest amount of variability across datasets. We therefore did not use this approach to reduce variable dimensionality.

2.4 Gene expression variability

To further investigate this disparity between females and males, we calculated the gene expression variability (GEV) between the two sexes in all cohorts. This analysis was inspired by a meta-analysis performed on physiological tissues showing a slightly lower gene expression variability in females compared to males (170). Considering the importance of GEV between sexes in literature and its association with tumour aggressiveness, we wanted to investigate its role in stage I,II and stage III CM, nevi and normal skin samples. For each gene we calculated the coefficient of variation across female samples only and across male samples only. The CVs were compared between the sexes using a paired Wilcoxon two-sample test. The mean CV of female and of males was reported including the Wilcoxon test p-value (Table 8).

Table 8: Average coefficients of variation obtained from the gene expression data of melanoma (LMC, TCGA, GSE53118) and non-melanoma (Nevi, Normal skin) datasets were calculated for both females and males in the stage stratified groups. A Wilcoxon (WC) paired test between females and males was performed on the total number of genes included in every subcategory. NSE = non sun exposed, SE = sun-exposed

	Groups	Mean CV	WC p-value
LMC	Females stage I-II	0.09739	0.00E+00
	Males stage I-II	0.10074	
	Females stage III	0.10642	0.00E+00
	Males stage III	0.09944	
TCGA	Females stage I-II	0.20385	1.79E-19
	Males stage I-II	0.21005	
	Females stage III	0.22972	2.47E-235
	Males stage III	0.21225	
GSE53118	Females stage III	0.04719	0.00E+00
	Males stage III	0.04477	
GSE65904	Females (Primary Melanoma)	0.16848	0.00E+00
	Males (Primary Melanoma)	0.20978	
GTEx	Females (NSE)	0.24439	5.98E-259
	Males (NSE)	0.24849	
	Females (SE)	0.24159	0.004940886
	Males (SE)	0.24164	
Nevi (GSE46517 + GSE3189)	Females	0.11753	0.00137
	Males	0.12157	

Significantly different CV was observed between melanoma samples of males and females in both stage groups (stage I-II and stage III CM), in all datasets. A consistently lower GEV was observed in early stage melanoma (stage I-II) in females compared to males, in both SE and NSE normal skin, and in Nevi samples. On the other hand, the opposite was observed in stage

III melanomas for LMC, TCGA and GSE53118 datasets, where males had a statistically significant lower GEV compared to females.

To further explore GEV in our largest cohort, LMC, density plots were drawn for CVs obtained after bootstrapping. Random groups of 3000 genes each (10,000 iterations) were created and the average CVs of each group in females and males for stage I-II and for stage III melanomas were plotted (Figure 8).

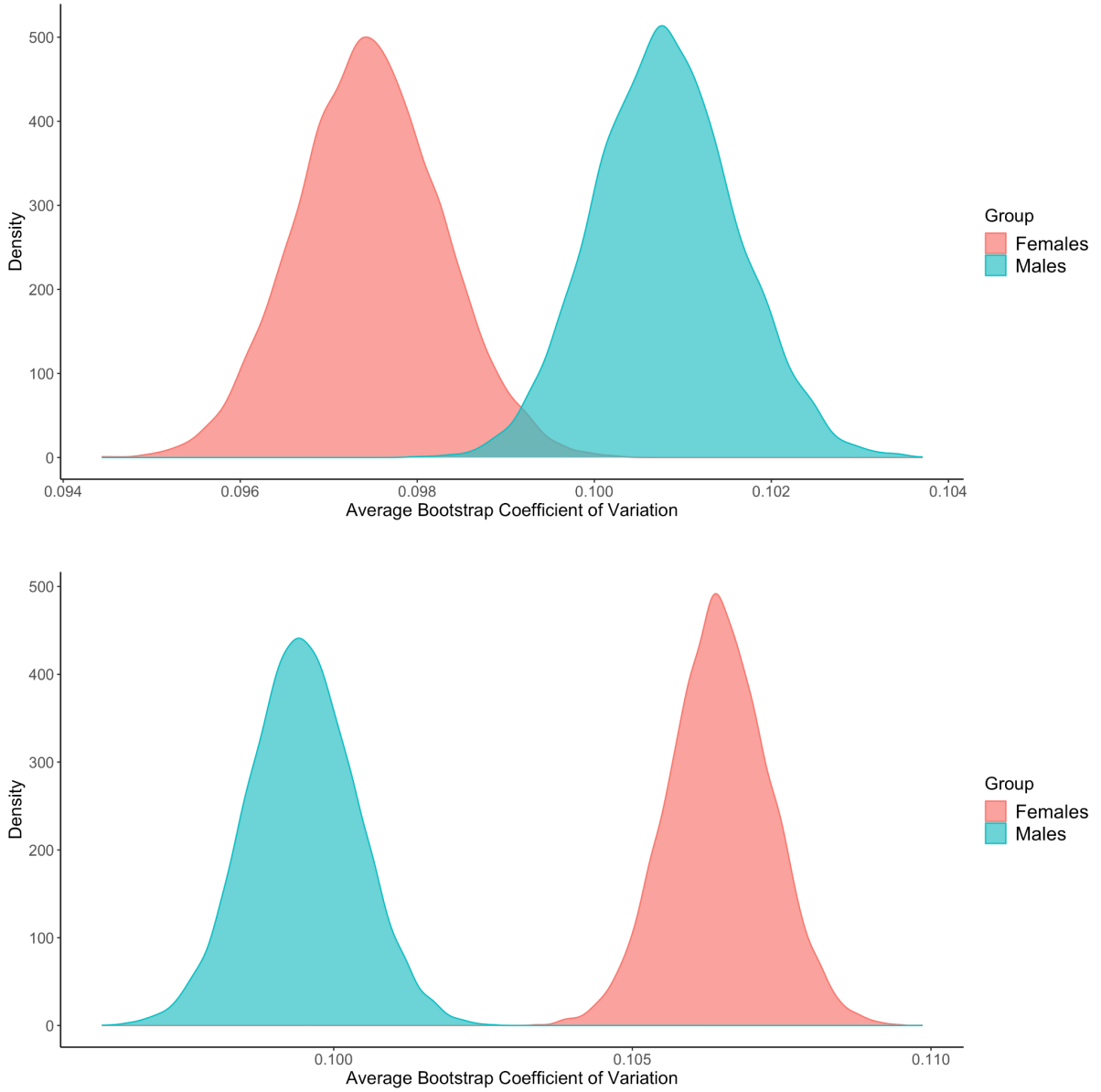


Figure 8: Density Plots of Average Bootstrap Coefficients of Variation from Gene Expression Data. Bootstrap analysis was performed on random subgroups of 3000 genes, repeated 10,000 times. The density plots show the average coefficients of variation (CVs) for females (red) and males (blue). The top plot represents the data for stage I-II, and the bottom plot represents the data for stage III melanomas.

In addition, to further investigate the GEV between females and males within multiple biological pathways, we choose 15 biological processes related to either immune, metabolism or cell cycle processes. Our goal was to select those pathways exhibiting the highest number of genes per pathway from the Gene Set enrichment analysis (GSEA) website. These pathways were contained in the gene ontology sets (C5) of the Human Molecular Signatures Database (MSigDB) and consisted of a total number of genes ranging from 240 to 1896 (Table 9).

Table 9: Gene Ontology Biological Pathways and Coefficient of Variation Analysis. The table lists 15 gene ontology biological pathways along with the total number of genes in each pathway. The median coefficients of variation (CV) among the total genes for females and males in stage I-II are provided. The Wilcoxon paired test (WC p-value) compares the CVs between females and males with stage I,II melanomas.

		Median CV		
General biological pathways	Total number of genes	Females stage I-II	Males stage I-II	WC p-value
GOBP_CELL_CYCLE	1702	0.0848	0.0892	2.20E-16
GOBP_ACTIVATION_OF_IMMUNE_RESPONSE	302	0.1024	0.1053	6.97E-05
GOBP_PROTEIN_PROCESSING	240	0.0934	0.0956	1.09E-06
GOBP_TISSUE_DEVELOPMENT	1896	0.0991	0.1033	2.20E-16
GOBP_CELLULAR_RESPONSE_TO_STRESS	1857	0.0828	0.0877	2.20E-16
GOBP_PHOSPHORYLATION	1747	0.0914	0.0959	2.20E-16
GOBP_CELL_MOTILITY	1677	0.0967	0.1005	2.20E-16
GOBP_REGULATION_OF_TRANSPORT	1699	0.0922	0.0965	2.20E-16
GOBP_REGULATION_OF_CELL_DEATH	1575	0.0890	0.0933	2.20E-16
GOBP_DNA_METABOLIC_PROCESS	987	0.0817	0.0859	2.20E-16
GOBP_CELL_CELL_SIGNALING	1587	0.0956	0.1012	2.20E-16
GOBP_VESICLE_MEDIATED_TRANSPORT	1477	0.0868	0.0932	2.20E-16
GOBP_DNA_REPAIR	570	0.0794	0.0834	1.86E-14
GOBP_CELLULAR_RESPONSE_TO_DNA_DAMAGE_STIMULUS	849	0.0811	0.0856	2.20E-16

These analyses demonstrate that the consistently lower GEV observed in females extends beyond the mentioned datasets (including all genes) and the randomly generated 3000-gene subsets through bootstrapping. This trend persists even within randomly selected biological pathways (Table 9), highlighting the ubiquitous nature of the lower GEV in female stage I-II melanomas across various datasets and pathways and the necessity of its further investigation/consideration. We therefore decided to consider sex as an effect modifier instead of a confounder variable to adjust for.

2.5 Class comparison analysis

Differential expression between the sexes was analysed on the three melanoma datasets as well as on the Nevi and normal skin datasets (Chrysanthou et al. 2022 (171) - Additional file 1). Details on the number of overexpressed genes in each of the sexes can be seen in Table 10. Interestingly, the number of differentially expressed genes between sexes showed a decreasing trend from normal skin to nevi and from early-stage to late-stage melanomas. Moreover, unlike the normal skin dataset and the melanoma stage I, II groups, very few differentially expressed genes were observed between sexes within the NEVI datasets.

Table 10: Up-regulated autosomal and sex chromosome genes in females and males. The table summarizes the number of up-regulated autosomal and sex chromosome genes identified through class comparison analysis across various datasets. The melanoma gene expression datasets (LMC, TCGA, GSE53118) are further stratified by stage. For the GTEx datasets, data are provided for non-sun exposed (NSE) and sun-exposed (SE) conditions.

		Up-regulated autosomal chromosomes genes		Up-regulated autosomal sex chromosomes genes	
	Stage	**Females**	**Males**	**Females**	**Males**
LMC	I, II, III	45	24	67	54
	I, II	27	17	64	4
	III	0	0	0	0
TCGA	I, II, III	39	18	17	3
	I, II	77	37	23	5
	III	39	21	16	2
GSE53118	III	0	0	5	0
Nevi (GSE46517+GSE3189)	N/A	0	0	3	0
GTEx NSE	N/A	604	939	N/A	N/A
GTEx SE	N/A	924	1018	N/A	N/A

The overexpressed genes within each sex stratified by stage were selected for gene ontology analysis by DAVID. Gene ontology analysis on autosomal chromosome overexpressed genes within the groups of interest from the LMC dataset showed significant enrichment for the fat cell differentiation process in females stage I, II. On the other hand, overexpressed genes in males stage I, II were significantly enriched, mainly in immune response and protein regulation (Chrysanthou et al. 2022 - Tables S20 and S21).

2.6 Sex-stratified immune deconvolution analysis using xCell

xCell is a computational method used to estimate the presence of different immune cell types within a complex tissue sample, based on gene expression data. This method, unlike other immune deconvolution methods using only one reference gene expression profile, has multiple cell-signatures for every immune cell type. In addition, the output does not represent absolute proportions of immune cells like other techniques but provides an enrichment score. The enrichment score from xCell quantifies the relative abundance of various cell types in a tissue sample based on specific gene expression profiles and a curated gene set matrix. This analysis was used to compare the enrichments of immune cell subtypes in: 1. female vs male stage I-II CM; 2. female vs male stage III CM.

The immune cell subtypes highly enriched in female compared to male stage I-II CM are T cell CD4+ Th2, T cell CD4+ naive, the general stroma score, mast cells, M2 macrophages, Hematopoietic stem cells, granulocyte-monocyte progenitors, eosinophils and B cells plasma. These immune subtypes are part of the adaptive immune response, the inflammation and tissue maintenance process, and hematopoietic and stem cell maintenance functions. As for the subtypes that are highly enriched in male stage I-II CM, these include T cell regulatory (Tregs), T cell CD8+ naive, T cell CD4+ Th1, T cell CD4+ memory, T cell CD4+ effector memory, T cell CD4+ central memory, neutrophils, myeloid dendritic activated cells, monocytes, M1 macrophages and generally macrophages, common lymphoid progenitors, class-switched memory B cells and B cell memory. Overall, these immune cells in males participate in Regulatory and memory T cell response, inflammatory response and phagocytosis, immune cell development and antibody production. There are a total of 24 differentially expressed immune subtypes with adjusted p-value lower than 0.05, out of 36 given by xCell, in LMC stage I-II melanoma (Figure 9).

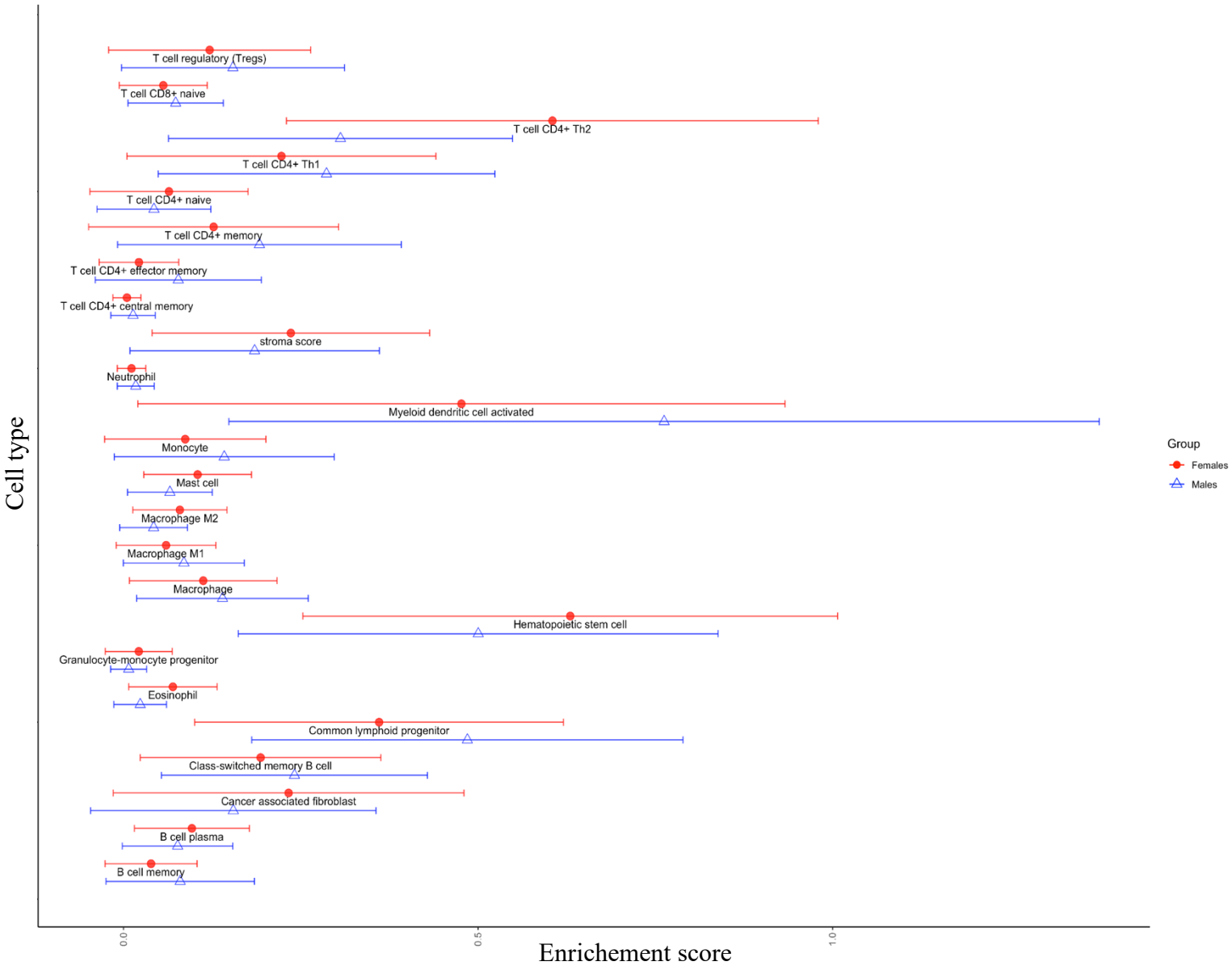
In addition, the analysis was repeated to compare the number of differentially enriched immune cells in female vs male stage III CM in the LMC cohort. Interestingly, only 9 out of the total 36 immune cell types were differentially enriched between the two sexes in stage III. Five cell types were more enriched in females than males: T cell CD8+ naive, T cell CD4+ Th2, T cell CD4+ Th1, Hematopoietic stem cells and B cell naive. As for the immune cell types enriched in males, these include: T cell gamma delta, T cell CD4+ naive, T cell CD4+ effector memory and common lymphoid progenitor cells (Figure 9). In stage III, females exhibit enrichment in

diverse T cell subsets participating in the adaptive immune response. However, male enriched immune cell types participate in immune cell development processes. In addition, to further verify that our observed differences were not due to the different number of cancer cells present between females and males, we used another immune deconvolution package in R called EPIC (estimate proportion of immune and cancer cells). EPIC analysis on the LMC dataset showed that the differences observed in gene expression variability were not attributable to differences in the sampling of the tissues, as the cells labelled by EPIC as "other cells", generally considered to be cancer cells, did not have any significant differences between the groups (Appendix Figure 2).

The analysis was also performed on GTEx normal skin, either not sun-exposed or sun-exposed, with 12 and 6 out of 36 immune cell types being differentially enriched between male and female skin in SE and NSE samples, respectively (Appendix Figure 3 and Figure 4).

These results underscore the complexity of the immune landscape both in stage I-II and stage III melanomas and underline differences between the two sexes, further raising the question about how these variations between the two sexes might influence disease behaviour, overall survival, response to therapy and overall prognosis in a sex-specific manner.

A.


T cell regulatory (Tregs)
T cell CD8+ naive
T cell CD4+ Th2
T cell CD4+ Th1
T cell CD4+ naive
T cell CD4+ memory
T cell CD4+ effector memory
T cell CD4+ central memory
stroma score
Neutrophil
Myeloid dendritic cell activated
Monocyte
Mast cell
Macrophage M2
Macrophage M1
Macrophage
Hematopoietic stem cell
Granulocyte-monocyte progenitor
Eosinophil
Common lymphoid progenitor
Class-switched memory B cell
Cancer associated fibroblast
B cell plasma
B cell memory
Cell type
0.0
0.5
1.0
Enrichement score
Group
Females
Males

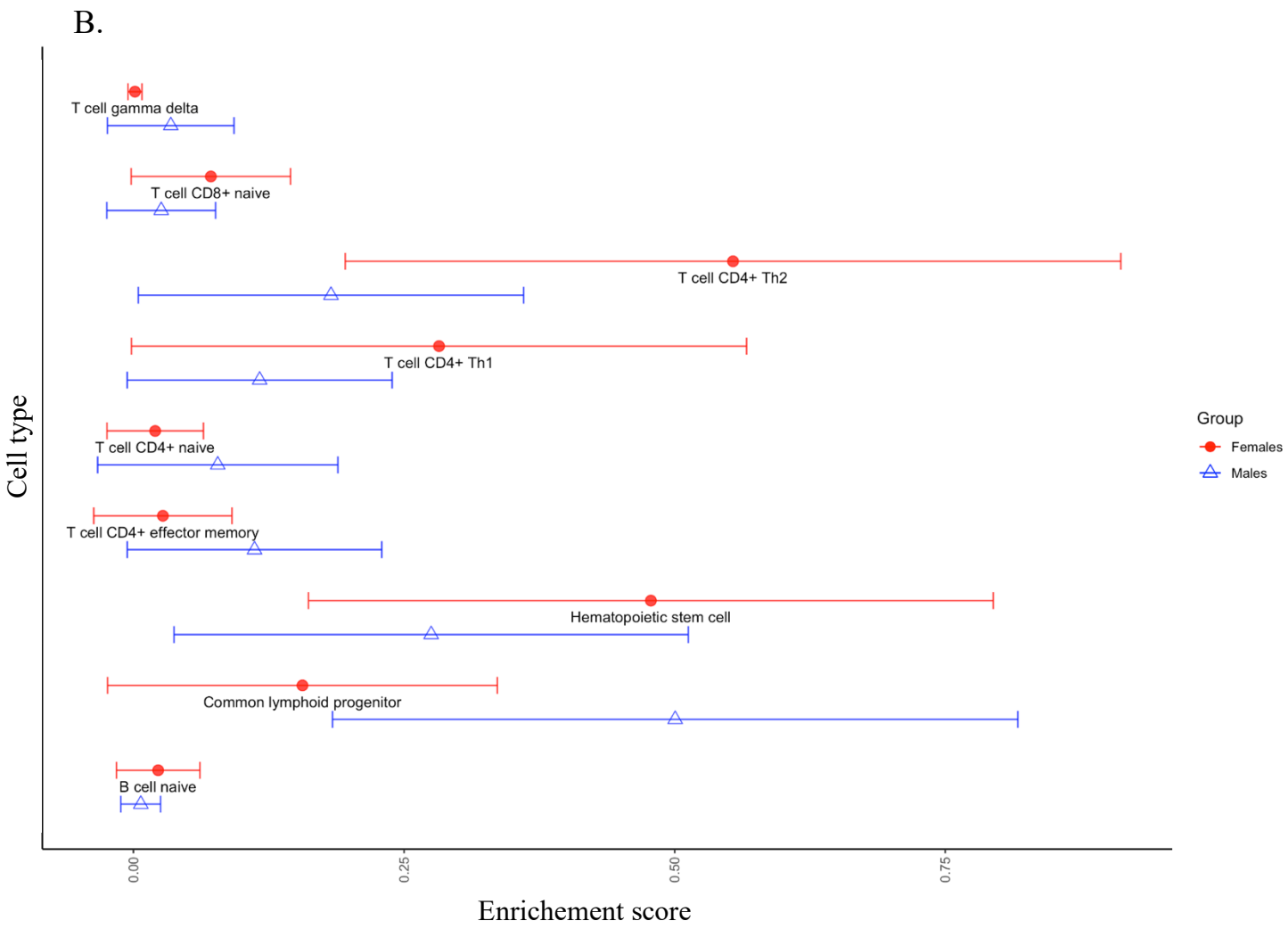


Figure 9: Differential Enrichment of Immune Cell Subtypes in LMC Female and Male Stage I-II (A) stage III (B). Differentially enriched immune cell subtypes between female and male patients with CM from the LMC dataset. The enrichment scores of the significantly different immune cell types are plotted, with females represented by red circles and males by blue triangles. The error bars indicate the variability of the enrichment scores within each group.

2.7 Multivariable Cox Regression Analysis of Immune Cell Subtypes

In order to fully evaluate the role of these immune cell subtypes in the context of survival analysis, we performed a multivariable Cox regression analysis on each cell subtype. This analysis, adjusted for age and stage, used xCell derived enrichment scores as predictors rather than gene expression values. The analysis was carried out in female stage I-II, male stage I-II,

female stage III and male stage III melanomas. Notably, among these subgroups, only female stage I-II samples exhibited a statistically significant association (adjusted p-value <0.05) between immune cell subtypes and survival outcomes (Table 11).

Eight reported immune subtypes alongside the immune and microenvironment scores emerged as statistically significant predictors. All identified subtypes had negative beta coefficients, indicating an association with favourable prognosis (Table 11).

Table 11: Multivariable Cox Regression Analysis in Female Stage I-II Melanomas. The results shown are obtained from a multivariable Cox regression analysis for female patients with stage I-II melanomas from the LMC dataset. The analysis adjusts for age and stage, using enrichment scores provided by xCell as input variables. The table includes the beta coefficient, hazard ratio (HR), 95% confidence interval (CI), and adjusted p-value (Adj. P-value) for each immune cell subtype.

	Females LMC stage I-II			
Immune cell subtypes and scores	**Beta Coefficient**	**HR**	**95% CI**	**Adj. P-value**
immune score	-0.8290	0.4365	[0.27 - 0.69]	0.0106
microenvironment score	-0.7325	0.4807	[0.32- 0.71]	0.0106
T cell CD8+	-1.6415	0.1937	[0.07 - 0.55]	0.0157
T cell CD8+ central memory	-1.4646	0.2312	[0.09 - 0.58]	0.0157
Macrophage	-4.0423	0.0176	[0.001-0.22]	0.0157
Macrophage M1	-6.3003	0.0018	[0.0002 - 0.01]	0.0157
Myeloid dendritic cell activated	-0.8148	0.4427	[0.25 - 0.79]	0.0338
Myeloid dendritic cell	-2.6518	0.0705	[0.01 - 0.49]	0.0372
B cell	-1.0171	0.3616	[0.17 - 0.77]	0.0401
Monocyte	-3.3809	0.0340	[0.002 - 0.47]	0.0444

Summary chapter 5

Sex-specific differences in gene expression and immune cell composition were investigated in stage I,II and stage III CM, normal skin, and nevi samples from multiple datasets (LMC, TCGA, GSE53118, GTEx and Nevi) using a combination of statistical data processing languages based on different computational methods: Kaplan-Meier survival analysis, principal component analysis (PCA), GEV assessment and differential gene expression. Results of Cox regression analysis demonstrated a notably worse OS for males in stage I-II CM. Females and males were hardly separated by PCA, primarily because principal components explained only a small fraction of the variance in all datasets. We observed significant differences in GEV with females having a lower GEV than males for both the early-stage melanomas and the normal skin, while males had a more substantial median difference in late-stage melanomas. Bootstrap analysis further supported the GEV findings. Differential expression analysis revealed a decreasing tendency in the number of differentially expressed genes from normal skin to nevi and from early-stage to late-stage melanomas. Next, using xCell for immune deconvolution, we discovered sex-based differences in the enrichment of various types of immune cells during early-stage (I-II) and advanced stage (III) melanomas. On the multivariable Cox regression analysis in female stage I-II melanomas, several immune cell subtypes were significantly associated with survival outcomes. Our results emphasize the importance of considering sex as a biological variable in cutaneous melanoma (CM) studies. This consideration could help elucidate how sex influences the biology of skin melanoma and may also indicate potential differences in response to immune checkpoint inhibitors based on sex.

Chapter 6: Discovery of sex-specific prognostic gene biomarkers in stage I-II CM.

The main objective of this chapter is to identify sex-dependent survival gene biomarkers in CM. Given the previous findings on sex-related differences in survival outcome and gene expression (variability), we aim at discovering female-specific and male-specific survival biomarkers. This analysis was performed in female-only and male-only sample sets adjusting for age and stage. Before that, we included the entire set of stage I,II patients and include the interaction term between gene expression and sex as covariate. Advanced statistical methods, including LASSO penalized and bi-directional stepwise Cox regression analysis, were used to build the sex-specific prognostic models.

3.1 Filtering out genes utilising multivariable Cox regression analysis

In order to filter-out any non-survival related genes, we performed a multivariable Cox regression analysis on each gene adjusting for age, stage and sex. Genes with Benjamini-Hochberg (BH) adjusted p-value less than 0.05 were considered for further analyses. First, we considered stage I, II and III combined groups of females and males, and then we analysed: females only, males only; stage I-II only and stage III only (Table 12).

Table 12: Multivariable Cox Regression Analysis on the LMC Melanoma Cohort. This table presents the results of a multivariable Cox regression analysis on the LMC melanoma cohort. The analysis was performed on combined groups of females and males, as well as separately for females and males, stratifying by stage and age. The "Adjusted for" column indicates the covariates each multivariable Cox model was adjusted for. The table shows the number of genes with an adjusted p-value < 0.05 for each combination.

		LMC	
Groups	**Adjusted for**	**Stage**	**Adjusted. p-value < 0.05**
Females + Males	Age, Stage and Sex	I,II,III	1603
Females	Age and Stage		1475
Males	Age and Stage		2
Females + Males	Age, Stage and Sex	I-II	351
Females	Age and Stage		1119
Males	Age and Stage		0
Females + Males	Age and Sex	III	20
Females	Age		0
Males	Age		0

The number of statistically significant genes obtained from each sex individually displayed a substantial dissimilarity. Female genes were more abundant (Table 12) and show an overall higher statistical significance, both in combined stage I-II and in stage I-II-III compared to male ones (Figure 11B). The same Cox regression analysis was then performed on TCGA and GSE53118 datasets, but no survival-related genes satisfying the adjusted p-value cut-off were obtained.

Analyses were repeated by randomly selecting, among females, subsamples with the same male sample size (256 samples), and results did not change. Ten subgroups containing 256 female samples were created, each including 57 events. The average number of adjusted p-value significant genes was 839.5, ranging from 320 to 1653. (Figure 10) .

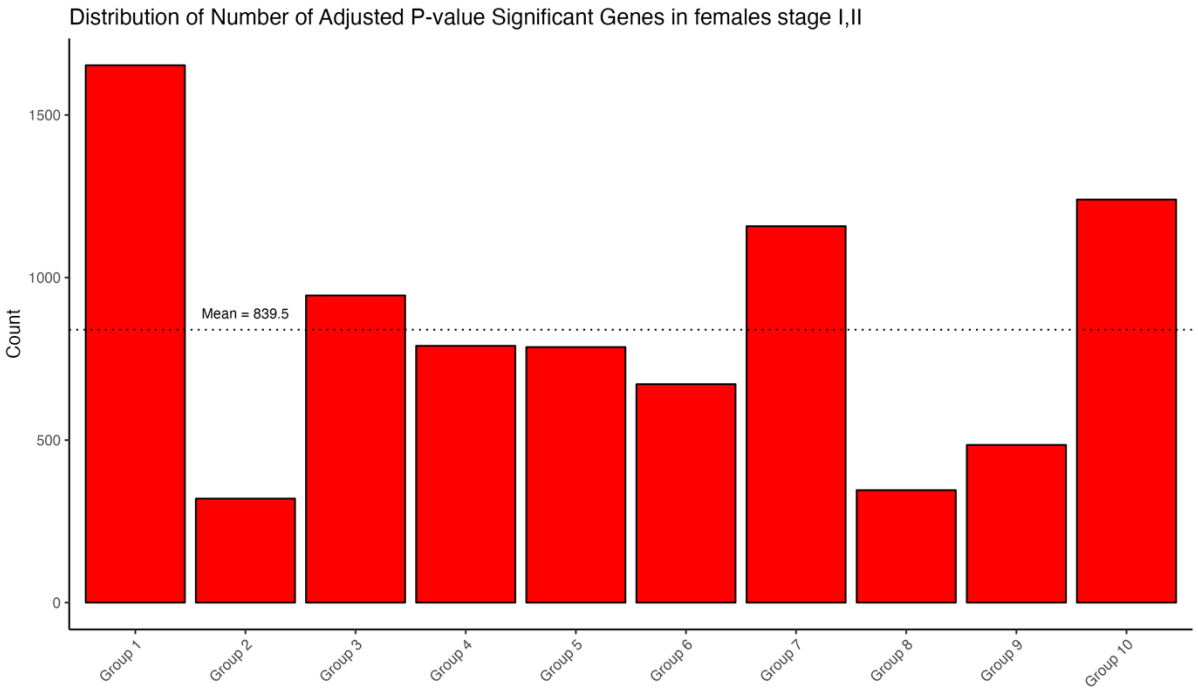


Figure 10: Distribution of genes with significant p-values in adjusted models ($p < 0.05$) in females with stage I-II cancer. The female cohort, consisting of 311 samples, was divided into 10 groups of 256 samples each, with the same proportion of event as the total cohort. The bar plot shows the count of significant genes (adjusted p-value < 0.05) in each group. The dashed horizontal line represents the mean count of significant genes across all groups.

Concerning stage III, there were no sex-specific statistically significant genes, but there were 20 when females and males were combined. Figure 11A presents the 351 genes obtained from the multivariable Cox regression analysis on stage I-II genes adjusted for age, stage and sex intersected with the 1119 genes obtained from the female only analysis, showing that 254 genes are still significant in the female population. In addition, the 351 genes' beta coefficients (the natural logarithm of the hazard ratios) and adjusted p-values extracted from the female-only and male-only analysis (Figure 11B), show that the significance only holds true for the female population. No gene for males was indeed found above the dashed horizontal line that corresponds to adjusted p-value = 0.05.

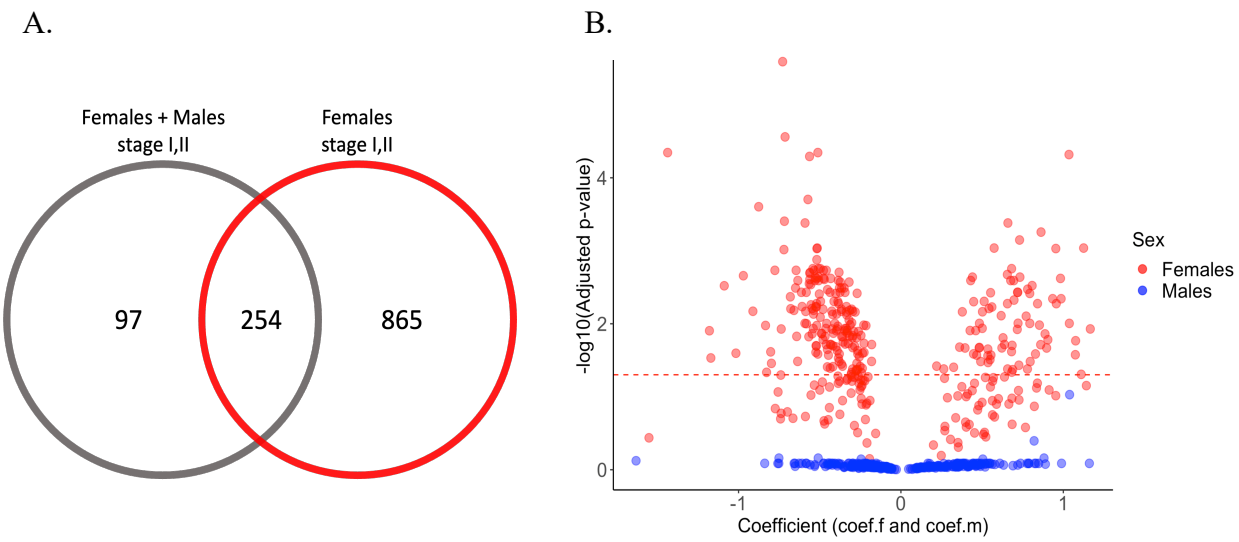


Figure 11: Multivariable Cox regression analysis results for the genes in stage I-II melanomas. A) Venn diagram representing the common and distinct statistically significant genes obtained for females + males stage I-II or females only stage I-II. B) Volcano plots representing multivariable Cox regression coefficients (x axis) and -log10 of the adjusted p-values (y axis) extracted from the female only (red points) and male only (blue points) analyses, for the 351 statistically significant genes in females + males. The horizontal dashed line corresponds to adjusted p-value = 0.05.

Table 13: Multivariable Cox Regression Analysis of Age, Stage, BRAF, and NRAS in LMC Stage I-II Melanomas. Table presents the results of a multivariable Cox regression analysis evaluating the impact of age, stage, BRAF, and NRAS mutations on survival in LMC stage I-II melanomas, separately for females (top) and males (bottom). The analysis adjusts for age and stage in both groups.

LMC Females stage I,II adjusted for age and stage

	beta coefficient	HR [95% CI]	p-value
Age	0.02	1.03 [1.01 - 1.04]	9.11E-03
Stage	1.15	3.14 [1.77 - 5.57]	8.68E-05
BRAF	0.15	1.17 [0.73 - 1.87]	0.5
NRAS	-0.14	0.87 [0.49 - 1.55]	0.63

LMC Males stage I,II adjusted for age and stage

	beta coefficient	HR [95% CI]	p-value
Age	0.04	0.96 [1.02 - 1.07]	3.69E-04
Stage	0.78	2.17 [1.28 - 3.70]	4.05E-03
BRAF	0.62	1.85 [1.18 - 2.92]	0.0008
NRAS	-0.24	0.79 [0.44 – 1.39]	0.41

Two separate multivariable Cox regression models were performed on each sex 1; including age, stage and BRAF and 2; including age, stage and NRAS (Table 13). Age and stage were both associated with worst survival in both sexes in stage I-II. These results revealed no significant association with overall survival in stage I-II female melanomas either for BRAF or NRAS. However, while in male melanomas NRAS was not significantly associated with survival, BRAF was significantly associated with poor overall survival (p-value = 0.008).

3.2 Investigation of the gene*sex interaction in a multivariable Cox regression analysis

With the aim of further investigating if the effect of gene expression on survival outcomes differs between females and males, the gene*sex interaction term added in the multivariate Cox regression survival analysis already adjusted for age and stage (Figure 12). Out of the 29,000 genes in the LMC dataset there were no adjusted p-value (<0.05) significant genes. However, when the cut-off was relaxed to a raw p-value < 0.001, there were 63 significant genes associated with survival in a sex-specific manner (Figure 13).

To investigate these 63 genes in a more sex-specific way, a multivariable Cox regression analysis was performed on females only and males only. While in females, all 63 genes were raw p-value significant with strong coefficients, in males only 22 genes were raw p-value significant and had lower in general coefficients compared to the female analysis.

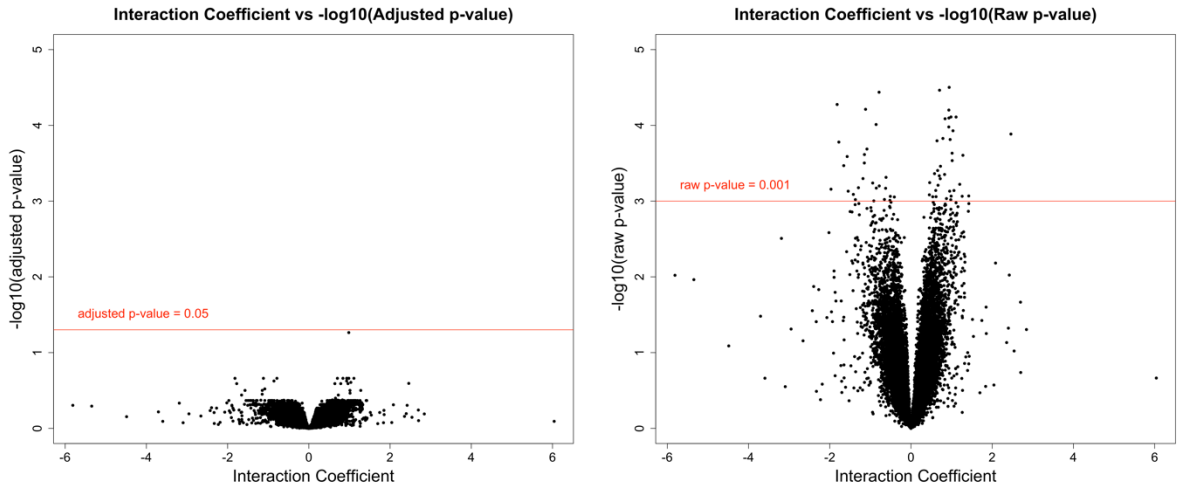


Figure 12 : Interaction Term Analysis of Gene Impact on Survival in Stage I-II CM. The figure shows the results of a multivariable Cox regression analysis in terms of the interaction gene*sex. The left panel displays a volcano plot of interaction coefficients vs. -log10(adjusted p-value) with no significant genes below the 0.05 threshold (red line), the right panel shows interaction coefficients vs. -log10(raw p-value), revealing 63 significant genes with sex-specific associations at the 0.001 threshold (red line).

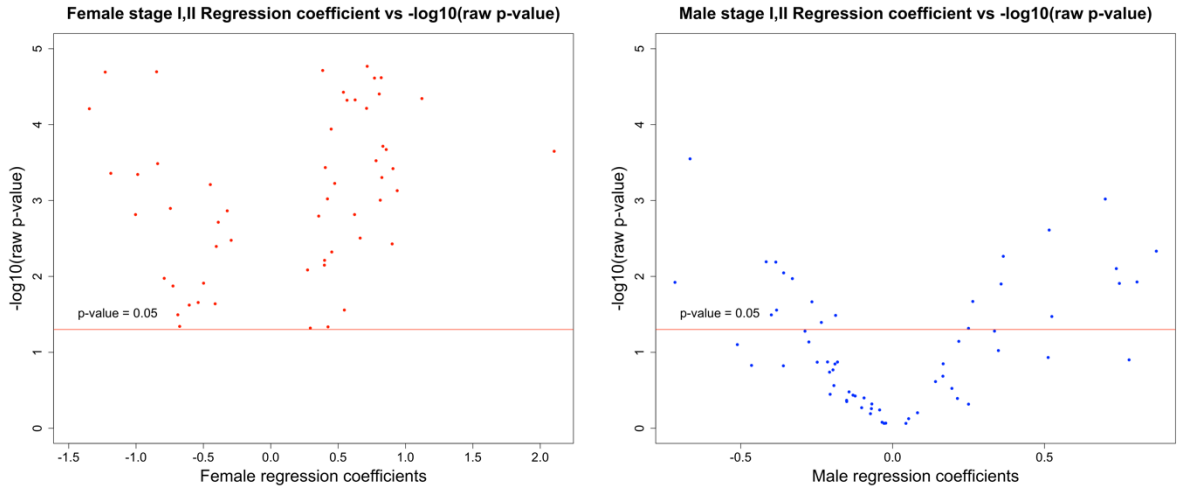


Figure 13: Regression Coefficients vs -log10(raw p-value) in Female and Male separately in stage I, II CM. This figure illustrates the regression coefficients plotted against -log10(raw p-value) for female and male stage I, II melanoma patients. The left panel displays the female regression coefficients in red, while the right panel shows the male regression coefficients in blue. Both plots include a red horizontal line representing the raw p-value threshold of 0.05.

3.3 Discovery of female-specific prognostic gene signature

In order to identify female-specific survival biomarkers, we first applied multivariable Cox regression analysis. As was previously shown in Table 12, by adjusting for age and stage we

found 1119 genes associated with female-specific survival. To build a prognostic model while avoiding overfitting, we used a 10-fold cross validation approach, while censoring at 8-years, by partitioning our sample group into 10 smaller groups of approximately 31 samples, maintaining in each subgroup the same proportion of events as in the total sample. During each iteration, we trained a LASSO penalised Cox regression model (using cv.glmnet) on 9/10 of the subgroups and evaluated its performance on the remaining one. A cross-validation LASSO Cox regression selects survival-influencing covariates, while penalising non important survival covariates with a zero coefficient. Age and stage were always included as input variables.

Our focus was on genes that consistently showed non-zero coefficients across all ten iterations. In each cycle of this internal validation process, we assessed performance metrics such as the UNO AUC, UNO C-index and observed/expected ratio. We also calculated the baseline survival rate, which indicates the chances of survival when all covariates are set to zero. Since certain factors, at zero, could be unrealistic, we standardised by scaling our data. The average baseline survival at 8-years was 81%, which makes sense as early detection of melanoma is associated with long term survival outcomes. An UNO AUC average value of 0.75 was obtained, which translates to the models having a moderate to strong predictive performance. However, the confidence interval spanning from 0.51 to 0.99 introduces uncertainty. For the UNO C-index, defined as the concordance probability estimate, we obtained an average value of 0.72 that proposes the model has a reasonably good predictive accuracy in ordering the samples paired according to their survival time. We obtained a mean ratio of observed/expected of 0.97, which suggests that the observed number of events is very close to the number of events predicted by the model (the closer to 1 the better the calibration of the model). However, the wide confidence interval (0.43 - 2.17) indicates considerable uncertainty (Table 14).

Table 14: Average performance metrics of 10-fold cross validation in female stage I-II melanomas. The table presents the average baseline survival rate, Uno AUC, UNO C-index, and observed/expected ratio, with respective confidence intervals, obtained from a 10-fold cross-validation of the LMC female stage I-II melanoma dataset. obs/exp = observed/ expected

Average (10-folds)	LMC - Females stage I-II
Baseline survival	0.8085
Uno AUC	0.7531 [0.5151 - 0.9909]
UNO C-index	0.7211
obs/exp ratio	0.9685 [0.4317 - 2.1741]

The genes with a non-zero coefficient all 10 times were: HLA class I histocompatibility antigen, alpha chain E (*HLA-E*), Ubiquitin like modifier activating enzyme 7 (*UBA7*) and Ubiquitin Like With PHD And Ring Finger Domains 1 (*UHRF1*). The three genes were then incorporated into a comprehensive multivariable Cox regression model (Figure 14) using the entire group consisting of 311 females. Neither age nor stage were selected by the ten models. The beta coefficients of *HLA-E* and *UBA7* were -0.38 and -0.29 respectively, which suggested a protective role (the higher their expression level the longer the patient survival) for these genes. Conversely, *UHRF1* had a beta coefficient of 0.63, indicating an association with worse prognosis.

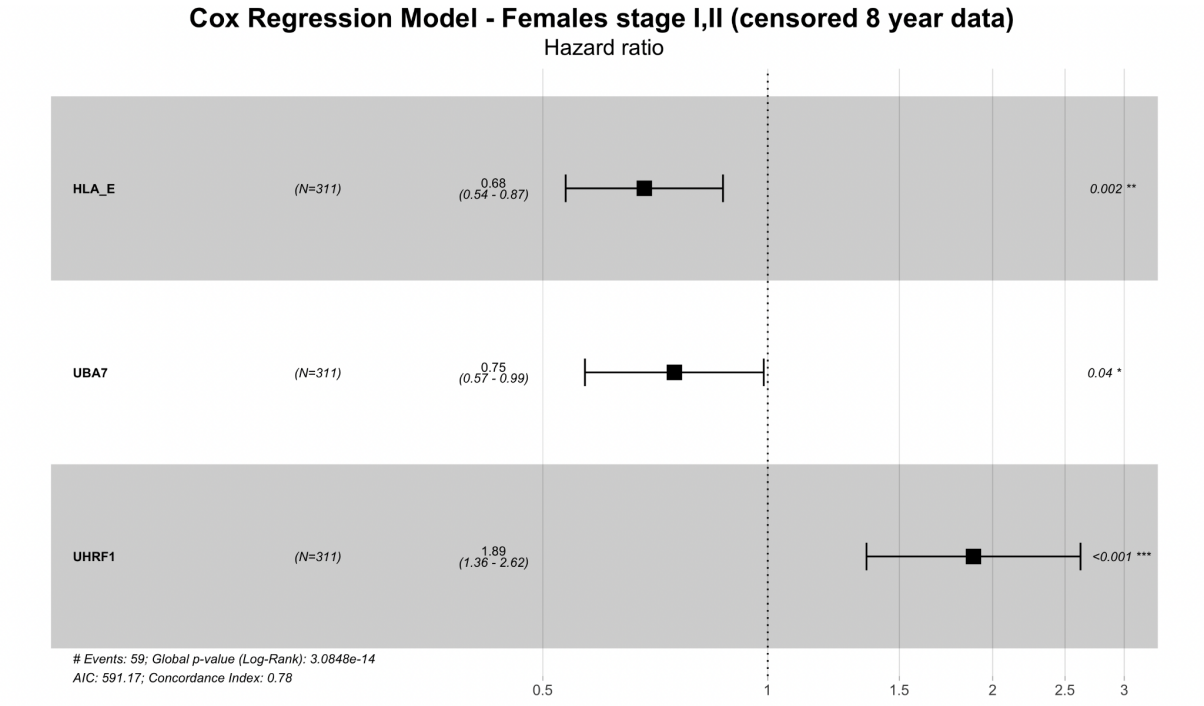


Figure 14: Multivariable Cox regression Hazard ratios and related confidence intervals of the three female-specific stage I-II survival covariates on the LMC cohort. The figure shows the hazard ratios and 95% confidence intervals for three female-specific survival genes covariates (HLA-E, UBA7, UHRF1) in stage I-II melanomas from the LMC cohort, with data censored at 8 years. The dotted line at 1 represents no effect. Significant p-values are indicated next to each covariate. The analysis includes 59 events, with a global p-value of 3.08e-14, AIC of 91.17, and a Concordance Index of 0.78.

3.4 Discovery of male-specific prognostic gene signature

Because of the absence of genes strongly associated with survival, we adapted a slightly different methodology to build a male-specific prognostic model. We set at 0.001 the raw p-value cut-off of the multivariable Cox regression analysis described in chapter 5 (Appendix

Table 1). Out of these 29 genes, we selected five with the higher Hazard ratios. The five chosen genes were Brain Expressed X-Linked 3 (*BEX3),* Integrin Subunit Alpha 4 *(ITGA4),* MYC Associated Zinc Finger Protein *(MAZ),* Quiescin Sulfhydryl Oxidase 1 *(QSOX1),* Splicing factor 3B subunit 3 *(SF3B3)*. All five genes had positive beta coefficients (1.04, 0.82, 1.25, 0.93, 0.88) which translates to an anti-survival role of these genes in our male population. In order to create, out of our 5 genes, age and stage, the most promising multivariable male-specific model, we used a 10-fold cross validation approach describe as done for females and within each iteration we performed a bi-directional stepwise analysis. This analysis started with a full model including all variables and systematically added or removed variables to improve the model fit, aiming to minimise the Akaike Information Criterion (AIC) value. The AIC is calculated based on the log-likelihood function and the number of parameters in the model: the lower the AIC the better the balance between the model's goodness of fit and its complexity.

Age, Stage, *BEX3* and *SF3B3* were selected 10 times. In each cycle, we assessed the same average performance metrics as for females: baseline survival, UNO AUC, UNO C-index and observed/expected ratio. The average baseline survival in males with stage I-II melanoma at year 8 was 0.71, signifying a 71% of survival at 8 years. As expected, this percentage is lower than that obtained for females. An average UNO AUC of 0.72 was obtained, indicating a moderate discriminatory ability of the model in distinguishing between individuals in ranking them by their survival time. Average UNO C-index was 0.69, suggesting a fair predictive accuracy in assigning higher predicted risks to those who experience events earlier in time. The average ratio of observed/expected was close to 1, indicating a reasonable fit of the model in aligning expected to observed events (Table 15).

Table 15: Average performance metrics of 10-fold cross validation in male stage I-II melanomas. The table presents the average baseline survival rate, Uno AUC, UNO C-index, and observed/expected ratio, with respective confidence intervals, obtained from a 10-fold cross-validation of the LMC male stage I-II melanoma dataset. obs/exp = observed/ expected

Average (10-folds)	**LMC - Males stage I-II**
Baseline survival	0.7128882
Uno AUC	0.7239 [0.4768 - 0.9711]
UNO C-index	0.6961
obs/exp ratio	1.0077 [0.4827 - 2.1043]

The four selected covariates (age, stage, BEX3 and SF3B3) were then used to create a multivariable Cox regression model on the whole population of males stage I-II (Figure 15).

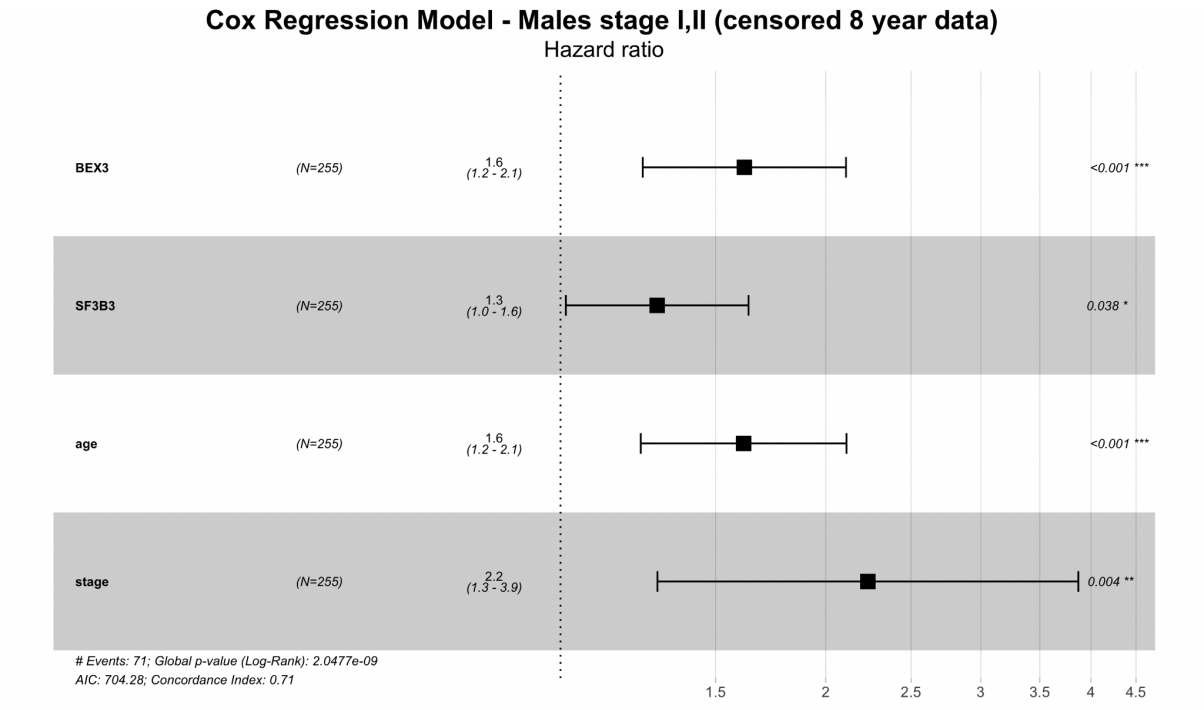


Figure 15: Multivariable Cox regression on LMC Hazard ratios of the four male-specific stage I-II survival covariates. The figure presents the hazard ratios and 95% confidence intervals for four male-specific survival covariates (BEX3, SF3B3, age, and stage) in stage I-II melanomas from the LMC cohort, with data censored at 8 years. The dotted line at 1 represents no effect. Significant p-values are indicated next to each covariate. The analysis includes 71 events, with a global p-value of 2.04e-09, AIC of 78.28, and a Concordance Index of 0.71.

3.5 Individual assessment of genes in the discovered sex-specific signatures

We reported individual survival performance through multivariable Cox regression metrics on each gene contained in sex-specific prognostic models for females and males and plotted Kaplan-Meier (KM) survival curves. In female stage I,II melanomas UHRF1 had a positive beta coefficient, indicating that it is associated with worse survival, while *HLA-E* and *UBA7* had negative coefficients, meaning that they are associated with better survival (Table 16). On the other hand, the two male specific genes (*BEX3* and *SF3B3*) showed much lower absolute coefficients and much higher p-values, with the raw p-value not even reaching 0.1.

Table 16: Individual gene evaluation in multivariable Cox regression (adjusted for age and stage) for females with stage I-II melanomas, LMC cohort. Evaluation of individual gene performance in a multivariable Cox regression model for females with stage I-II melanomas from the LMC cohort, adjusted for age and stage. Metrics include beta coefficient, raw p-value, and adjusted p-value.

	LMC Females stage I-II - adjusted for age and stage		
Genes	beta coefficient	raw p-value	adjusted p-value
UHRF1	**1.03498**	**9.82E-09**	**4.80E-05**
HLA-E	**-1.43685**	**5.99E-09**	**4.51E-05**
UBA7	**-0.87191**	**7.68E-09**	**4.51E-05**
BEX3	0.24846	0.28118	0.63987
SF3B3	0.35366	0.15008	0.48994

The KM survival curves of the five genes were plotted in females, stratifying patients according to the median expression of the genes, revealing a significant difference for *UHRF1* (p = 0.00038), *HLA-E* (p<0.0001) and *UBA7* (p<0.0001) and no significant differences for *BEX3* (p = 0.51) and SF3B3 (p = 0.42) (Figure 16). High expression of *HLA-E* and *UBA7* seem to be significantly associated with a longer survival. However, high expression of *UHRF1* is negatively implicated with the survival of females with stage I-II melanomas.

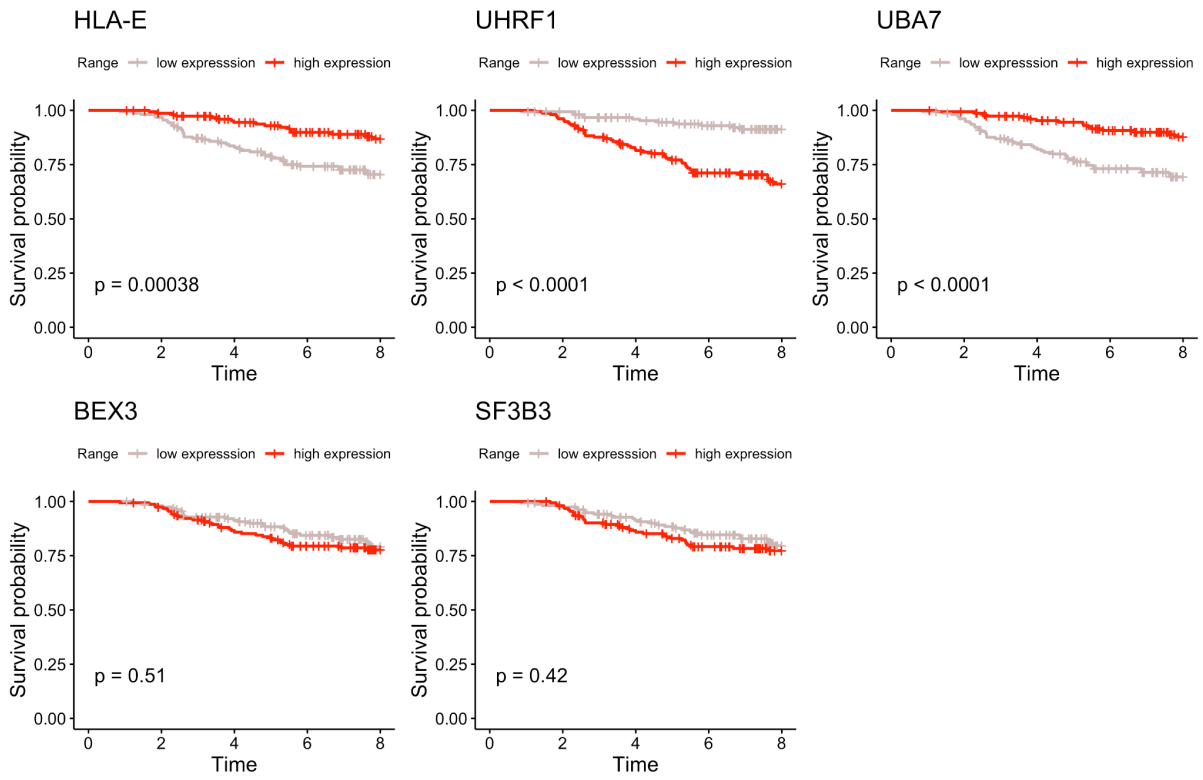


Figure 16: Kaplan-Meir curves of the five genes of interest in LMC females with stage I-II melanomas. Data is censored at year 8, and curves are created according to gene expression values (pink curve = below the median; red curve = above the median).

The five genes of interest were also assessed in the male cohort of stage I-II samples. Interestingly, *HLA-E*, a gene identified to be protective in females, had a close to significant raw p-value in males with stage I-II melanomas, suggesting a potential protective role of this gene in males as well. The beta coefficients of *BEX3* and *SF3B3* were both positive, associating these genes to negative survival, and their absolute values were by far higher than those of the three female selected genes (Table 17).

Table 17: Individual gene evaluation in multivariable Cox regression (adjusted for age and stage) for males with stage I-II melanomas, LMC cohort. Evaluation of individual gene performance in a multivariable Cox regression model for males with stage I-II melanomas from the LMC cohort, adjusted for age and stage. Metrics include beta coefficient, raw p-value, and adjusted p-value.

	LMC -Males stage I-II - adjusted for age and stage		
Genes	beta coefficient	raw p-value	adjusted p-value
UHRF1	0.14116	0.24298	0.92552
HLA-E	-0.53101	0.05317	0.89507
UBA7	0.08186	0.62563	0.97606
BEX3	**1.03876**	**3.19E-06**	**0.09357**
SF3B3	**0.88132**	**1.13E-04**	**0.69318**

The survival KM curves of the five genes of interest were also plotted in males with stage I-II melanomas, dividing patients according to the median expression value of each gene. The curves from the female-specific survival biomarkers *UHRF1*, *HLA-E* and *UBA7* were not separated at all. On the other hand, the high expression of *BEX3* was significantly associated with poor survival (log-rank test p-value = 0.015), whereas *SF3B3* expression was less able to separate patients into distinct prognostic groups (log-rank test p-value = 0.089). Only after year 5.5 the high expression of *SF3B3* seems to be associated with worse survival (Figure 17). However, this is not surprising as these genes were selected together during bidirectional stepwise analysis and combined with age and stage in the male-specific prognostic model.

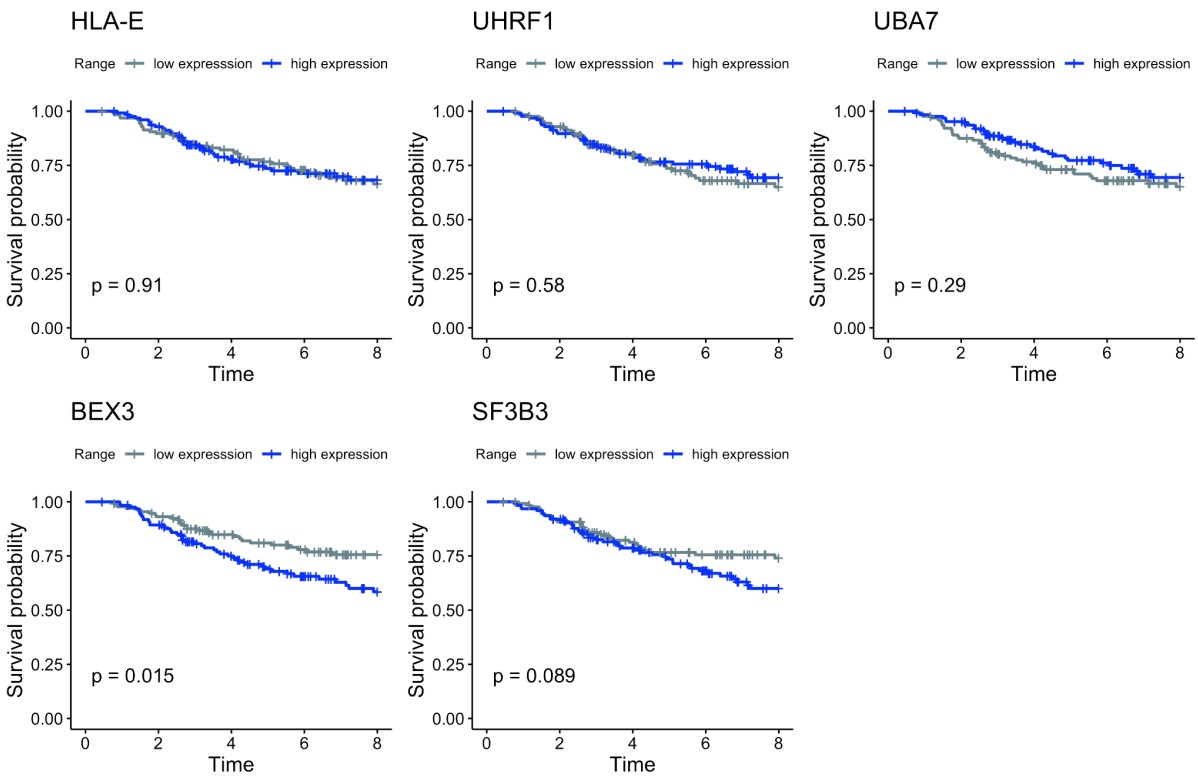


Figure 17: Kaplan-Meir curves of the five genes of interest in LMC males with stage I-II melanomas. Data is censored at year 8 and curves are created according to gene expression values (grey curve = below the median; blue curve = above the median).

The "Appendix Table 2" contains the TCGA multivariable Cox regression coefficients and raw p-values of females and males separately in stage I, II. None of the genes exhibit a significant p-value in either females or males. However, in females with stage I-II melanomas, most beta coefficients, except for *UBA7*, align in directionality with those observed in LMC. Meaning that they share the same sign, either positive or negative like the LMC obtained results.

In the analysis of TCGA males with stage I-II melanomas, the coefficients for the two male-specific genes exhibited coefficients in opposite directions (positive in LMC, negative in TCGA). As for the female-specific genes, *UHRF1* and *HLA-E* demonstrated coefficients in the same direction, but *UBA7* did not follow the same pattern.

The observed discrepancies might be due to the small sample dataset, different technology for expression profiling (microarrays in LMC vs RNA-seqs in TCGA) and contribute to the miscalibration observed for the models (especially the male one) when applied to the TCGA external dataset.

Summary chapter 6

Genes not significantly associated with sex-specific survival were filtered out and we found that in stage I,II female melanomas there is a much higher number of genes related with survival compared to male ones. The role of BRAF and NRAS mutation was evaluated in a sex specific manner, showing that BRAF mutation status is associated with survival only in males with stage I,II CM. The gene*sex interaction term was investigated using Cox regression, but no genes with statistically significant association, in terms of adjusted p-value, resulted from this analysis. Finally, for the selection of the sex-specific survival genes, LASSO penalised Cox regression and bi-directional stepwise analysis were utilised for females and males, respectively, with more relaxed cut-off values to choose the input genes in males. The female-specific model identified both protective (HLA-E, UBA7) and adverse (UHRF1) key genes and was validated internally (AIC = 91.17, Concordance Index 0.78). The male-specific model included two genes (BEX3, SF3B3), age and stage, and had good performance upon internal validation, as indicated by an AIC of 78.28 and a Concordance Index C of 0.71. Additionally, the discovered genes were assessed individually in each sex, showing once more a greater performance, in terms of both coefficients and p-values, in females compared to males stage I,II melanomas. Female-specific genes performed poorly in males, while male-specific genes demonstrated weak association with survival in females.

Chapter 7: Assessment of the identified sex-specific prognostic models on external cohorts.

This chapter focuses on validating the newly discovered sex-specific gene signatures of stage I,II CM. The aim of the analysis is to demonstrate that the gene signatures are robust and generalizable across external cohorts such as the TCGA and our own Biella cohort. Firstly, we assessed the overall more performant female-specific gene signature, obtained from LASSO Cox regression analysis, verifying its predictive power using metrics such as the Uno AUC, C-index and observed/expected ratio. These metrics provide understanding as to the discriminatory power, calibration and overall predictive performance. Similar validation steps were followed for the male-specific gene signature.

4.1 Validation of the female-specific gene signature

Is important to note that we did not compute average coefficients from the LASSO Cox regression because they encompass contributions not only from these three specific genes of interest, but also from all non-zero coefficients obtained throughout every iteration.

Table 18: Validation metrics of the female-specific survival model applied to Biella and TCGA datasets. Validation metrics of the multivariable Cox model for female-specific survival covariates in Biella and TCGA datasets (stage I-II). Metrics include Uno's AUC, UNO C-index, and observed/expected ratio with confidence intervals.

	Biella cohort - Females stage I-II	**TCGA - Females stage I-II**
Uno AUC	0.8036 [0.5736 - 1]	0.7953 [0.5006 - 1]
UNO C-index	0.7638	0.7161
obs/exp ratio	0.3044 [0.1267 - 0.7313]	0.5843 [0.2922 - 1.1684]

The survival model was then applied to the Biella (Appendix table 5). and TCGA cohorts for females with stage I-II melanomas and showed promising performance (Table 18). Due to the unavailability of follow up information in TCGA, the survival model was validated on data censored at year 3, unlike the Biella cohort that had follow up info for up to 8 years. UNOs AUC estimates indicated a fair ability to discriminate between females with or without event, with values of 0.8036 for the Biella cohort and 0.7953 for TCGA, although some uncertainty due to confidence intervals. UNO C index suggested higher effective ranking of survival times

in the Biella cohort (0.7638) than in the TCGA cohort (0.7161). However, the observed/expected ratios suggest calibration issues while aligning predicted and observed event rates, with both datasets having a value lower than 1, meaning fewer events than expected. These findings demonstrate promising discrimination metrics but emphasise the need for further investigation ideally with cohorts including more samples and events. TCGA validation metrics are poorer than the Biella ones.

Overall, while the validations in both external cohorts showed decent discriminatory and performance ability, the wide confidence intervals across numerous metrics indicated a level of uncertainty. Further investigation is recommended, especially in larger cohorts.

4.2 Validation of male-specific gene signature

The validation of the multivariable Cox regression model in the two external cohorts was performed by evaluating the same performance metrics as in the internal validation: UNO AUC, UNOC-index, and obs/exp ratio (Table 19). Across these metrics, the Biella cohort (Appendix table 5). demonstrated notably stronger performance compared to the TCGA dataset. UNO AUC was equal to 0.85 in the Biella cohort, signifying a strong ability in distinguishing between individuals with events compared to those without an event, while in TCGA it was equal to 0.57, indicating a very weak discriminatory ability. Again, the ability to rank pairs of individuals based on the survival time was notably lower in TCGA (0.52) compared to the Biella cohort (0.86). Concerning the ratio of obs/exp, both external cohorts performed similarly, with a value lower than 1 signifying that the observed events were fewer than what the model anticipated, indicating potential overestimation of the model's predictions.

Table 19: Validation metrics of the male-specific survival model applied to Biella and TCGA datasets. Validation metrics of the multivariable Cox model for male-specific survival covariates in Biella and TCGA datasets (stage I-II). Metrics include Uno's AUC, UNO C-index, and observed/expected ratio with confidence intervals.

	Biella cohort - Males stage I-II	**TCGA - Males stage I-II**
Uno AUC	0.85 [0.6596 - 1]	0.5729 [0.1652 - 0.9805]
UNO C-index	0.8649	0.5246
obs/exp ratio	0.5397 [0.2026 - 1.4379]	0.5837 [0.2191 - 1.5551]

To conclude, the validation outcomes of the male-specific survival biomarkers performed poorer compared to the female ones. These results were somewhat anticipated, given that, overall, 1; the initial multivariable Cox regression results were statistically poorer in males and 2; the male population's higher GEV introduces unclear results in a number of analyses. The findings underscore poor to moderate performance strength and considerable uncertainty.

Summary chapter 7

This chapter's goal was to validate the newly discovered sex-specific gene signatures in external datasets, TCGA and Biella. Although results are promising especially for females, it is important to mention that the sample size of the validation datasets has a limited number of samples. The 3-gene female signature prognostic values was assessed using metrics such as the UNO AUC, C-index and observed/expected ratio. The female-specific validation performed well, especially in the Biella dataset. In contrast, validation of the male-specific gene signature showed substantial variation in performance between the two external datasets, with stronger validation metrics in the Biella one. These results were to some extent expected given the initial multivariable Cox regression outcomes and the GEV results.

This chapter highlights the strengths and limitations of the identified sex-specific biomarkers, in terms of robustness and generalizability to external dataset. To improve generalization and predicted power of these biomarkers, further investigation with larger cohorts is needed.

Chapter 8: Functional analysis of the discovered sex-specific gene biomarkers

This chapter aims to expand and further understand the functional implications of the discovered female and male specific gene biomarkers. Initially, through correlation analysis, the genes positively or negatively correlated with our genes of interest were assessed. This analysis was performed separately in females and males, and the top genes are plotted followed by a functional gene ontology analysis. The enriched biological processes facilitate the interpretation and understanding of the functions related to the genes of interest. Additionally, a correlation analyses between the gene expression values of the discovered biomarkers and the xCell enrichment scores were performed to elucidate any immune cell-related functions. Furthermore, network analysis identified the most important genes related to our genes of interest outside the context of CM. Subsequently, extraction of single cell data and KM plots based on the TCGA from the Protein Atlas highlighted: 1) the cells expressing higher levels of the genes of interest and 2) the ability of the selected genes to stratify TCGA CM patients in groups with significantly different overall survival.

5.1 Assessment of correlation of the genes of interest with all other genes

We calculated Spearman correlation and adjusted the p-value based on the number of tests performed using the Benjamini-Hochberg (BH) method. This analysis was used to identify the most positively and negatively correlated genes with our genes of interest. In stage I-II female melanomas we correlated individually, *UHRF1*, *HLA-E* and *UBA7* with the other 29,352 genes in the cohort. Also, the two male genes, *BEX3* and *SF3B3*, were correlated with the remaining 29,353 genes in the male stage I-II cohort. The adjusted p-value threshold was set at 0.05 to select significantly positively and negatively correlated genes in each analysis. The top 200 positively and negatively correlated genes of each analysis were plotted (Figure 11) and uploaded in DAVID bioinformatic tool to assess the biological processes (BP5) for which they were significantly enriched. All top 200 genes had correlated test adjusted p-value lower than 0.05.

The plots showing Spearman correlation coefficients and corr.test p-values of the top 200 positively correlated genes and the top 200 negatively correlated genes for each of the five genes of interest (evaluated within sex-specific datasets) confirm, once more, an overall better

"performance" in the female-population, with all correlated genes having lower p-values and higher correlation coefficients compared to the genes correlated to male-specific genes (Figure 18). This one more reflects the lower gene expression variability observed in females.

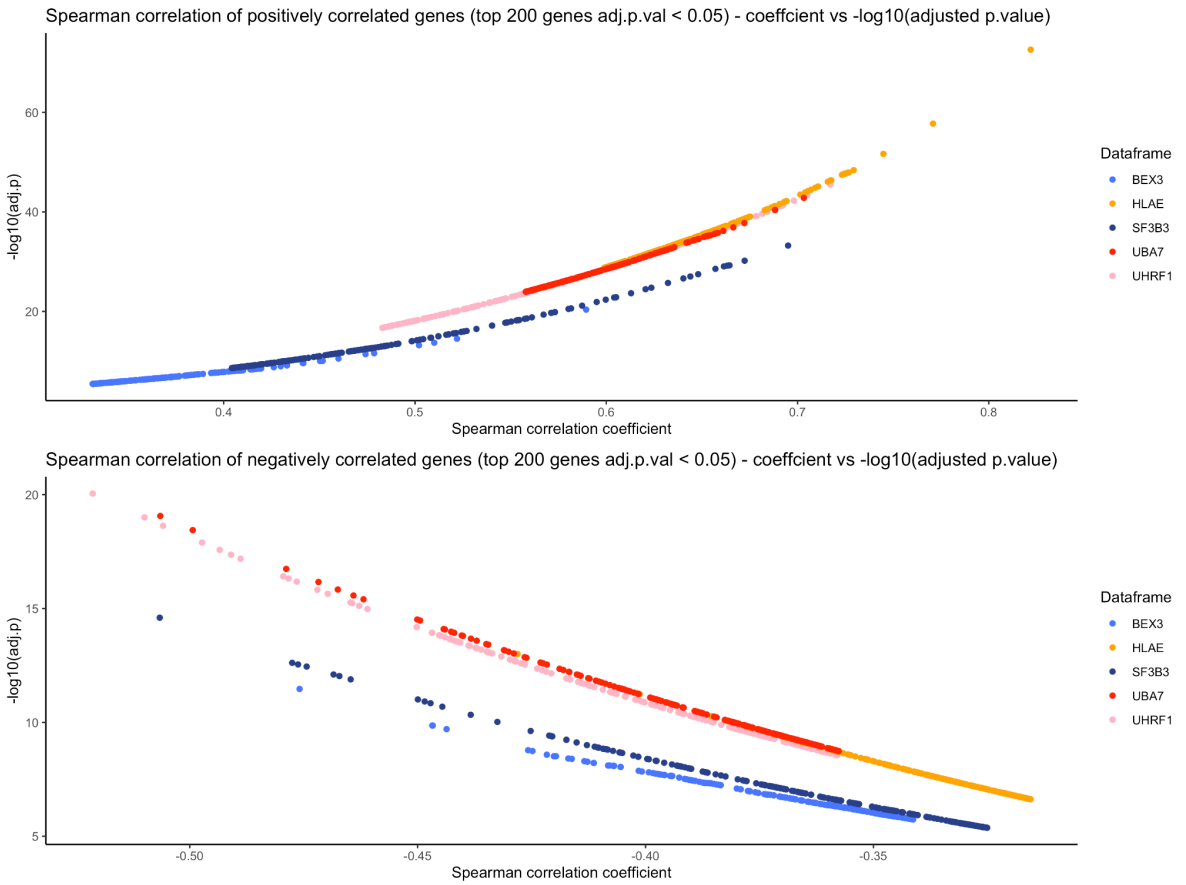


Figure 18: Spearman correlation coefficients plotted against the -log10 adjusted p-value of the top 200 positively correlated and the top 200 negatively correlated genes, each colour representing one gene of interest. Correlation with UHRF1, HLA-E and UBA7 was calculated in the LMC female stage I-II dataset, while correlation withBEX3 and SF3B3 was calculated in the LMC male stage I-II dataset.

We then applied the DAVID functional enrichment tool to retrieve the biological processes (level 5 of the gene ontology) for which the lists of (anti)correlated genes were enriched. The threshold for the Bonferroni adjusted enrichment p-value was set to 0.05 and the top biological processes significantly overrepresented within the list of genes positively and negatively correlated with the genes of interest were reported. If less than 10 processes were reported, it was due to the limited number of biological processes with significant enrichment in the analysis.

5.2 Female- specific genes

UHRF1

Ubiquitin-like with PHD and ring finger domains 1 (*UHRF1*) is a protein-coding gene that is involved in DNA methylation maintenance and epigenetic regulation (172). The genes that show a positive correlation with *UHRF1* are involved in processes like controlling the cell cycle during mitosis, dividing the nucleus, separating sister chromatids and transitioning between different phases of the cell cycle. On the other hand, the genes that have a negative correlation with *UHRF1* are significantly enriched only in the signalling pathway of cell surface receptors and cell migration. However, enrichment scores and p-values are significantly poorer compared to the positively correlated processes (Figure 19).

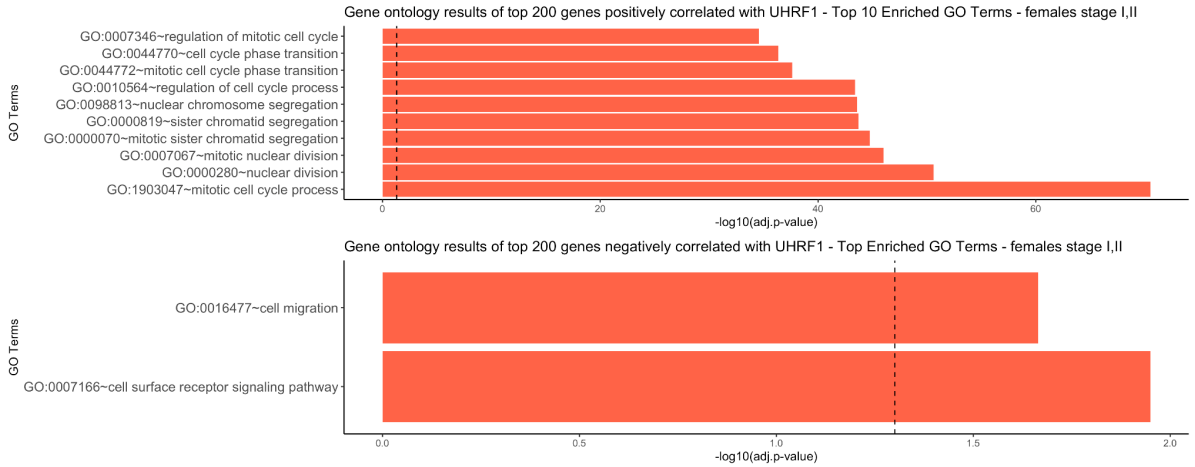


Figure 19: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for UHRF1. The BPs overrepresented within the top200 genes positively (top) and negatively (bottom) correlated with UHRF1 in LMC female stage I-II melanomas are represented. The histogram length reflects the statistical significance of the enrichment.

HLA-E

Major histocompatibility complex, class I, E (*HLA-E*) is a protein-coding gene that encodes a non-classical MHC class I molecule that presents peptides derived from the leader sequences of other MHC class I molecules. It is expressed in various tissues and cell types, including placenta, endothelial cells, and immune cells. It presents peptides to natural killer (NK) cells and certain T cells to regulate immune responses.

Genes that revealed positive correlation with *HLA-E* were found to be enriched in immune response regulation processes, including lymphocyte activation, regulation of cell adhesion

leukocytes T cell activation and proliferation of lymphocytes and mononuclear cells. Genes negatively correlated with *HLA-E* were found to be enriched in BPs associated with ribosome assembly and metabolism of non-coding RNA (ncRNA). These findings provide insights into how *HLA-E* plays a role in modulating a variety of immune related mechanisms (Figure 20).

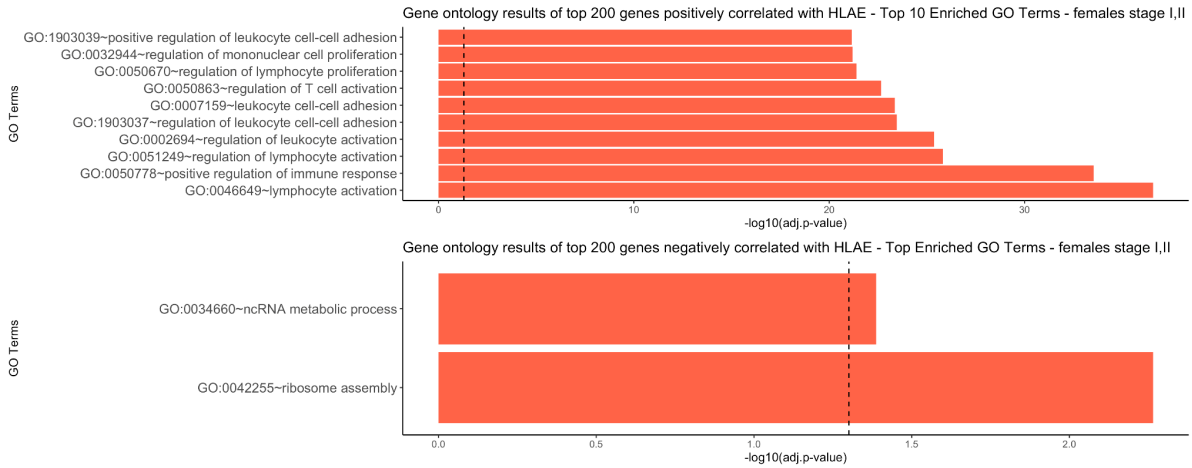


Figure 20: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for HLA-E. The BPs overrepresented within the top200 genes positively (top) and negatively (bottom) correlated with *HLA-E* in LMC female stage I-II melanomas are represented. The histogram length reflects the statistical significance of the enrichment.

UBA7

Ubiquitin-like modifier activating enzyme 7 (*UBA7*) is a protein-coding gene that encodes an enzyme that activates the ubiquitin-like protein ISG15. It is expressed in various tissues and cell types, especially in response to interferon stimulation. It is involved in innate immunity and antiviral defence (173). It has been linked to various diseases, such as viral infections, inflammatory disorders, and cancer (174,175).

Genes with a positive correlation with *UBA7* were found to be involved in regulatory processes related to immune response. These processes include activation of lymphocytes, proliferation of leukocytes and lymphocytes adhesion between leukocytes and regulation of T cell activation.

Conversely, genes exhibiting a negative correlation with *UBA7* were significantly associated with DNA metabolic processes, cell cycle processes such as division and segregation of chromosomes metabolic pathways involving nucleotides, biogenesis of ribosomal subunits, maintenance of telomeres and segregation of sister chromatids during mitosis (Figure 21).

The findings highlight the immune relationship of positively correlated biological processes, further supporting the anti-tumorigenic role of *UBA7* in females stage I-II. In addition, the negatively correlated BPs with *UBA7* support a cell cycle pro-tumorigenic role.

Figure 21: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for UBA7. The BPs overrepresented within the top200 genes positively (top) and negatively (bottom) correlated with *UBA7* in LMC female stage I-II melanomas are represented. The histogram length reflects the statistical significance of the enrichment.

5.3 Male-specific genes

BEX3

Unlike females-specific genes that had both positively and negatively significant BPs in every analysis, male-specific genes *BEX3* and *SF3B3* only had significant BPs associated with the top 200 negatively correlated genes.

Brain expressed X-linked 3 (*BEX3*). It is a protein-coding gene that encodes a member of the brain expressed X-linked (BEX) family of proteins. It is expressed in various tissues and cell types, especially in the brain and nervous system. It has been implicated in various cancers and neurological disorders (176,177).

BEX3 was identified in our analysis as a gene with high expression significantly associated with worst survival. Interestingly, the BPs significantly overrepresented within the list of negatively correlated genes are involved in immune processes. Some of these processes are lymphocyte activation, leukocyte differentiation and proliferation, haematopoiesis and positive

regulation of cytokine production (Figure 22). These results further support the pro-tumorigenic/ anti-survival role of BEX3 in male stage I-II melanomas.

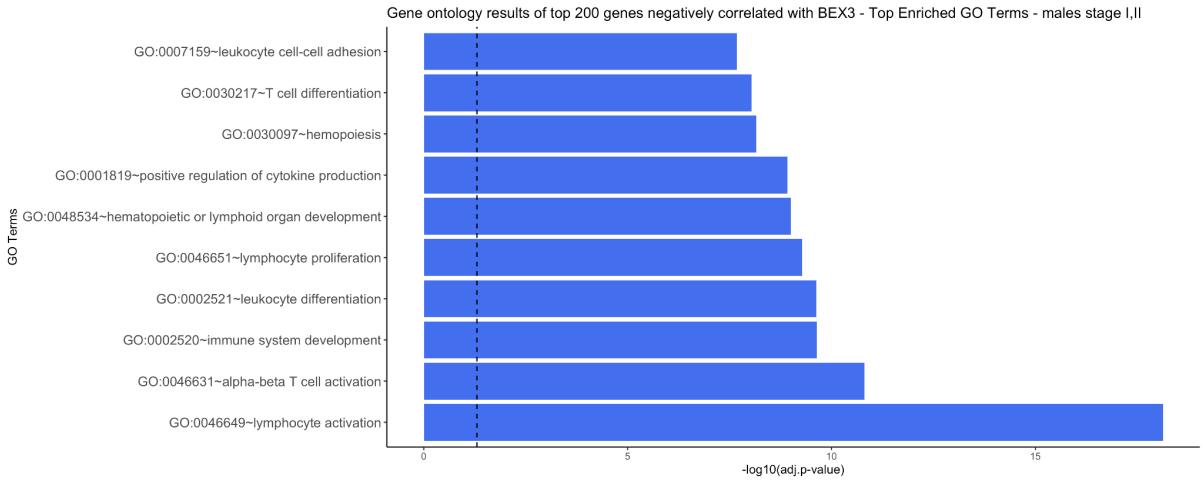


Figure 22: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for BEX3. The BPs overrepresented within the top200 genes negatively correlated with BEX3 in LMC male stage I-II melanomas are represented. The histogram length reflects the statistical significance of the enrichment.

SF3B3

Splicing factor 3b subunit 3 (*SF3B3*). It is a protein-coding gene that encodes a subunit of the splicing factor 3b (SF3B3) complex, which is involved in pre-mRNA splicing. Aberrant splicing has been associated with cancer development (178,179).

Like *BEX3*, *SF3B3* only had significantly overrepresented BPs for the negatively correlated genes, which resulted in associated with cellular mechanisms such as chromatin organisation, regulation of RNA splicing, protein, amide, and peptide transport. Additionally, these genes were involved in the regulation of mRNA metabolic processes (Figure 23).

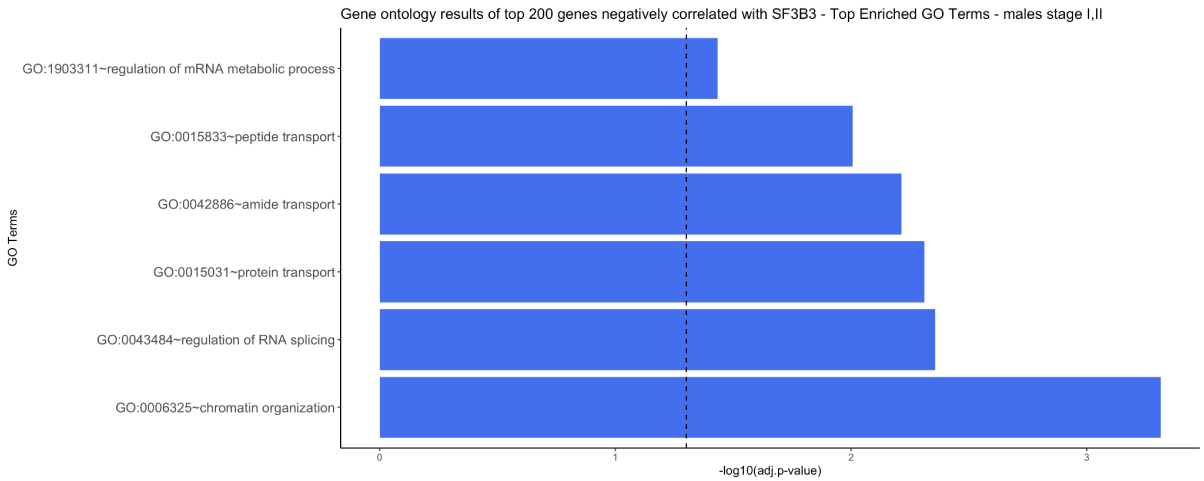


Figure 23: Gene ontology biological processes (level 5) obtained from DAVID functional enrichment tool for SF3B3. The BPs overrepresented within the top 200 genes negatively correlated with *SF3B3* in LMC male stage I-II melanomas are represented. The histogram length reflects the statistical significance of the enrichment.

5.4 Immune deconvolution using xCell

Considering our up to now results, especially the two immune related survival biomarkers identified in females, and wanting to further investigate the immune subtypes enrichment in each sample we performed an immune deconvolution analysis called xCell. This analysis, as previously reported, yields enrichment scores of 36 immune cell types.

To thoroughly assess the connection between our sex-specific genes and xCell's enrichment scores, we conducted Spearman correlation analyses. These assessments linked the enrichment scores obtained from xCell in female and male datasets with the expression values of sex-specific genes in the corresponding datasets. These analyses were carried out in a different context with respect to survival analysis where the genes were originally identified, and were performed separately for females (*UHRF1*, *HLA-E*, and *UBA7*) and males (*BEX3* and *SF3B3*).

The correlation results were plotted through a comprehensive heatmap (Figure 24), where columns refer to genes and rows to xCell immune cell-types. Each rectangle represents the correlation coefficient between enrichment scores and expression values.

Interestingly, *HLA-E* and *UBA7* that were identified by survival analysis in female stage I-II melanomas to be associated with better survival, are generally positively correlated with a number of immune cell subtypes. Oppositely, *UHRF1* that was identified by survival analysis as associated with worst survival in female stage I-II melanomas, was generally negatively

correlated with a number of immune cells. The only exceptions that were significantly positively correlated with UHRF1 were T cell CD4+ Th2 and T cell CD4+ Th1.

Regarding the two male-specific genes, *SF3B3* and *BEX3*, that were both identified as being linked with poor overall survival, their expression was generally negatively correlated with immune cell enrichment scores. *BEX3* had a number of significantly correlated immune cell types, unlike *SF3B3* that had no significant correlation but generally had negative Spearman correlation coefficients.

These results, to a certain extent, confirm our findings both in females and males, as in both sexes genes identified to be associated with poor survival are negatively correlated with immune cell subtypes and the two genes identified with favourable survival are positively correlated with immune cell subtypes.

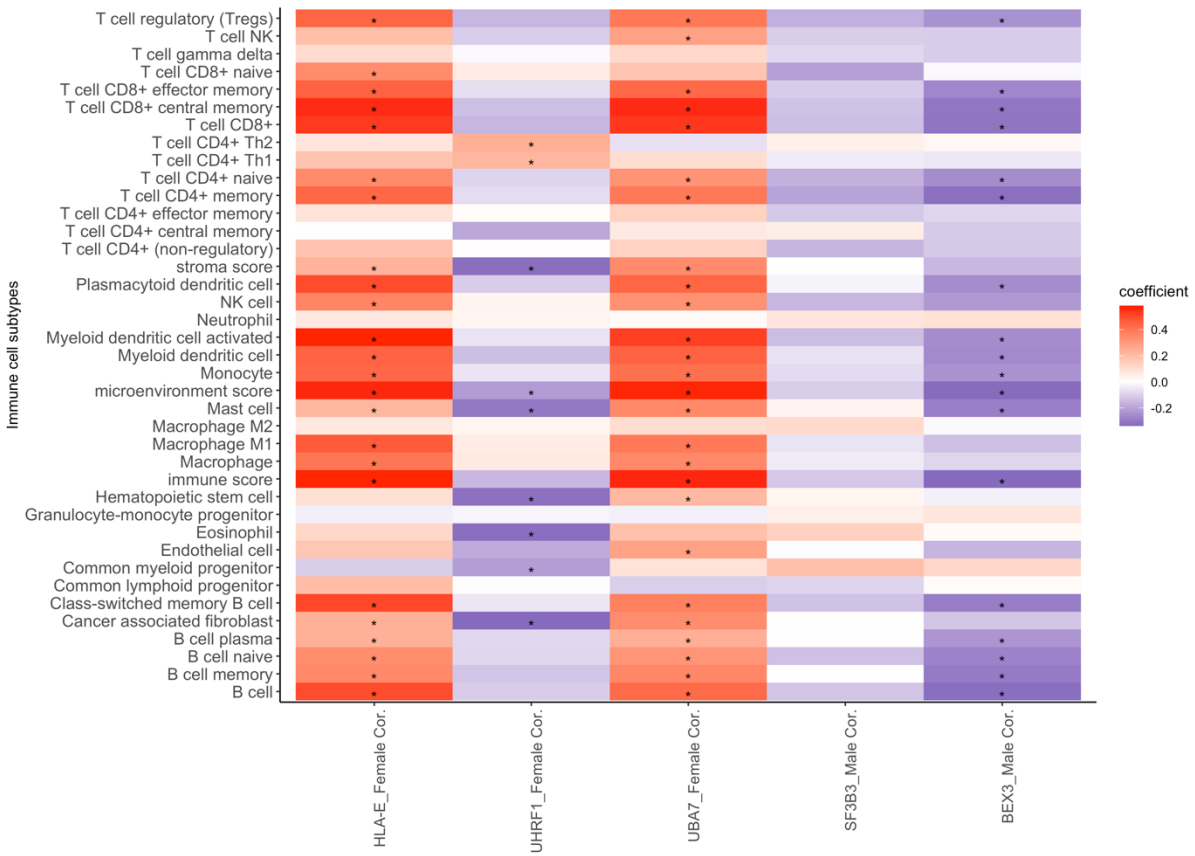


Figure 24: Correlation between gene expression of *HLA-E*, *UHRF1* and *UBA7* and xCell enrichment scores in female stage I-II melanomas and between gene expression of *SF3B3* and *BEX3* and xCell enrichment scores in male stage I-II melanomas. Red represents positive and blue negative correlation. Stars (*) represent correlations with adjusted p-value <0.05.

5.5 Network analysis

In this section we utilised the power of GeneMANIA, a bioinformatics tool known for its ability to predict gene functions and perform network analysis. GeneMANIA creates networks that show how different genes and/or proteins are related, helping us understand gene interactions, functional relationships and biological pathways linked to input genes. Two separate networks were built, one for female gene proteins UHRF1, HLA-E and UBA7 (Figure 25) and a second one for male gene BEX3 and SF3B3.

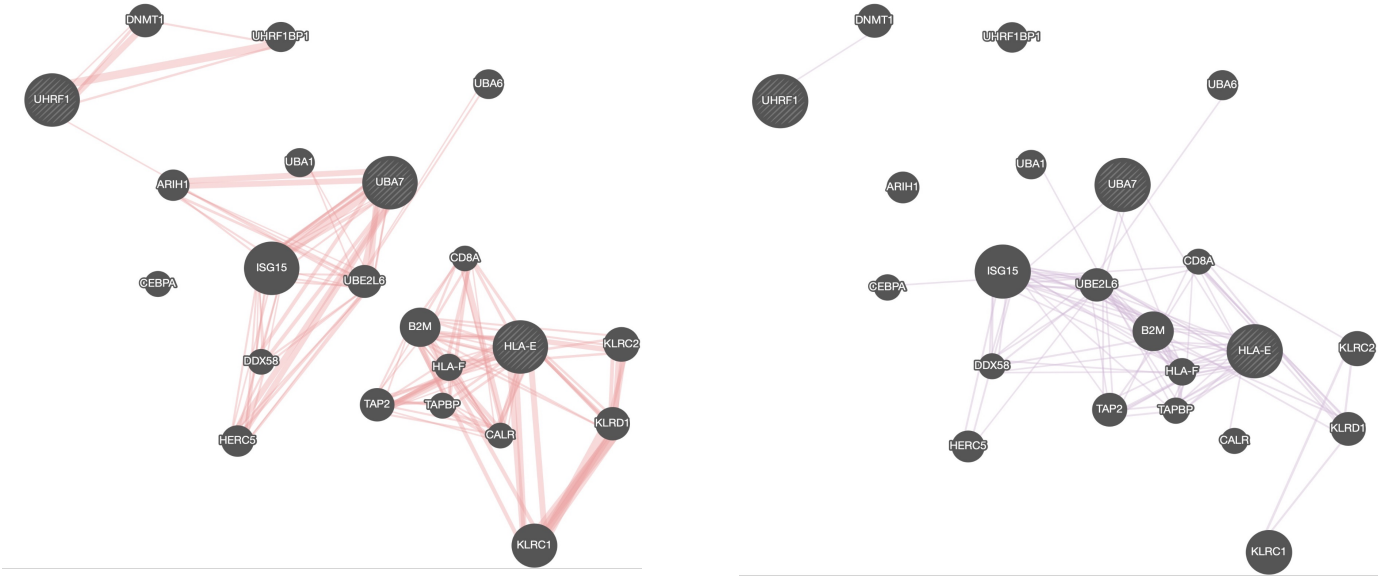


Figure 25: Network created using GeneMANIA
for females-identified genes, UHRF1, HLA-E and UBA7; in pink the physical interactions between the proteins encoded by our genes of interest and their direct interaction genes. On the right, the purple lines represent the genes that are co-expressed among the network.

While the proteins encoded by the two immune related genes, HLA-E and UBA7, have no physical interactions, UHRF1 seems to physically interact with ARIH1 which physically interacts with UBA7, proposing a potential action of UHRF1 on UBA7. In addition, regarding the co-expressed information, we observe HLA-E and UBA7 being co-expressed, which makes sense as they both participate in multiple immune related processes.

We also extracted network functions with a False-Discovery Rate (FDR) lower than 0.05 and that referred to more than 4 genes in the network. All reported functions in Appendix Table 3 are mostly related to HLA-E and the proteins that physically interact with it.

The selected functions exhibited a pronounced involvement in antigen processing and presentation, emphasising their roles in immune responses. Functions such as antigen binding, antigen processing, and presentation via major histocompatibility complex (MHC) class I, alongside various membrane-related activities, highlighted their significance in cellular compartments like the endoplasmic reticulum and vesicles. Moreover, their impact on immune regulation was evident, showcasing roles in lymphocyte-mediated immunity, regulation of cytokine production, and the modulation of type I interferon production.

The network analysis obtained from the male-specific markers highlights *YWHAE* as a gene linking *BEX3* and *SF3B3* (Figure 26). It is a highly conserved gene coding for a protein suggested to participate in several biochemical activities, such as signal transduction, cell division and regulation of insulin sensitivity, and has also been implicated in diseases such as lung cancer (180).

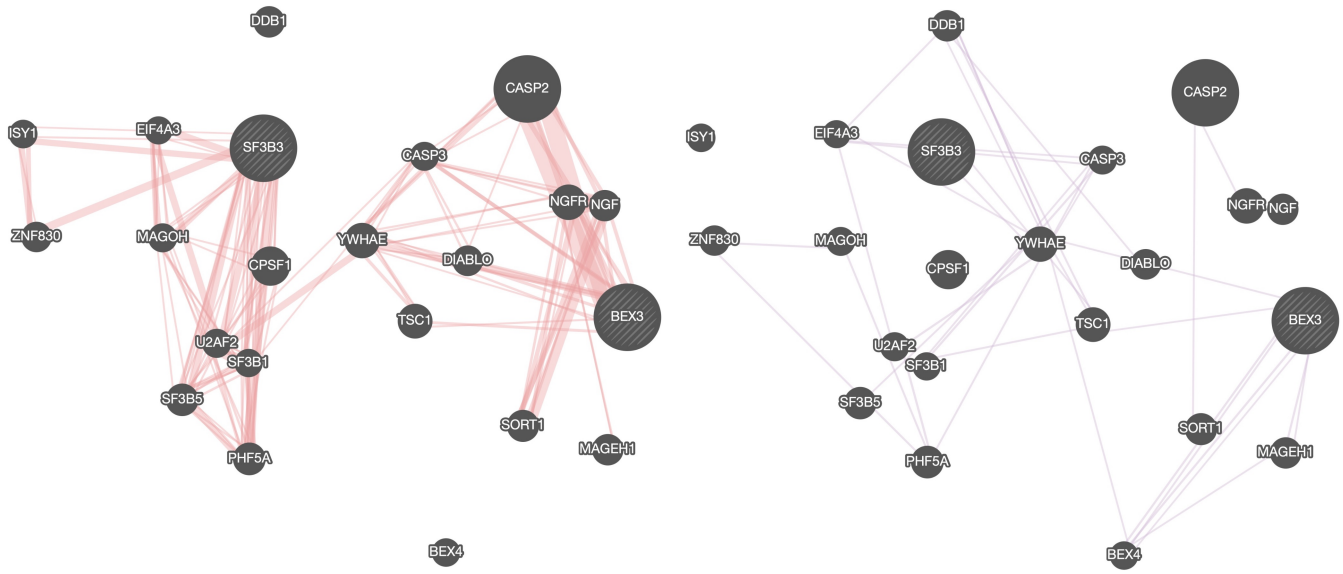


Figure 26: Network created using GeneMANIA for males-identified genes, BEX3 and SF3B3. In pink the physical interactions between the proteins encoded by our genes of interest and their direct interaction proteins. On the right, the purple lines represent the genes that are co-expressed among the network.

The functions related to this network analysis encompassed critical activities, such as RNA splicing (particularly involving U2-type spliceosomal complexes), nucleocytoplasmic transport, and protein export from the nucleus. Additionally, the genes demonstrated associations with the regulation of apoptotic processes, indicating potential roles in cell death mechanisms. Their involvement in nucleic acid transport and RNA localization further underscored their participation in cellular trafficking and molecular localization processes. Moreover, these genes exhibited regulatory influence on enzymatic activities, particularly in peptidase and endopeptidase functions, hinting at their involvement in protein breakdown and cellular regulation (Appendix Table 4) .

5.6 Protein Atlas analysis

The Protein Atlas was used as it offers information crucial for the understanding of protein and RNA expression, single-cell details, and survival outcomes across diverse tissues. Through this platform, various biological aspects of genes are accessible, encompassing RNA and protein expression profiles alongside single-cell insights. Unfortunately, there is no way to stratify protein expression, RNA expression and single cell information based on sex or stage. Only when plotting the KM survival curves there is the option to select a specific sex or stage. Protein and RNA expression of each gene can be found in Appendix Figures 5-9. The 3-year survival high or low percentage refers to the percentage of alive patients with melanomas having higher or lower expression than the selected cut-off.

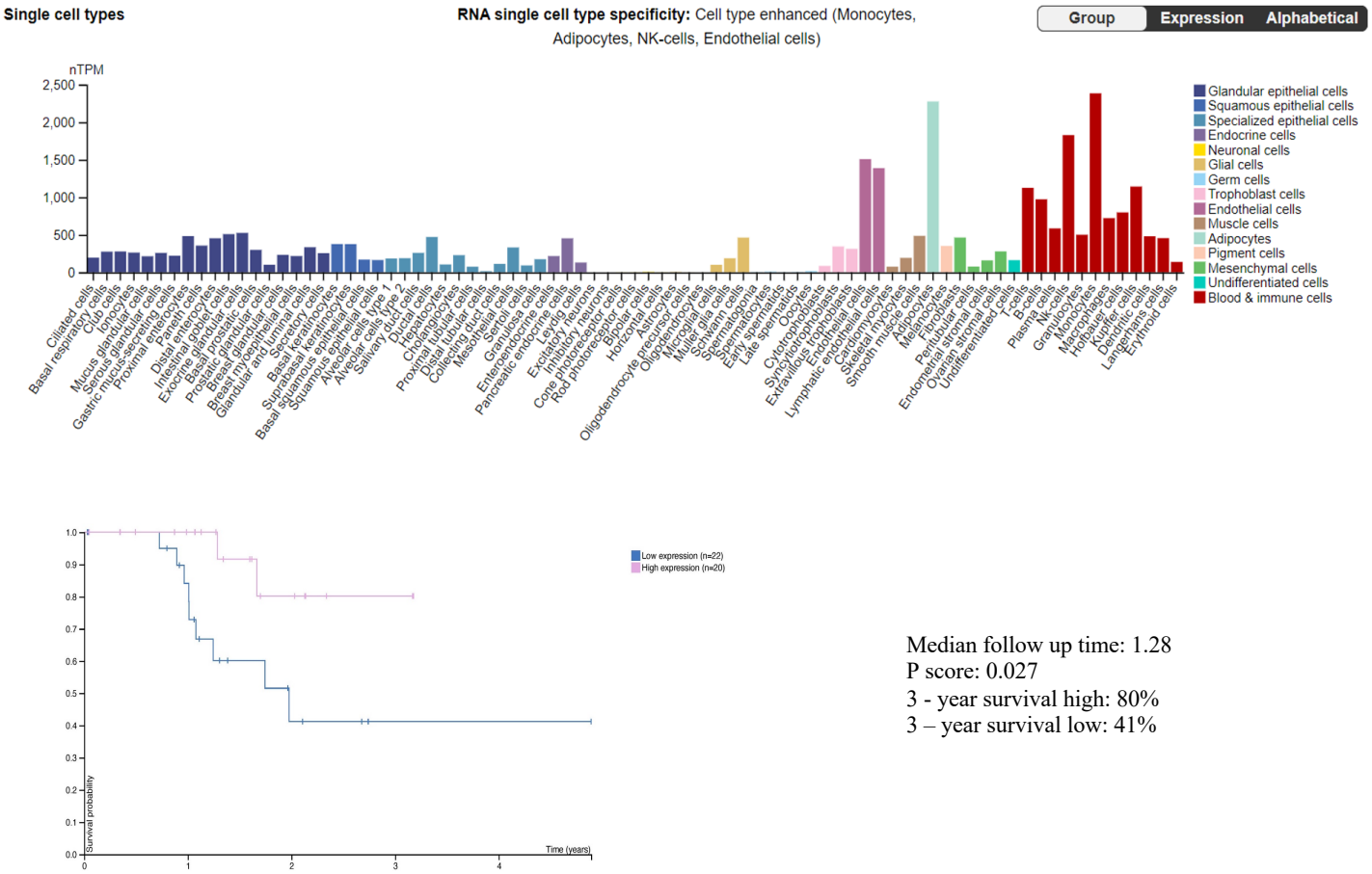


Figure 27: The top plot represents HLA-E normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to HLA-E mRNA expression levels in female primary melanomas from TCGA.

Interestingly, *HLA-E* seems to be highly expressed, under physiological conditions, across numerous immune cells such as monocytes and NK cells. Its high expression is also observed in adipocytes and endothelial cells (Figure 27, top). Regarding the KM curve, high expression of HLA-E in females with primary melanoma seems to be associated with favourable survival, with a p-value of 0.027 (Figure 27, bottom). Protein expression of HLA-E has a high score across lung, lymph nodes and tonsil. HLA-E has low RNA tissue specificity; However, it seems to have quite higher expression in lung and spleen tissue, compared to all the others (Appendix Figure 5).

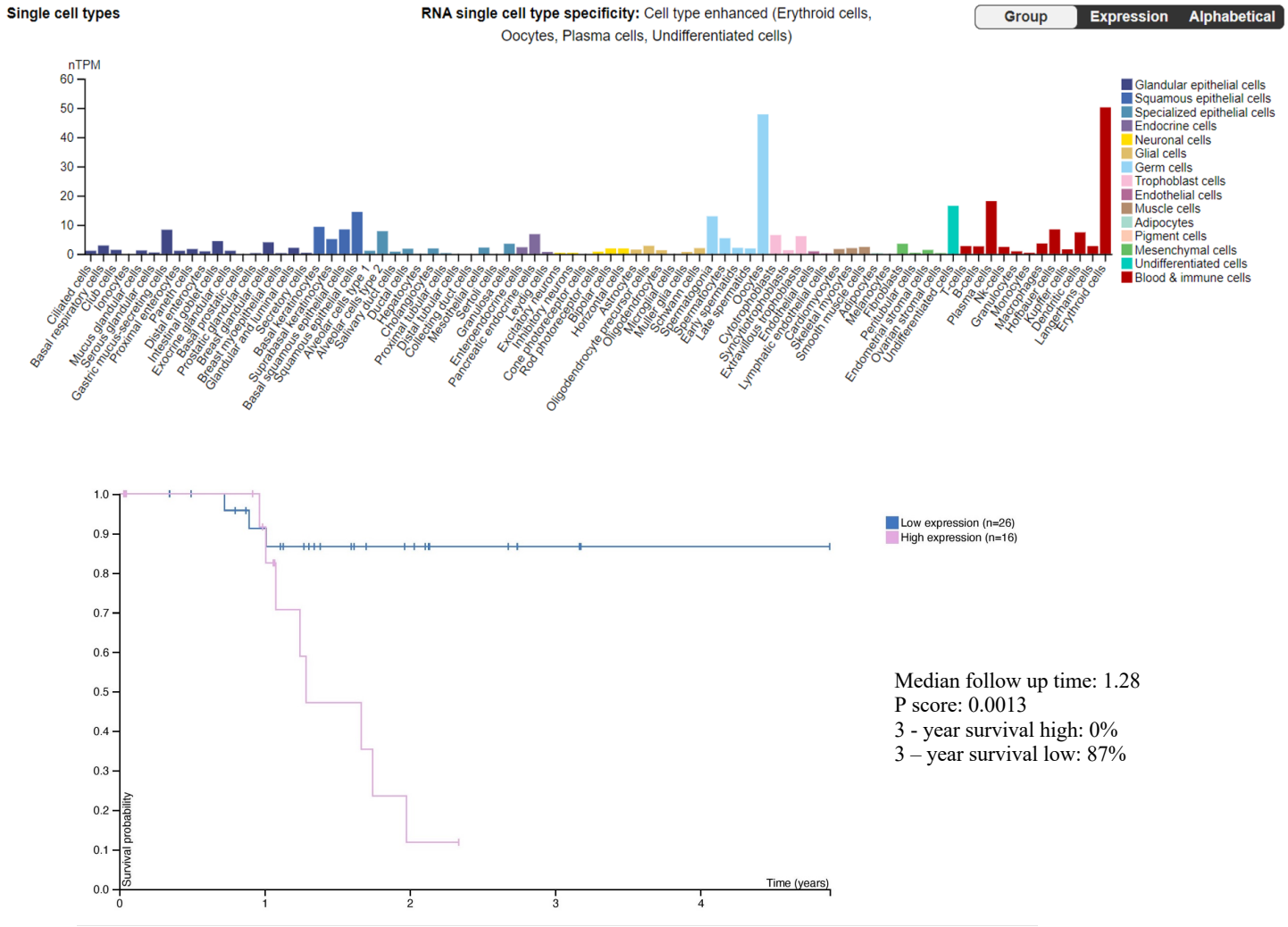


Figure 28: Top plot represents UHRF1 normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to UHRF1 mRNA expression levels in female primary melanomas from TCGA.

Single cell type results for *UHRF1* generally seem to show no specificity across many cell types. However, *UHRF1* expression is enhanced in erythroid cells, oocytes, plasma cells and undifferentiated cells. Female specific survival plot obtained from Protein Atlas based on TCGA associates high expression of *UHRF1* with poor survival, p-value = 0.0013 (Figure 28). Protein expression of *UHRF1* is high only in ovary and thymus tissues and medium across placenta, spleen, tonsil and testis. As for RNA, *UHRF1* protein expression is enriched in bone marrow and lymphoid tissue (Appendix Figure 6).

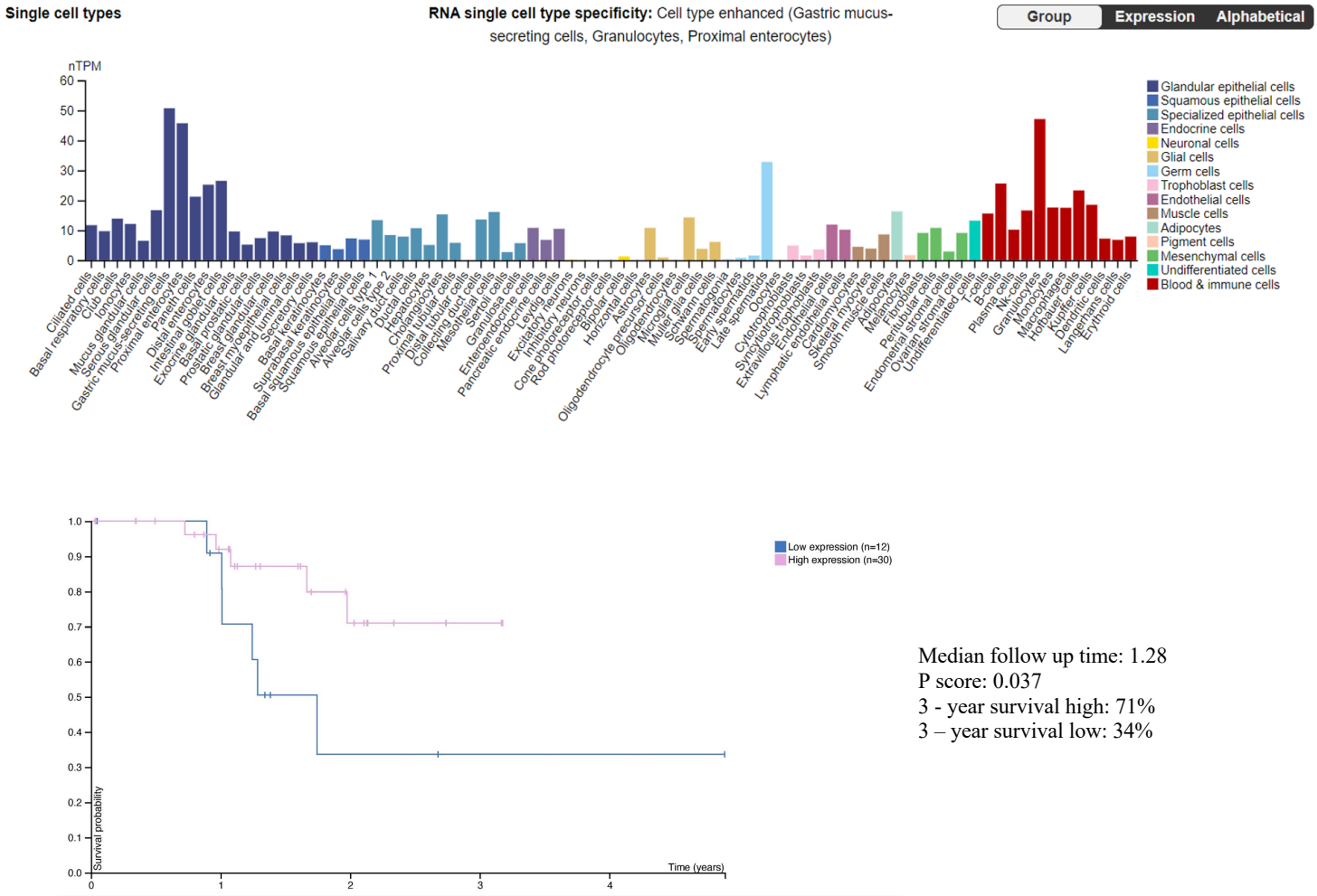


Figure 29: Top plot represents UBA7 normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to UBA7 mRNA expression levels in female primary melanomas from TCGA.

According to protein Atlas, RNA single cell type specificity analysis evidenced that *UBA7* expression is enriched in gastric mucus secreting cells, granulocytes and proximal enterocytes. The survival curve made on females with primary melanoma, significantly (p-value = 0.037) associates high expression of *UBA7* with favourable overall survival (Figure 29).

High *UBA7* protein levels were observed in cerebellum, caudate, kidney, testis, fallopian tube, endometrium and placenta, while average expression of *UBA7* was observed in several tissue types. In addition, *UBA7* RNA expression generally showed low tissue specificity across many tissue types, except for spleen where it was more expressed than compared to other tissue types (Appendix Figure 7).

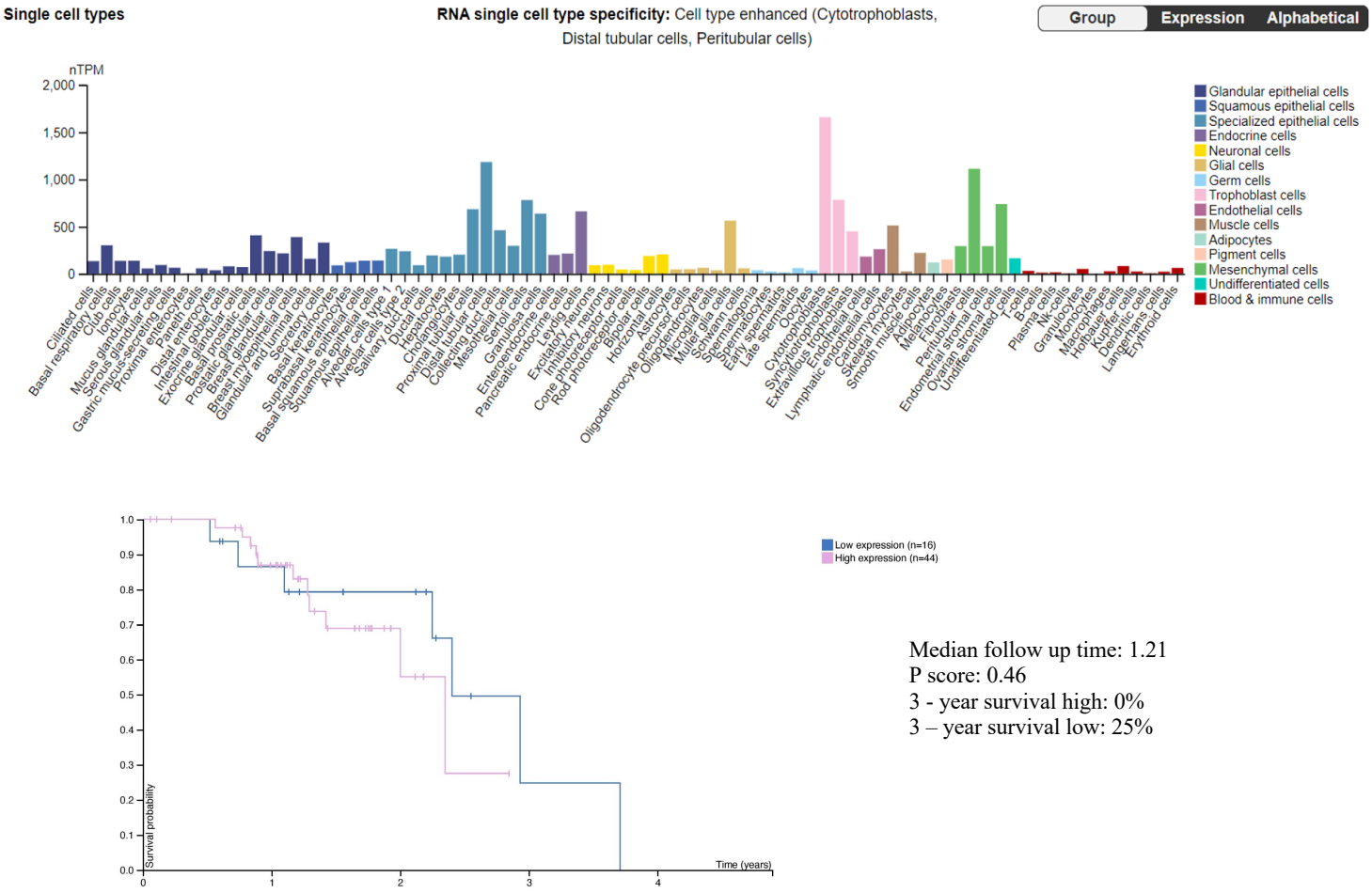


Figure 30: Top plot represents BEX3 normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to BEX3 mRNA expression levels in male primary melanomas from TCGA.

BEX3 is enriched in cytotrophoblasts, distal tubular cells and peritubular cells. KM survival curves on primary melanoma male patients are not significantly associated with survival (Figure 30). *BEX3* protein expression is highly expressed in many tissue types including many brain subregions, the thyroid gland, lungs, stomach, kidney, in many female and male specific tissues and tonsils. Its RNA expression is enhanced in the choroid plexus and many other brain regions (Appendix Figure 8).

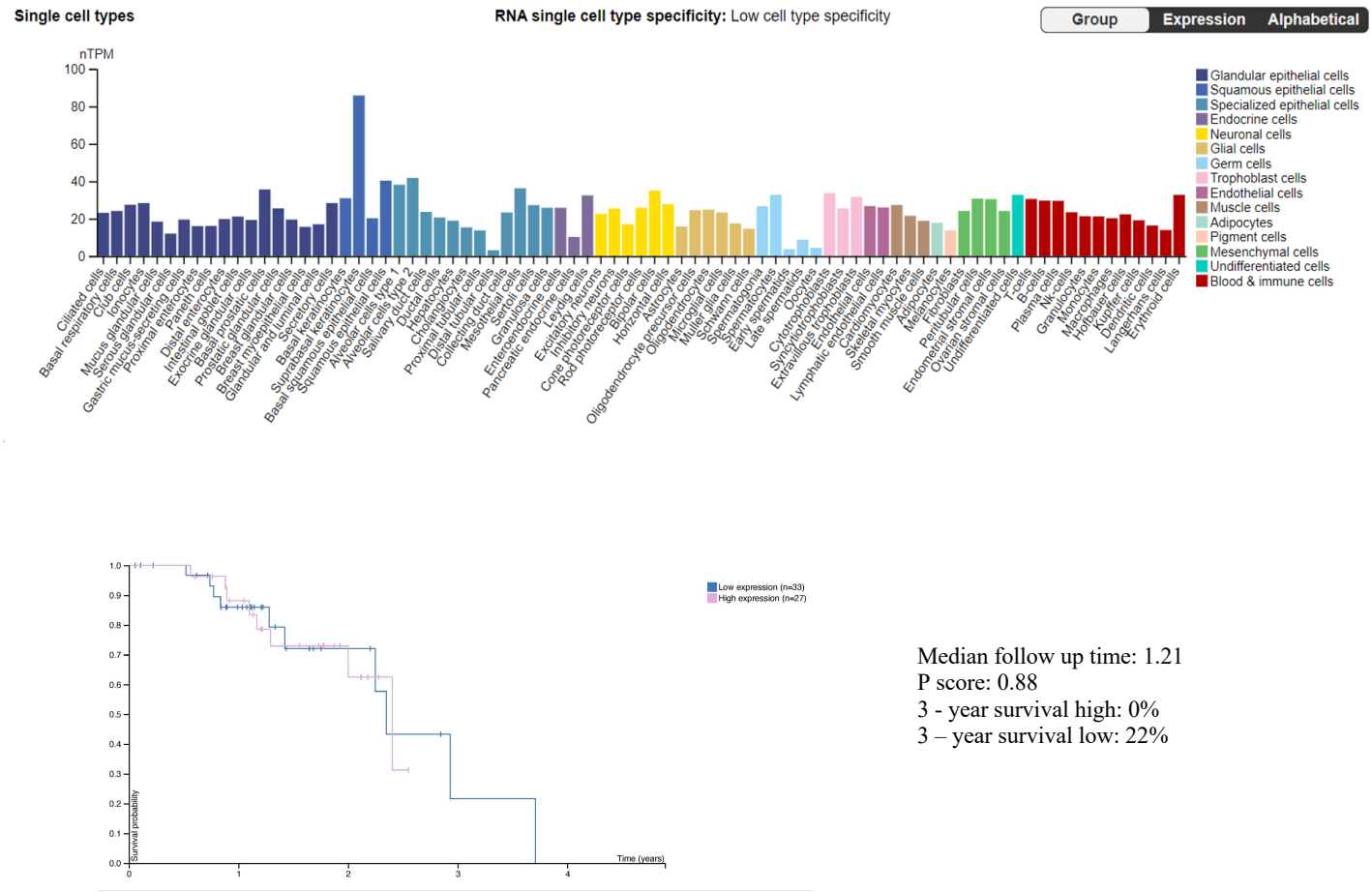


Figure 31: Top plot represents *SF3B3* normalised RNA expression in each cell subtype. The bottom plot shows Kaplan-Meier curves where patients are separated according to *SF3B3* mRNA expression levels in male primary melanomas from TCGA.

SF3B3 has an overall low single cell type specificity unlike all other genes, but has a high expression in Suprabasal keratinocytes. The KM curve shows no significant association between survival and the expression of this gene (Figure 31). In addition, protein and RNA expression of this gene has no tissue specificity and seems to be equally expressed in almost all tissues (Appendix Figure 9).

Summary chapter 8

This chapter provided a detailed characterization of identified sex-specific genes in stage I-II CM by evaluating the correlations with other genes, and also investigating their possible functional roles. Applying Spearman correlation, we determined positively and negatively correlated genes to be included as a downstream analysis of all gene sets considered. Enrichment analysis using the DAVID bioinformatics tool, identified significantly associated

biological processes with the top correlated genes. Furthermore, UHRF1, HLA-E and UBA7 were all associated with immune-related processes in females, indicating that these genes could facilitate or inhibit the progression of cancer. Given that in male patients BEX3 and SF3B3 were associated with lymphocyte activation process and RNA splicing, the oncogenic activity linked to poor survival outcomes is reinforced.

We also found that genes associated with better survival in females (HLA-E and UBA7) predominantly correlated positively with immune cell subtypes, while genes associated with worse survival patterns (UHRF1 in females; BEX3 and SF3B3 in males) showed negative correlations with immune cell enrichment scores.

By GeneMANIA network analysis, we gained insights into genes that physically interact or are co-expressed with the sex-specific genes of interest, identifying possible mechanisms of action by which these interaction partners may influence melanoma progression and patient survival. Last, Protein Atlas analysis aided in exploring single cell expression patterns and survival probability (based on TCGA) based on our genes of interest.

Chapter 9: Discussion

Cutaneous melanoma (CM) is the most lethal type of skin cancer and due to its high ability to metastasize, its most effective treatment is early detection(181–183). This PhD thesis takes an innovative approach to uncover sex-specific survival biomarkers for early-stage CM, acknowledging the higher incidence and mortality rate in males compared to females (184–186), which implies underlying biological variations that need to be further explored (187). Furthermore, biological differences involving hormone levels, genetic expression and the immune system can highly influence progression and treatment of melanoma (188,189). Androgen and oestrogen levels influence melanoma growth and affect immune responses differently in women and men, resulting in different outcomes between the two sexes (190,191). Current skin melanoma studies and prognostic tools do not adequately account for these sex disparities, potentially leading to less effective management of patients (192–194). The aim of the thesis was to explore the interplay of sex and stage in this disease and discover sex-specific prognostic survival biomarkers for early stage CM.

The impact of this research lies in the potential of these biomarkers to stratify patients according to sex into high and low-risk groups, thereby aiding in tailored follow-up care in clinical settings.

Based on literature, most existing CM research either does not account for or adjust for sex. Researchers typically consider sex as a confounding variable, but no melanoma study has stratified by sex. This could potentially overshadow significant differences between the two sexes, leading to inaccurate findings. This study aimed to address this research gap by testing whether stratifying by sex provides more accurate and clinically translatable prognostic information. Instead of treating sex as a confounding variable, sex was considered as an effect modifier.

The key aspects that were evaluated before looking for sex-specific prognostic biomarkers are:

- established gene biomarkers from CM literature, observing that in early-stage CM, their predictive ability in female datasets was much higher compared to male ones;
- gene expression variability, finding out that in early-stage disease female gene expression is significantly less variable than that of males;
- the number and strength of survival predictors, discovering that the biomarkers resulting from sex adjustment or the inclusion of an interaction between sex and gene expression are no longer significant in male-only datasets and differ from those retrieved from sex stratification;
- immune cell subtype differences, highlighting the already known specificities of female and male melanoma microenvironments.

These initial findings reinforced the need for sex stratification to gain more precise prognostic information, leading to better-tailored treatment strategies and follow-up for CM patients.

The larger sample size, the longer follow-up, and availability of early-stage samples in the LMC cohort made it more suitable for our analyses; therefore, many of our results were based on it. In addition, we used the SKCM-TCGA dataset, which has smaller sample size, different tissue origin (frozen instead of FFPE) and generally thicker melanomas, as this cohort also performed other relevant analyses on every sample. Considering that melanoma thickness is one of the principal prognostic factors, TCGA results might not be directly comparable with LMC ones. As for the GSE65904 dataset, frozen tissue samples were used but exact information about staging other than "primary melanoma" was not reported, therefore it might include some stage III samples. On the other hand, GSE53118 contains only stage III samples from frozen lymph nodes, which differ in the tissue origin from the other datasets. However, this dataset was used solely in the investigation of gene expression variability, diversifying our findings beyond skin-based samples. The normal skin and nevi datasets were used to validate our findings, where appropriate, in non-pathological conditions.

In order to evaluate already established CM biomarkers obtained from literature, we assessed them in a mixed population of females and males, as well as in a sex stratified manner on our LMC cohort. We analysed 25 single genes, two gene signatures with 8 and 26 genes each, and

six classes of genes provided by unsupervised clustering. All the selected single-gene and multidimensional signatures extended across various stages of melanoma, including late-stage and metastatic disease. They were identified in samples containing both females and males because to date, there are no sex-specific tissue biomarkers identified for any stage of melanoma. These genes are involved in diverse pro-tumorigenic and anti-tumorigenic pathways. For almost all the assessed genes or gene panels, especially in early-stage disease, the predictive ability was higher in female samples compared to male ones, even surpassing the performance in the combined female and male dataset. This raises concerns about the accuracy of the predictions made with the biomarkers proposed so far, especially in males. And since males are usually at higher risk of recurrence, it further highlights the urgent need of sex-specific prognostic biomarkers

Kaplan–Meier curves stratified by sex obtained from the analysed datasets, confirmed that females have better overall survival in melanoma than males when stages I, II, and III are taken together, finding corroborated by two other studies (188,195). In stage-specific subsets, it was revealed that the impact of sex on survival is stronger in earlier stages (I-II) compared to stage III.

The superior survival observed in females aligns with emerging evidence suggesting sex-specific factors that influence disease progression and treatment response. There are multiple hypotheses that could partly explain this consistent difference, including difference in the immune system between the two sexes. Our hypothesis involves the X-inactivation phenomenon, where randomly silenced genes on the X-chromosome, along with epigenetic modifications that can vary based on gene expression, could potentially explain this "protective" ability in females (196,197).

As the stronger impact of sex was obtained when stages I-II were combined (HR=1.59, CI 1.16 - 2.19, p=0.004) and no consistent difference was found between stage I and II, we used them together as stage I-II for the rest of the analysis. The consistent finding that females exhibit better survival in early stage CM hinders underlying variations between the two sexes and could potentially drive formal re-assessing of multiple prognostic biomarkers in a more sex-specific way.

Despite the favourable observed survival benefit in females of early-stage melanoma, PCA was not useful for segregating females and males, or stages, or patients with different survival outcomes, based on PC1 and PC2, as minimal discrepancies were observed among groups, with highly overlapping ellipses. When sex-specific PCA among all stages was performed, there was slightly higher variance explained by PC1 in females compared to males. Additionally, when repeatedly performing PCA among all datasets analysed, a relatively low variance was captured by PC1. Furthermore, the cumulative sum of the variance percentages from PC1 to PC5 accounted for one-only quarter of the total variance in LMC. The remaining three-quarters unaccounted for may suggest sex-specific complexities that would be excluded from our analysis if we focused solely on the first principal components. Therefore, we opted against using PCA ensuring a more inclusive approach for all genes.

We proceeded with multivariable Cox regression analysis to focus on survival associated variables. Including the gene*sex interaction term in the analysis did not reveal any genes with an opposite association with survival based on sex. However, when the analysis was conducted without the interaction term, a substantial number of genes were found to be associated with survival. Subsequent separate analyses for each sex, adjusting only for age and stage, showed a notable difference: 1475 genes met the criteria for females, while only 2 did so for males. This indicates that the initial combined results were largely influenced by the female population. When this analysis was repeated in stage I-II only, again there was an astonishingly big difference observed in the number of statistically significant genes in females compared to males. Although sample size was not the same in females and males, with 311 females and 255 males with stage I-II melanomas, this discrepancy did not significantly account for the observed difference in the multivariable Cox regression analysis results. Analyses were repeated on random subsets of females with equal sample size as males, and the number of statistically significant genes again was much higher, ranging from 320 to 1653. Moreover, the bigger sample size in females reflected an approximately equal number of events occurring in males, which is more important in survival analysis than sample size. We then argued that biological variability could be the reason for this disparity, as it is well known (but not well accounted for) that the two sexes exhibit distinct biological responses due to hormonal, genetic or other physiological differences (29). These inherent manifestations could be a consequence of divergent gene expression profiles, but of course this would require more investigation on each variation individually. In addition, molecular mechanisms regarding melanoma might affect

the two sexes differently (24). An example might be the response of the immune system, which is known for being different between the two sexes (30,31).

Measuring the variability in gene expression across male and female samples, we aimed to assess sex-related differences in GEV. Interestingly, all tested melanoma datasets were in concordance regarding the GEV difference observed within the sex and stage subcategories. More specifically, a consistently lower GEV in female samples was found in stage I-II melanomas from the LMC, TCGA and GSE65904 datasets. On the other hand, gene expression was more variable in stage III female samples in the LMC, TCGA, and GSE53118 datasets. Stage III GEV was higher than stage I-II GEV in both sexes. It is interesting to note that these differences are observed not only on cutaneous melanoma biopsies but also on the lymph node biopsies from dataset GSE53118. Normal skin and common acquired nevi samples were also analysed to evaluate any difference in GEV between the sexes. Indeed, in both datasets, female samples showed a slightly lower GEV, as also observed by Itoh and Arnold 2015 (32), pointing out that, when it comes to GEV, normal skin and nevi samples are more similar to the early stages of melanoma. In the LMC cohort this difference is perceived when considering all 29,000 genes, when choosing random 3000 gene groups or genes involved in specific biological pathways, such as immune, cell cycle and metabolism related ones. We, therefore, hypothesise that the higher number of genes associated with survival in females compared to male stage I-II samples may be explained by the larger gene expression heterogeneity in males. Hence, the same explanation can be applied to the fact that no genes significantly associated with survival were obtained in stage III samples, neither for females or males, as in both sexes GEV was found to be significantly higher in stage III than in stage I-II samples. This goes in line with the well-reported positive correlation of melanoma heterogeneity with stage progression (143). Consequently, increased genetic instability in the tumour causes transcriptomic and proteomic diversity, which in turn allows microenvironment-driven or cell-intrinsic phenotype-switching, allowing reversible switching between different phenotypes of proliferative and invasive potentials (198,199).

In order to look for sex-specific survival biomarkers, we chose not to start by selecting differentially expressed genes (DEGs) between females and males. The intent was to avoid sole focus on these genes, mainly found on the X and Y chromosomes, recognising that other genes not distinctly differentially expressed may still play a prognostic role in one or both sexes. Notably, in colorectal cancer and glioblastoma multiforme cancers (200,201), multiple genes

with similar expression levels across sexes exhibited varying associations with survival. Our focus was conducting female-specific and male-specific multivariable Cox regression analysis, while accounting for age and stage.

We then proceeded with penalised LASSO Cox regression in females and bidirectional stepwise regression in males, primarily because the female cohort revealed a higher number of statistically significant outcomes compared to males. LASSO Cox regression serves as a technique for variable selection and shrinkage in Cox proportional hazards model (202), whereas bidirectional stepwise analysis fits the best regression model by adding or subtracting variables and finally estimating the lower AIC. The bidirectional stepwise model was used to build a survival model for males, utilising the best predictors out of the seven (five genes plus age and stage) available variables associated with survival, being aware that they had lower statistical power than the variables available for females. Both for females and males, we implemented a 10-fold cross validation approach in order to avoid overfitting and chose the variables retained by all the ten models to fit a final Cox regression model using all the samples available (311 for females, 255 for males). These two final models were then tested on two independent external cohorts.

A limitation of our study is the small number of samples in the validation cohorts and the scarcity of gene expression datasets available online of early-stage melanoma. To assess the performance of the two prognostic survival models, not only we used the UNO area under the time-dependent curve proposed by Uno et al. (2007), which is more appropriate than the standard receiver operating characteristic (ROC) curve AUC, but also expanded our evaluation by including multiple metrics, such as the c-index and the observed/expected ratio (160,203). These additional validation metrics were employed to provide a comprehensive assessment of our models' performance in terms of predictive ability across the validation cohorts. In addition, baseline survival calculation and reporting is also very important as it allows for appropriate validation, model assessment and reproducibility as proposed by DJ Mclernon et al 2022 (204). Further, the more comprehensive evaluation of the model, allows for transparent reporting of the validity of predictions and enhances its potential clinical utility to support clinical decision making.

One of the major decisions that needs to be made when validating a survival predictive model is censored time. As early stage melanoma has generally a long survival, with localised

melanoma of Stage 0, Stage I and Stage II having a 5-year survival rate of 98.4% (205), we extended to the maximum censoring time of 8-years based on our LMC and Biella cohorts available follow-up time. Unfortunately, for the TCGA female and male datasets, because of their limited follow up information, we could only predict survival at 2 years for males and 3 years for females.

The baseline survival of our female stage I-II LMC cohort, calculated by averaging the 10 baseline survival values obtained through cross-validation at 8 years, was 0.81. This implies a 19% probability of mortality (1-0.81), which, as expected, is lower than the calculated probability of mortality in LMC males stage I-II, which is equal to 29%. Our predictive models further confirm what is known in literature, i.e. that males exhibit an increased mortality rate compared to females (206).

In stage I-II female melanomas, three genes always had non zero coefficients persistently across the 10 penalised Cox regression models: *HLA-E*, *UHRF1* and *UBA7*. Utilising these three genes, we constructed a multivariable Cox regression model including all samples. This approach allowed us to calculate beta coefficients from multivariable Cox regression on all 311 females. Without relying on cross validation averages of shrinked coefficients, the model was then validated on two external independent cohorts, the Biella and the TCGA cohort performing well according to UNO AUC (0.804 and 0.795) and UNO C-index (0.763 and 0.716).

Collectively, our analyses focusing on *HLA-E* and *UBA7*, both associated with favourable survival, unveiled interesting findings. When exploring the top genes positively correlated with our genes of interest through Spearman correlation, we found various biological processes especially linked to the immune system. Additionally, numerous immune cell subtype enrichment scores predicted by xCell, when correlated with the expression of these genes, displayed a consistent and notably positive Spearman correlation coefficient. In contrast, the third gene, *UHRF1*, revealed a different pattern, with negative correlation between its gene expression and xCell predictions. Indeed, the top genes positively associated with *UHRF1* contributed to pathways related to cell proliferation and tumour progression.

HLA-E - major histocompatibility complex, class I, E - has a role in both innate and adaptive immunity as it is an important modulator of Natural killer (NK) cells and cytotoxic T

lymphocytes (CTL) (207). This gene has been already found significantly associated with favourable survival in other tumours, such as renal cell carcinoma (208) and glioblastoma (209). Our female-specific multivariable Cox regression analysis in stage I-II female melanomas, revealed a favourable association with survival for monocytes, macrophages and M1 macrophages, indicated by xCell enrichment scores. Notably, *HLA-E* is up-regulated in monocytes to macrophage differentiation (210). Moreover, the presence of adipose tissue beneath the skin (211), proximate to melanoma sites and including adipocytes which were identified by Protein Atlas as overexpressing *HLA-E*, suggests a possible connection and potential interaction between immune cells, melanocytes and the nearby adipose environment. In addition, given the recognised higher proportion of body fat in females compared to males (212) and its mild protective role in females only (213), we hypothesise that the increased presence of adipose tissue, particularly subcutaneous adipocytes, might underlie the protective role of *HLA-E* particularly in females. Further supporting its immune-related function, *HLA-E* is highly expressed by various immune cells including myeloid dendritic cells, B cells, NK cells and T cells (214). Among those immune cell subtypes, B cell and myeloid dendritic cells were also found to be correlated with favourable survival in our research. B cells are known to be enriched in adulthood and old age in females (23), as well as known to be correlated with increased patient survival (215) in melanoma patients. Myeloid dendritic cells (MDC) are antigen-presenting cells responsible for presenting antigens to T cells (216).

UBA7 - ubiquitin-like modifier-activating enzyme 7 - is an enzyme involved in interferon-stimulated gene 15 conjugation (ISG15), which is reported as a negative regulator in IFN immunity and antitumor defence (207), to promote genome stability (217). It was also identified as a protective biomarker in terms of survival in females and known as a tumour suppressor in breast cancer (175) and to improve uveal melanoma OS (218). UBA7 is also known as an interferon-stimulating (ISG) gene that is related to T cell CD8+ infiltration. UBA7 loss is associated with immune-evasion in breast cancer (175). Interestingly, genes negatively correlated with UBA7 were enriched in cell proliferation processes, further supporting underlying immune escape mechanisms. CD8+ T cells were identified by multivariable Cox regression to be positively associated with OS in female stage I-II melanomas. The findings in breast cancer related to UBA7, its positive correlation with genes associated with immune related processes and our female-specific analysis further propose this gene as having a female-specific protective role.

UHRF1 - Ubiquitin-like, containing PHD and RING finger domains 1 - is an essential gene for maintaining DNA methylation. Together with *DNMT1*, it is enriched in DNA replication foci during S phase, further implicating its essential role in cell proliferation (219). Existing literature on *UHRF1* further supports our findings, as it is a well-known gene for its role in tumorigenesis in multiple cancers including adenocarcinoma (220), breast cancer (221) and osteosarcoma (222). High *UHRF1* expression dysregulates immune infiltration in neutrophils, eosinophils and M2 macrophages in hepatocellular carcinoma (223). UHRF1 gene expression was found to be positively correlated with both CD4+ Th2 and Th1 T cells by our analysis. While Th2 are known for their tumour promoting role, Th1 are known for their anti-tumour response through the production of IFNg and IL-2 (224). However, both immune cells are highly plastic and can change phenotype according to their microenvironment (225,226).

Male-specific results in various analyses exhibited lower performance compared to female ones. One exception is BRAF mutation status, which showed significant association with overall survival, after adjustment for age and stage, only in males. This result is interesting, since BRAF mutation has always been considered a prognostic factor without any distinction between females and males (125,227). Specifically, before the advent of BRAF inhibitors, it was a negative prognostic factor (Hazard ratio higher than 1), while when BRAF mutated melanomas began to be treated with targeted therapy (in combination or not with immunotherapy), the association with survival drastically changed because accessibility to personalised treatment improved overall survival. Since the LMC cohort was collected before the advent of targeted therapy for melanoma, and we analysed stage I-II melanomas that do not routinely receive adjuvant treatment, patients' outcome was not influenced by treatment options. If we were to consider BRAF mutation in a prognostic model to apply to newly diagnosed melanomas, its impact would be different due to the big changes in treatment options that occurred in the last years. For this reason, and also since the adjusted p-value was not lower than 0.001, we decided not to include this variable in our male-specific analyses.

The general lower performance of male-specific results was particularly evident in the statistical significance across tests, from the initial multivariable Cox regression to model validation, individual evaluation of identified genes, Spearman correlation analysis, DAVID analysis, multivariable Cox with xCell enrichment scores, and comparative analysis of existing melanoma biomarkers. These findings, highlighting higher statistical performance of female-specific markers, can again be linked to differences in GEV, higher in early stage male

melanomas compared to early stage female ones, which aligns with known slightly lower GEV in multiple non-pathological tissues in females (170). Moreover, a study examining interindividual GEV in 43 tissues from GTEx revealed genes with sex differences involved in diverse biological functions, cell types, and potentially sex-biased diseases (228). These results go in line with Ellis's century-old "greater male variability hypothesis" (229,230) though initially unrelated to gene expression but encompassing various physical, psychological, intellectual and genetic traits. Despite historical exclusion of females from studies due to presumed variability, recent research, including mouse studies contradicting female variability notions, particularly in behavioural and neuroscience research, highlight the need to reevaluate assumptions about female variability in scientific studies (231–234).

Validation of the model combining the two male genes *BEX3* and *SF3B3* with age and stage was only prominent in the Biella cohort, but not significant in TCGA, neither individually nor combined in the prognostic model.

BEX3 - Brain Expressed X-Linked 3- under physiological conditions interacts with a death domain (p75NTR) and mediates apoptosis in neural cells (235). It is also known to contribute to cisplatin chemoresistance in nasopharyngeal carcinomas (177). In our analysis, this gene is negatively correlated with almost all immune cell subtypes and the genes negatively correlated with BEX3 are involved in immune related processes, suggesting an immune inhibition in males. The negative correlation between immune cell enrichment and *BEX3* was also observed in Glioblastoma multiforme (236).

SF3B3 - Splicing Factor 3b Subunit 3 - is a gene that participates in protein splicing responsible for protein diversity, chromatin modification and transcription (237). It has been seen to promote the inclusion of exon 14 on *EZH2* which contributes greatly to its role in renal cancer promotion (238). *SF3B3*-knockdown in prostate cancer resulted in effective growth inhibition (239). This gene is overexpressed in melanomas compared to nevi samples, suggesting a pro-tumorigenic function (240). Its overall pro-tumorigenic role was confirmed in our study only in males, both in combination with *BEX3*, age and stage, and taken individually. In addition, its expression was overall negatively correlated with immune cell enrichments. Indeed, it was recently found to form a complex with AGO3 in the nucleus, to regulate inflammatory diseases by restraining type 2 immunity (179). Another important information about *SF3B3* is that it is

a target of compounds with anticancer and splicing inhibitor functions, such as spliceostatin A, which makes it interesting from the therapeutic point of view.

The three-gene signature identified as specific to females were individually assessed in male stage I-II patients. Similarly, the male-specific two-gene signature was evaluated individually within the female stage I-II cohort. However, neither of these signatures revealed any significant association with overall survival, further supporting the sex-specificity of our signatures.

Immune deconvolution analysis by xCell revealed 24 differentially enriched immune cell types between females and males in stage I-II melanomas, while only 9 immune cell types exhibited distinct enrichment patterns in stage III. This disparity in immune cell type variations between sexes across stages I-II and III parallels our findings in the class comparison between gene expression profiles of the two sexes, where statistically significant gene expression differences present in stage I-II but not in stage III (171). Also, after running multivariable Cox regression analysis on the xCell enrichments scores in all subgroups, only female stage I-II melanomas had 10 significant immune subtypes that positively correlated with overall survival, while no significant immune subtypes were related with survival in male stage I-II melanomas, female and male stage III melanomas. It is well known that adult females generally mount stronger innate and adaptive immune responses compared to males. Both in adulthood and in old age CD4/CD8 ratios and CD4+ T cells are generally higher in females (23). This information aligns with the xCell enriched scores, as 8 out of 24 differentially enriched immune cells in stage I-II are various subtypes of CD4 and CD8 T cells.

Macrophages exhibited differential enrichment between sexes in stage I-II melanomas, but only correlated with improved overall survival in females. Specifically, subtypes “M1” and “Macrophages” generally demonstrated significant association with favourable overall survival in female stage I-II melanomas, even if we found a higher enrichment of these cell types in males. In contrast, M2 enrichment was higher in female versus male stage I-II melanomas, yet not significantly linked to survival in either sex. Nonetheless, further investigation is needed, given the pro-tumorigenic nature of M2-like macrophages and the pro-inflammatory/anti-tumorigenic role of M1 in numerous cancers (241,242). In addition, CD8+ T cells, known to infiltrate tumours that respond better to immune checkpoint blockade therapy (243), are associated with longer survival in female stage I-II melanomas and are more enriched in female

stage I-II melanomas. On the other hand, they are more enriched in male compared to female stage III melanomas. This divergence prompts an exploration into whether sex-specific variations in immune cell composition might underlie the observed differences in responsiveness to immune checkpoint inhibitors. Specifically, could the increased presence of CD8+ T cells in stage III males, together with their overall less active immune response and high antigenicity contribute to their comparatively superior response to these therapies? The elevated levels of CD8+ T cells in earlier stages among females (244) suggest the possibility that initiating immune checkpoint inhibitors at an early stage of disease in females may lead to more favourable treatment outcomes (139,151). These hypotheses necessitate comprehensive exploration to ascertain their implications in treatment response and potential therapeutic strategies.

An intriguing validation of the survival roles attributed to our five genes emerges from correlation analyses involving both gene expression values and xCell-derived enrichment scores across patients. Notably, beyond their survival implications, the genes *HLA-E* and *UBA7*, exhibit significant positive correlations with multiple immune cell subtypes. Conversely, the genes *UHRF1* in stage I-II female melanomas, and *BEX3* and *SF3B3* in stage I-II male melanomas, previously identified as anti-survival genes, display an overall negative correlation with immune cell subtypes.

In examining the gene network comprising *HLA-E*, *UBA7* and *UHRF1*, while direct interactions are absent in the female network, a robust co-expression is evident between *HLA-E* and *UBA7* with the most implicated genes being *ISG15*, *HLA-F*, *TAP2* and *CD8A*. Interestingly, these genes participate in interferon response (245), immune signalling (246), antigen pre-processing, antigen presentation (247), immune surveillance and trigger cytotoxic T cell response (248,249). Additionally, *UHRF1* exhibits both interaction and co-expression with *DNMT1*, highlighting their DNA methylation maintenance role. *DNMT1* is a gene that is characterised as a Hallmark for melanoma (250,251), well known for its pro-tumorigenic association with this disease, further confirming a non-protective role for *UHRF1*. The male network analysis revealed that the two genes seem to have more direct interactions within their network and to also be co-expressed. The gene in the centre of this male network is *YWHAE*, which belongs to the 14-3-3 protein family, mediates signal transduction (252) and is associated with many malignancies, such as colon (253), lung (254) and liver cancer (255).

Upon close examination of our findings, it is clear that most differences between the sexes are visible in the early stages of the disease. This highlights a distinct response in the body's early reaction to melanoma, showing varied gene expressions, immune cell compositions, and survival markers between females and males. These differences raise an important question: should we adjust our approach to focus on sex-specific methods in how we manage patients' follow-up and consider treatment options? The significant distinctions observed in gene expression, survival markers, and immune cell makeup during the initial phases of melanoma onset suggest the need for tailored care. Recognizing the impact of sex-specific details on disease progression and response to treatments might refine our understanding and potentially enhance patient outcomes.

Chapter 10: Strength, limitations, and future directions

Strengths:

- Used an innovative approach treating sex as an effect modifier and not a confounding variable in a non-sex-specific cancer.
- Identified sex-specific prognostic biomarkers in early-stage disease.
- Uncovered lower gene expression variability in females across multiple early-stage melanoma and normal skin datasets.
- Applied comprehensive validation metrics beyond the AUC evaluations.
- Revealed an overall stronger performance in the analyses carried out on female compared to male stage I-II melanomas.

Limitations:

- Had limited sample size in external validation cohorts
- Faced difficulties in finding and accessing gene expression data of early-stage melanoma.
- Was challenged in validation due to varying follow-up times in different cohorts.
- Did not encompass the overall mutation burden, nor any epidemiologic factors such as potential exposures or tumour sites in melanoma.
- The male-specific gene biomarkers were weaker and not completely independent on age and stage.

Implications for future research:

1. This study aims to encourage future studies in cancer and other diseases to consider sex-stratified analysis rather than 1; not considering sex 2; or only consider sex as a confounding variable.
2. Inspire the search for additional sex-specific biomarkers across various cancers and diseases for personalised medicine.
3. Further explore, in terms of sex-specific gene expression, variability and immune cell subtype differences in other cancers.
4. Deepen the understanding of sex-specific molecular mechanisms underlying disease progression.

This study suggests a **fundamental change** in considering sex-specificity in research methodologies and emphasises the potential for sex-specific biomarkers in personalised and targeted medicine.

Chapter 11: Conclusion

This innovative thesis explores areas within melanoma research related to how sex influences early-stage disease. By challenging the historical confounding perception of sex, this study applies it as an effect modifier, unveiling unprecedented insights into sex-specific prognostic biomarkers.

The identification of sex-specific genes—*HLA-E*, *UHRF1*, and *UBA7* in females, and *BEX3* and *SF3B3* in males—proposes a new distinct sex-specific molecular prognostic landscape for early-stage melanoma. Unexpected disparities in gene expression variability across sexes and stages highlight the differences between females and males, particularly present in early-stage disease. Moreover, the profound implications arising from this research go beyond simple biomarker identification. The approach to melanoma survival analysis is redefined, highlighting the need to consider sex-specific complexities for tailored follow-up and clinical interventions.

These revelations, notably the distinct differences in gene prediction patterns between sexes in early-stage melanoma, are in hope of marking a new beginning in sex-specific personalised medicine. The implications extend beyond prognostication, urging a re-evaluation of the methodology around discovery of prognostic biomarkers, thereby propelling the field toward more sex-precise approaches.

References

1. Melanoma Research Alliance [Internet]. [cited 2023 Dec 21]. Understanding Melanoma Staging. Available from: https://www.curemelanoma.org/about-melanoma/melanoma-staging/understanding-melanoma-staging/

2. Brown A, Karkaby L, Perovic M, Shafi R, Einstein G. Sex and Gender Science: The World Writes on the Body. Curr Top Behav Neurosci. 2023;62:3–25.

3. Greaves L, Ritz SA. Sex, Gender and Health: Mapping the Landscape of Research and Policy. Int J Environ Res Public Health. 2022 Jan;19(5):2563.

4. National Academies of Sciences, Engineering, and Medicine; Division of Behavioral and Social Sciences and Education; Committee on National Statistics; Committee on Measuring Sex, Gender Identity, and Sexual Orientation. Measuring Sex, Gender Identity, and Sexual Orientation [Internet]. Becker T, Chin M, Bates N, editors. Washington (DC): National Academies Press (US); 2022 [cited 2024 Feb 28]. (The National Academies Collection: Reports funded by National Institutes of Health). Available from: http://www.ncbi.nlm.nih.gov/books/NBK578625/

5. Clayton JA. Studying both sexes: a guiding principle for biomedicine. FASEB J. 2016 Feb;30(2):519–24.

6. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, Brinton RD, Carrero JJ, DeMeo DL, et al. Sex and gender: modifiers of health, disease, and medicine. Lancet Lond Engl. 2020 Aug 22;396(10250):565–82.

7. Zucker I, Beery AK. Males still dominate animal studies. Nature. 2010 Jun 10;465(7299):690.

8. Yoon DY, Mansukhani NA, Stubbs VC, Helenowski IB, Woodruff TK, Kibbe MR. Sex bias exists in basic science and translational surgical research. Surgery. 2014 Sep;156(3):508–16.

9. Beery AK, Zucker I. Sex Bias in Neuroscience and Biomedical Research. Neurosci Biobehav Rev. 2011 Jan;35(3):565–72.

10. Issues USPHSTF on WH. Women's Health: Report of the Public Health Service Task Force on Women's Health Issues. U.S. Department of Health and Human Services, Public Health Service; 1985. 420 p.

11. Dayton A, Exner EC, Bukowy JD, Stodola TJ, Kurth T, Skelton M, et al. BREAKING THE CYCLE: ESTROUS VARIATION DOES NOT REQUIRE INCREASED SAMPLE SIZE IN THE STUDY OF FEMALE RATS. Hypertens Dallas Tex 1979. 2016 Nov;68(5):1139–44.

12. Meziane H, Ouagazzal AM, Aubert L, Wietrzych M, Krezel W. Estrous cycle effects on behavior of C57BL/6J and BALB/cByJ female mice: implications for phenotyping strategies. Genes Brain Behav. 2007;6(2):192–200.

13. Gomes HL, Graceli JB, Gonçalves WLS, dos Santos RL, Abreu GR, Bissoli NS, et al. Influence of gender and estrous cycle on plasma and renal catecholamine levels in rats. Can J Physiol Pharmacol. 2012 Jan;90(1):75–82.

14. Seeman MV. Psychopathology in Women and Men: Focus on Female Hormones. Am J Psychiatry. 1997 Dec;154(12):1641–7.

15. Austad SN, Fischer KE. Sex Differences in Lifespan. Cell Metab. 2016 Jun 14;23(6):1022–33.

16. Ginter E, Simko V. Women live longer than men. Bratisl Lek Listy. 2013;114(2):45–9.

17. Bredella MA. Sex Differences in Body Composition. Adv Exp Med Biol. 2017;1043:9–27.

18. Ben Mansour G, Kacem A, Ishak M, Grélot L, Ftaiti F. The effect of body composition on strength and power in male and female students. BMC Sports Sci Med Rehabil. 2021 Nov 28;13(1):150.

19. Vogel JA, Friedl KE. Body fat assessment in women. Special considerations. Sports Med Auckl NZ. 1992 Apr;13(4):245–69.

20. St. Pierre SR, Peirlinck M, Kuhl E. Sex Matters: A Comprehensive Comparison of Female and Male Hearts. Front Physiol. 2022 Mar 22;13:831179.

21. Wilkinson NM, Chen HC, Lechner MG, Su MA. Sex Differences in Immunity. Annu Rev Immunol. 2022 Apr 26;40:75–94.

22. Wilson C. ARE THERE SEX DIFFERENCES IN THE IMMUNE SYSTEM? New Sci 1971. 2023 Feb 4;257(3424):40.

23. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016 Oct;16(10):626–38.

24. Wu BN, O'Sullivan AJ. Sex Differences in Energy Metabolism Need to Be Considered with Lifestyle Modifications in Humans. J Nutr Metab. 2011;2011:391809.

25. Tarnopolsky MA. Gender differences in metabolism; nutrition and supplements. J Sci Med Sport. 2000 Sep 1;3(3):287–98.

26. Xin J, Zhang Y, Tang Y, Yang Y. Brain Differences Between Men and Women: Evidence From Deep Learning. Front Neurosci [Internet]. 2019 [cited 2023 Dec 21];13. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2019.00185

27. Bots SH, Peters SAE, Woodward M. Sex differences in coronary heart disease and stroke mortality: a global assessment of the effect of ageing between 1980 and 2010. BMJ Glob Health. 2017 Mar 1;2(2):e000298.

28. Barnes PJ. Sex Differences in Chronic Obstructive Pulmonary Disease Mechanisms. Am J Respir Crit Care Med. 2016 Apr 15;193(8):813–4.

29. Bushnell CD, Chaturvedi S, Gage KR, Herson PS, Hurn PD, Jiménez MC, et al. Sex differences in stroke: Challenges and opportunities. J Cereb Blood Flow Metab. 2018 Dec;38(12):2179–91.

30. Zhu D, Montagne A, Zhao Z. Alzheimer's Pathogenic Mechanisms and Underlying Sex Difference. Cell Mol Life Sci CMLS. 2021 Jun;78(11):4907–20.

31. Kautzky-Willer A, Leutner M, Harreiter J. Sex differences in type 2 diabetes. Diabetologia. 2023 Jun;66(6):986–1002.

32. Kattah AG, Garovic VD. Understanding sex differences in progression and prognosis of chronic kidney disease. Ann Transl Med. 2020 Jul;8(14):897–897.

33. Cooper KM, Delk M, Devuni D, Sarkar M. Sex differences in chronic liver disease and benign liver lesions. JHEP Rep Innov Hepatol. 2023 Nov;5(11):100870.

34. Zheng Z, Zhao W, Zhou Q, Yang Y, Chen S, Hu J, et al. Sex differences in depression, anxiety and health-promoting lifestyles among community residents: A network approach. J Affect Disord. 2023 Nov 1;340:369–78.

35. Feldman S, Ammar W, Lo K, Trepman E, van Zuylen M, Etzioni O. Quantifying Sex Bias in Clinical Studies at Scale With Automated Data Extraction. JAMA Netw Open. 2019 Jul 3;2(7):e196700.

36. Ji H, Niiranen TJ, Rader F, Henglin M, Kim A, Ebinger JE, et al. Sex Differences in Blood Pressure Associations With Cardiovascular Outcomes. Circulation. 2021 Feb 16;143(7):761–3.

37. Fadiran EO, Zhang L. Effects of Sex Differences in the Pharmacokinetics of Drugs and Their Impact on the Safety of Medicines in Women. In: Harrison-Woolrych M, editor. Medicines For Women [Internet]. Cham: Springer International Publishing; 2015 [cited 2024 Mar 11]. p. 41–68. Available from: https://doi.org/10.1007/978-3-319-12406-3_2

38. Soldin O, Mattison D. Sex Differences in Pharmacokinetics and Pharmacodynamics. Clin Pharmacokinet. 2009;48(3):143–57.

39. Rubin JB. The spectrum of sex differences in cancer. Trends Cancer. 2022 Apr 1;8(4):303–15.

40. Kim HI, Lim H, Moon A. Sex Differences in Cancer: Epidemiology, Genetics and Therapy. Biomol Ther. 2018 Jul;26(4):335–42.

41. Vera R, Juan-Vidal O, Safont-Aguilera MJ, de la Peña FA, del Alba AG. Sex differences in the diagnosis, treatment and prognosis of cancer: the rationale for an individualised approach. Clin Transl Oncol. 2023;25(7):2069–76.

42. Cook MB, McGlynn KA, Devesa SS, Freedman ND, Anderson WF. Sex Disparities in Cancer Mortality and Survival. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2011 Aug;20(8):1629–37.

43. Jackson SS, Marks MA, Katki HA, Cook MB, Hyun N, Freedman ND, et al. Sex disparities in the incidence of 21 cancer types: Quantification of the contribution of risk factors. Cancer. 2022;128(19):3531–40.

44. Liu J, Morgan M, Hutchison K, Calhoun VD. A study of the influence of sex on genome wide methylation. PloS One. 2010 Apr 6;5(4):e10028.

45. Singmann P, Shem-Tov D, Wahl S, Grallert H, Fiorito G, Shin SY, et al. Characterization of whole-genome autosomal differences of DNA methylation between men and women. Epigenetics Chromatin. 2015 Oct 19;8:43.

46. García-Calzón S, Perfilyev A, de Mello VD, Pihlajamäki J, Ling C. Sex Differences in the Methylome and Transcriptome of the Human Liver and Circulating HDL-Cholesterol Levels. J Clin Endocrinol Metab. 2018 Dec 1;103(12):4395–408.

47. Hall E, Volkov P, Dayeh T, Esguerra JLS, Salö S, Eliasson L, et al. Sex differences in the genome-wide DNA methylation pattern and impact on gene expression, microRNA levels and insulin secretion in human pancreatic islets. Genome Biol. 2014 Dec 3;15(12):522.

48. Davegårdh C, Hall Wedin E, Broholm C, Henriksen TI, Pedersen M, Pedersen BK, et al. Sex influences DNA methylation and gene expression in human skeletal muscle myoblasts and myotubes. Stem Cell Res Ther. 2019 Jan 15;10(1):26.

49. McCormick H, Young PE, Hur SSJ, Booher K, Chung H, Cropley JE, et al. Isogenic mice exhibit sexually-dimorphic DNA methylation patterns across multiple tissues. BMC Genomics. 2017 Dec 13;18(1):966.

50. Kundakovic M, Tickerhoof M. Epigenetic mechanisms underlying sex differences in the brain and behavior. Trends Neurosci [Internet]. 2023 Nov 15 [cited 2023 Dec 21]; Available from: https://www.sciencedirect.com/science/article/pii/S0166223623002254

51. McCarthy MM, Auger AP, Bale TL, De Vries GJ, Dunn GA, Forger NG, et al. The Epigenetics of Sex Differences in the Brain. J Neurosci. 2009 Oct 14;29(41):12815–23.

52. Dunford A, Weinstock DM, Savova V, Schumacher SE, Cleary JP, Yoda A, et al. Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias. Nat Genet. 2017 Jan;49(1):10–6.

53. Tukiainen T, Villani AC, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. Nature. 2017 Oct 11;550(7675):244–8.

54. Yousefzadeh MJ, Zhao J, Bukata C, Wade EA, McGowan SJ, Angelini LA, et al. Tissue specificity of senescent cell accumulation during physiologic and accelerated aging of mice. Aging Cell. 2020 Mar;19(3):e13094.

55. Ng M, Hazrati LN. Evidence of sex differences in cellular senescence. Neurobiol Aging. 2022 Dec;120:88–104.

56. Ruhland MK, Loza AJ, Capietto AH, Luo X, Knolhoff BL, Flanagan KC, et al. Stromal senescence establishes an immunosuppressive microenvironment that drives tumorigenesis. Nat Commun. 2016 Jun 8;7:11762.

57. Taneja V. Sex Hormones Determine Immune Response. Front Immunol. 2018 Aug 27;9:1931.

58. Harding AT, Heaton NS. The Impact of Estrogens and Their Receptors on Immunity and Inflammation during Infection. Cancers. 2022 Feb 12;14(4):909.

59. Motomura K, Miller D, Galaz J, Liu TN, Romero R, Gomez-Lopez N. The effects of progesterone on immune cellular function at the maternal-fetal interface and in maternal circulation. J Steroid Biochem Mol Biol. 2023 May 1;229:106254.

60. Hall OJ, Klein SL. Progesterone-based compounds affect immune responses and susceptibility to infections at diverse mucosal sites. Mucosal Immunol. 2017 Sep;10(5):1097–107.

61. Lai JJ, Lai KP, Zeng W, Chuang KH, Altuwaijri S, Chang C. Androgen Receptor Influences on Body Defense System via Modulation of Innate and Adaptive Immune Systems. Am J Pathol. 2012 Nov;181(5):1504–12.

62. Ben-Batalla I, Vargas-Delgado ME, von Amsberg G, Janning M, Loges S. Influence of Androgens on Immunity to Self and Foreign: Effects on Immunity and Cancer. Front Immunol [Internet]. 2020 [cited 2023 Dec 21];11. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2020.01184

63. De Francia S, Berchialla P, Armando T, Storto S, Allegra S, Sciannameo V, et al. Colorectal cancer chemotherapy: can sex-specific disparities impact on drug toxicities? Eur J Clin Pharmacol. 2022;78(6):1029–38.

64. Singh S, Parulekar W, Murray N, Feld R, Evans WK, Tu D, et al. Influence of sex on toxicity and treatment outcome in small-cell lung cancer. J Clin Oncol Off J Am Soc Clin Oncol. 2005 Feb 1;23(4):850–6.

65. Morgan AM, Lo J, Fisher DE. How does pheomelanin synthesis contribute to melanomagenesis?: Two distinct mechanisms could explain the carcinogenicity of pheomelanin synthesis. BioEssays News Rev Mol Cell Dev Biol. 2013 Aug;35(8):672–6.

66. Mitra D, Luo X, Morgan A, Wang J, Hoang MP, Lo J, et al. An ultraviolet-radiation-independent pathway to melanoma carcinogenesis in the red hair/fair skin background. Nature. 2012 Nov 15;491(7424):449–53.

67. Elder DE, Bastian BC, Cree IA, Massi D, Scolyer RA. The 2018 World Health Organization Classification of Cutaneous, Mucosal, and Uveal Melanoma: Detailed Analysis of 9 Distinct Subtypes Defined by Their Evolutionary Pathway. Arch Pathol Lab Med. 2020 Apr;144(4):500–22.

68. Saginala K, Barsouk A, Aluru JS, Rawla P, Barsouk A. Epidemiology of Melanoma. Med Sci. 2021 Oct 20;9(4):63.

69. SEER [Internet]. [cited 2024 Jul 15]. Melanoma of the Skin - Cancer Stat Facts. Available from: https://seer.cancer.gov/statfacts/html/melan.html

70. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74(1):12–49.

71. Arnold M, Singh D, Laversanne M, Vignat J, Vaccarella S, Meheus F, et al. Global Burden of Cutaneous Melanoma in 2020 and Projections to 2040. JAMA Dermatol. 2022 May 1;158(5):495–503.

72. Rebecca VW, Sondak VK, Smalley KSM. A Brief History of Melanoma: From Mummies to Mutations. Melanoma Res. 2012 Apr;22(2):114–22.

73. Sober AJ, Balch CM, Thompson JF, Kirkwood JM. A History of Melanoma: From Hunter to Morton. In: Balch C, Atkins M, Garbe C, Gershenwald J, Halpern AC, Kirkwood JM, et al., editors. Cutaneous Melanoma [Internet]. Cham: Springer International Publishing; 2019 [cited 2023 Dec 21]. p. 1–19. Available from: https://doi.org/10.1007/978-3-319-46029-1_69-1

74. Silvers DN, Gorham JD. Observations on a melanoma by William Norris, M.D., a country practitioner of the early 19th century. Am J Dermatopathol. 1982 Oct;4(5):421–4.

75. Roses DF, Harris MN, Gumport SL. Surgery for Primary Cutaneous melanoma. Dermatol Clin. 1985 Apr 1;3(2):315–26.

76. Seebode C, Lehmann J, Emmert S. Photocarcinogenesis and Skin Cancer Prevention Strategies. Anticancer Res. 2016 Mar;36(3):1371–8.

77. Cancer Today [Internet]. [cited 2024 May 23]. Available from: https://gco.iarc.who.int/today/

78. Nikolaou V, Stratigos AJ. Emerging trends in the epidemiology of melanoma. Br J Dermatol. 2014 Jan;170(1):11–9.

79. Linos E, Swetter SM, Cockburn MG, Colditz GA, Clarke CA. Increasing burden of melanoma in the United States. J Invest Dermatol. 2009 Jul;129(7):1666–74.

80. Waseh S, Lee JB. Advances in melanoma: epidemiology, diagnosis, and prognosis. Front Med. 2023;10:1268479.

81. Duarte AF, Sousa-Pinto B, Azevedo LF, Barros AM, Puig S, Malvehy J, et al. Clinical ABCDE rule for early melanoma detection. Eur J Dermatol EJD. 2021 Dec 1;31(6):771–8.

82. Holmes GA, Vassantachart JM, Limone BA, Zumwalt M, Hirokane J, Jacob SE. Using Dermoscopy to Identify Melanoma and Improve Diagnostic Discrimination. Fed Pract. 2018 May;35(Suppl 4):S39–45.

83. Grossarth S, Mosley D, Madden C, Ike J, Smith I, Huo Y, et al. Recent Advances in Melanoma Diagnosis and Prognosis Using Machine Learning Methods. Curr Oncol Rep. 2023 Jun;25(6):635–45.

84. Marghoob AA, Koenig K, Bittencourt FV, Kopf AW, Bart RS. Breslow thickness and clark level in melanoma: support for including level in pathology reports and in American Joint Committee on Cancer Staging. Cancer. 2000 Feb 1;88(3):589–95.

85. Reed RJ, Martin P. Variants of melanoma. Semin Cutan Med Surg. 1997 Jun;16(2):137–58.

86. Papageorgiou C, Apalla Z, Manoli SM, Lallas K, Vakirlis E, Lallas A. Melanoma: Staging and Follow-Up. Dermatol Pract Concept. 2021 Jul;11(Suppl 1):e2021162S.

87. Pathak S, Zito PM. Clinical Guidelines for the Staging, Diagnosis, and Management of Cutaneous melanoma. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 [cited 2023 Dec 21]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK572149/

88. Melanoma Research Alliance [Internet]. [cited 2023 Dec 21]. Breslow Depth and Clark Level. Available from: https://www.curemelanoma.org/about-melanoma/melanoma-staging/breslow-depth-and-clark-level/

89. Sadurní MB, Meves A. Breslow thickness 2.0: Why gene expression profiling is a step toward better patient selection for sentinel lymph node biopsies. Mod Pathol. 2022 Nov;35(11):1509–14.

90. Colloby PS, West KP, Fletcher A. Observer variation in the measurement of Breslow depth and Clark's level in thin cutaneous melanoma. J Pathol. 1991 Mar;163(3):245–50.

91. Prade M, Sancho-Garnier H, Cesarini JP, Cochran A. Difficulties encountered in the application of Clark classification and the Breslow thickness measurement in cutaneous melanoma. Int J Cancer. 1980 Aug;26(2):159–63.

92. Keung EZ, Gershenwald JE. The eighth edition American Joint Committee on Cancer (AJCC) melanoma staging system: implications for melanoma treatment and care. Expert Rev Anticancer Ther. 2018 Aug;18(8):775–84.

93. Bartlett EK, Karakousis GC. Current staging and prognostic factors in melanoma. Surg Oncol Clin N Am. 2015 Apr;24(2):215–27.

94. Gerami P, Busam K, Cochran A, Cook MG, Duncan LM, Elder DE, et al. Histomorphologic assessment and interobserver diagnostic reproducibility of atypical spitzoid melanocytic neoplasms with long-term follow-up. Am J Surg Pathol. 2014 Jul;38(7):934–40.

95. Farmer ER, Gonin R, Hanna MP. Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists. Hum Pathol. 1996 Jun;27(6):528–31.

96. Corona R, Mele A, Amini M, De Rosa G, Coppola G, Piccardi P, et al. Interobserver variability on the histopathologic diagnosis of cutaneous melanoma and other pigmented skin lesions. J Clin Oncol Off J Am Soc Clin Oncol. 1996 Apr;14(4):1218–23.

97. Kim RH, Meehan SA. Immunostain use in the diagnosis of melanomas referred to a tertiary medical center: a 15-year retrospective review (2001-2015). J Cutan Pathol. 2017 Mar;44(3):221–7.

98. Hudson AR, Smoller BR. Immunohistochemistry in diagnostic dermatopathology. Dermatol Clin. 1999 Jul;17(3):667–89, x.

99. Prieto VG, Shea CR. Use of immunohistochemistry in melanocytic lesions. J Cutan Pathol. 2008 Nov;35 Suppl 2:1–10.

100. Orosz Z. Melan-A/Mart-1 expression in various melanocytic lesions and in non-melanocytic soft tissue tumours. Histopathology. 1999 Jun;34(6):517–25.

101. Schwechheimer K, Zhou L. HMB45: a specific marker for melanoma metastases in the central nervous system? Virchows Arch Int J Pathol. 1995;426(4):351–3.

102. Henze G, Dummer R, Joller-Jemelka HI, Böni R, Burg G. Serum S100--a marker for disease monitoring in metastatic melanoma. Dermatol Basel Switz. 1997;194(3):208–12.

103. Hartman ML, Czyz M. MITF in melanoma: mechanisms behind its expression and activity. Cell Mol Life Sci. 2015;72(7):1249–60.

104. Willis BC, Johnson G, Wang J, Cohen C. SOX10: A Useful Marker for Identifying Metastatic Melanoma in Sentinel Lymph Nodes. Appl Immunohistochem Mol Morphol. 2015 Feb;23(2):109.

105. Alomari AK, Ustun B, Aslanian HR, Ge X, Chhieng D, Cai G. Endoscopic ultrasound-guided fine-needle aspiration diagnosis of secondary tumors involving the pancreas: An institution's experience. CytoJournal. 2016;13:1.

106. Nielsen PS, Riber-Hansen R, Steiniche T. Immunohistochemical double stains against Ki67/MART1 and HMB45/MITF: promising diagnostic tools in melanocytic lesions. Am J Dermatopathol. 2011 Jun;33(4):361–70.

107. Heppt MV, Siepmann T, Engel J, Schubert-Fritschle G, Eckel R, Mirlach L, et al. Prognostic significance of BRAF and NRAS mutations in melanoma: a German study from routine care. BMC Cancer. 2017 Aug 10;17(1):536.

108. Compton LA, Murphy GF, Lian CG. Diagnostic Immunohistochemistry in Cutaneous Neoplasia: An Update. Dermatopathol Basel Switz. 2015;2(1):15–42.

109. Garg M, Couturier DL, Nsengimana J, Fonseca NA, Wongchenko M, Yan Y, et al. Tumour gene expression signature in primary melanoma predicts long-term outcomes. Nat Commun. 2021 Feb 18;12:1137.

110. Zager JS, Gastman BR, Leachman S, Gonzalez RC, Fleming MD, Ferris LK, et al. Performance of a prognostic 31-gene expression profile in an independent cohort of 523 cutaneous melanoma patients. BMC Cancer. 2018 Feb 5;18:130.

111. Gerami P, Cook RW, Wilkinson J, Russell MC, Dhillon N, Amaria RN, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. Clin Cancer Res Off J Am Assoc Cancer Res. 2015 Jan 1;21(1):175–83.

112. Thakur R, Laye JP, Lauss M, Diaz JMS, O'Shea SJ, Poźniak J, et al. Transcriptomic Analysis Reveals Prognostic Molecular Signatures of Stage I Melanoma. Clin Cancer Res. 2019 Dec 13;25(24):7424–35.

113. Liu Q, Peng Z, Shen L, Shen L. Prognostic and Clinicopathological Value of Ki-67 in Melanoma: A Meta-Analysis. Front Oncol. 2021 Sep 8;11:737760.

114. Kucher C, Zhang PJ, Pasha T, Elenitsas R, Wu H, Ming ME, et al. Expression of Melan-A and Ki-67 in desmoplastic melanoma and desmoplastic nevi. Am J Dermatopathol. 2004 Dec;26(6):452–7.

115. Ladstein RG, Bachmann IM, Straume O, Akslen LA. Prognostic importance of the mitotic marker phosphohistone H3 in cutaneous nodular melanoma. J Invest Dermatol. 2012 Apr;132(4):1247–52.

116. Ladstein RG, Bachmann IM, Straume O, Akslen LA. Ki-67 expression is superior to mitotic count and novel proliferation markers PHH3, MCM4 and mitosin as a prognostic factor in thick cutaneous melanoma. BMC Cancer. 2010 Apr 14;10:140.

117. Chorny JA, Barr RJ, Kyshtoobayeva A, Jakowatz J, Reed RJ. Ki-67 and p53 expression in minimal deviation melanomas as compared with other nevomelanocytic lesions. Mod Pathol Off J U S Can Acad Pathol Inc. 2003 Jun;16(6):525–9.

118. Ohsie SJ, Sarantopoulos GP, Cochran AJ, Binder SW. Immunohistochemical characteristics of melanoma. J Cutan Pathol. 2008 May;35(5):433–44.

119. Palmer SR, Erickson LA, Ichetovkin I, Knauer DJ, Markovic SN. Circulating serologic and molecular biomarkers in malignant melanoma. Mayo Clin Proc. 2011 Oct;86(10):981–90.

120. Chan SH, Chiang J, Ngeow J. CDKN2A germline alterations and the relevance of genotype-phenotype associations in cancer predisposition. Hered Cancer Clin Pract. 2021 Mar 25;19:21.

121. Masoomian B, Shields JA, Shields CL. Overview of BAP1 cancer predisposition syndrome and the relationship to uveal melanoma. J Curr Ophthalmol. 2018 Mar 22;30(2):102–9.

122. Robles-Espinoza CD, Harland M, Ramsay AJ, Aoude LG, Quesada V, Ding Z, et al. POT1 loss-of-function variants predispose to familial melanoma. Nat Genet. 2014 May;46(5):478–81.

123. Inamdar GS, Madhunapantula SV, Robertson GP. Targeting the MAPK Pathway in Melanoma: Why some approaches succeed and other fail. Biochem Pharmacol. 2010 Sep 1;80(5):624–37.

124. Maldonado JL, Fridlyand J, Patel H, Jain AN, Busam K, Kageshita T, et al. Determinants of BRAF mutations in primary melanomas. J Natl Cancer Inst. 2003 Dec 17;95(24):1878–90.

125. Castellani G, Buccarelli M, Arasi MB, Rossi S, Pisanu ME, Bellenghi M, et al. BRAF Mutations in Melanoma: Biological Aspects, Therapeutic Implications, and Circulating Biomarkers. Cancers. 2023 Jan;15(16):4026.

126. Yazdi AS, Palmedo G, Flaig MJ, Puchta U, Reckwerth A, Rütten A, et al. Mutations of the BRAF gene in benign and malignant melanocytic lesions. J Invest Dermatol. 2003 Nov;121(5):1160–2.

127. Banerji U, Affolter A, Judson I, Marais R, Workman P. BRAF and NRAS mutations in melanoma: potential relationships to clinical response to HSP90 inhibitors. Mol Cancer Ther. 2008 Apr 15;7(4):737–9.

128. Pham D (Daniel) M, Guhan S, Tsao H. KIT and Melanoma: Biological Insights and Clinical Implications. Yonsei Med J. 2020 Jul 1;61(7):562–71.

129. Silva-Rodríguez P, Fernández-Díaz D, Bande M, Pardo M, Loidi L, Blanco-Teijeiro MJ. GNAQ and GNA11 Genes: A Comprehensive Review on Oncogenesis, Prognosis and Therapeutic Opportunities in Uveal Melanoma. Cancers. 2022 Jun 22;14(13):3066.

130. Scolyer RA, Long GV, Thompson JF. Evolving concepts in melanoma classification and their relevance to multidisciplinary melanoma patient care. Mol Oncol. 2011 Apr;5(2):124–36.

131. Read J, Wadt KAW, Hayward NK. Melanoma genetics. J Med Genet. 2016 Jan;53(1):1–14.

132. Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, et al. TERT promoter mutations in familial and sporadic melanoma. Science. 2013 Feb 22;339(6122):959–61.

133. Lopez-Bergami P, Huang C, Goydos JS, Yip D, Bar-Eli M, Herlyn M, et al. Re-wired ERK-JNK signaling pathways in melanoma. Cancer Cell. 2007 May;11(5):447–60.

134. Amaral T, Sinnberg T, Meier F, Krepler C, Levesque M, Niessner H, et al. The mitogen-activated protein kinase pathway in melanoma part I – Activation and primary resistance mechanisms to BRAF inhibition. Eur J Cancer. 2017 Mar 1;73:85–92.

135. Lee C, Collichio F, Ollila D, Moschos S. Historical review of melanoma treatment and outcomes. Clin Dermatol. 2013;31(2):141–7.

136. Manzano JL, Layos L, Bugés C, de Los Llanos Gil M, Vila L, Martínez-Balibrea E, et al. Resistant mechanisms to BRAF inhibitors in melanoma. Ann Transl Med. 2016 Jun;4(12):237.

137. Arozarena I, Wellbrock C. Overcoming resistance to BRAF inhibitors. Ann Transl Med. 2017 Oct;5(19):387.

138. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med. 2017 Apr 19;9(1):34.

139. Knight A, Karapetyan L, Kirkwood JM. Immunotherapy in Melanoma: Recent Advances and Future Directions. Cancers. 2023 Feb 9;15(4):1106.

140. Thibodeau J, Bourgeois-Daigneault MC, Lapointe R. Targeting the MHC Class II antigen presentation pathway in cancer immunotherapy. Oncoimmunology. 2012 Sep 1;1(6):908–16.

141. Mooradian MJ, Sullivan RJ. Immunotherapy in Melanoma: Recent Advancements and Future Directions. Cancers. 2023 Jan;15(16):4176.

142. Wilson MA, Schuchter LM. Chemotherapy for Melanoma. Cancer Treat Res. 2016;167:209–29.

143. Grzywa TM, Paskal W, Włodarski PK. Intratumor and Intertumor Heterogeneity in Melanoma. Transl Oncol. 2017 Dec;10(6):956–75.

144. Scoggins CR, Ross MI, Reintgen DS, Noyes RD, Goydos JS, Beitsch PD, et al. Gender-related differences in outcome for melanoma patients. Ann Surg. 2006 May;243(5):693–8; discussion 698-700.

145. Dika E, Patrizi A, Lambertini M, Manuelpillai N, Fiorentino M, Altimari A, et al. Estrogen Receptors and Melanoma: A Review. Cells. 2019 Nov 19;8(11):1463.

146. Bhari N, Schwaertz RA, Apalla Z, Salerni G, Akay BN, Patil A, et al. Effect of estrogen in malignant melanoma. J Cosmet Dermatol. 2022;21(5):1905–12.

147. Allil PAA, Visconti MA, Castrucci AML, Isoldi MC. Photoperiod and testosterone modulate growth and melanogenesis of s91 murine melanoma. Med Chem Shariqah United Arab Emir. 2008 Mar;4(2):100–5.

148. Morvillo V, Lüthy IA, Bravo AI, Capurro MI, Portela P, Calandra RS, et al. Androgen receptors in human melanoma cell lines IIB-MEL-LES and IIB-MEL-IAN and in human melanoma metastases. Melanoma Res. 2002 Dec;12(6):529–38.

149. Samarkina A, Youssef MK, Ostano P, Ghosh S, Ma M, Tassone B, et al. Androgen receptor is a determinant of melanoma targeted drug resistance. Nat Commun. 2023 Oct 14;14(1):6498.

150. Irelli A, Sirufo MM, D'Ugo C, Ginaldi L, De Martinis M. Sex and Gender Influences on Cancer Immunotherapy Response. Biomedicines. 2020 Jul 21;8(7):232.

151. Wang S, Cowley LA, Liu XS. Sex Differences in Cancer Immunotherapy Efficacy, Biomarkers, and Therapeutic Strategy. Molecules. 2019 Sep 4;24(18):3214.

152. Dakup PP, Porter KI, Little AA, Zhang H, Gaddameedhi S. Sex differences in the association between tumor growth and T cell response in a melanoma mouse model. Cancer Immunol Immunother CII. 2020 Oct;69(10):2157–62.

153. Newton-Bishop JA, Beswick S, Randerson-Moor J, Chang YM, Affleck P, Elliott F, et al. Serum 25-hydroxyvitamin D3 levels are associated with breslow thickness at presentation and survival from melanoma. J Clin Oncol Off J Am Soc Clin Oncol. 2009 Nov 10;27(32):5439–44.

154. Akbani R, Akdemir KC, Aksoy BA, Albert M, Ally A, Amin SB, et al. Genomic Classification of Cutaneous Melanoma. Cell. 2015 Jun 18;161(7):1681–96.

155. Cirenajwis H, Ekedahl H, Lauss M, Harbst K, Carneiro A, Enoksson J, et al. Molecular stratification of metastatic melanoma using gene expression profiling: Prediction of survival outcome and benefit from molecular targeted therapy. Oncotarget. 2015 May 20;6(14):12297–309.

156. Mann GJ, Pupo GM, Campain AE, Carter CD, Schramm SJ, Pianova S, et al. BRAF mutation, NRAS mutation, and the absence of an immune-related expressed gene profile predict poor outcome in patients with stage III melanoma. J Invest Dermatol. 2013 Feb;133(2):509–17.

157. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

158. Dessau RB, Pipper CB. ["R"--project for statistical computing]. Ugeskr Laeger. 2008 Jan 28;170(5):328–30.

159. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med. 2011;30(10):1105–17.

160. Uno H, Cai T, Tian L, Wei LJ. Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. J Am Stat Assoc. 2007;102(478):527–37.

161. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017 Nov 15;18(1):220.

162. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol. 2008 Jun 27;9(1):S4.

163. M U, L F, Bm H, C L, P O, A M, et al. Proteomics. Tissue-based map of the human proteome. Science [Internet]. 2015 Jan 23 [cited 2023 Nov 20];347(6220). Available from: https://pubmed.ncbi.nlm.nih.gov/25613900/

164. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419.

165. Weinstein D, Leininger J, Hamby C, Safai B. Diagnostic and Prognostic Biomarkers in Melanoma. J Clin Aesthetic Dermatol. 2014 Jun;7(6):13–24.

166. Riker AI, Enkemann SA, Fodstad O, Liu S, Ren S, Morris C, et al. The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. BMC Med Genomics. 2008 Apr 28;1:13.

167. Revythis A, Shah S, Kutka M, Moschetta M, Ozturk MA, Pappas-Gogos G, et al. Unraveling the Wide Spectrum of Melanoma Biomarkers. Diagn Basel Switz. 2021 Jul 26;11(8):1341.

168. Mulder EEAP, Johansson I, Grünhagen DJ, Tempel D, Rentroia-Pacheco B, Dwarkasing JT, et al. Using a Clinicopathologic and Gene Expression (CP-GEP) Model to Identify Stage I-II Melanoma Patients at Risk of Disease Relapse. Cancers. 2022 Jun 9;14(12):2854.

169. Bellomo D, Arias-Mejias SM, Ramana C, Heim JB, Quattrocchi E, Sominidi-Damodaran S, et al. Model Combining Tumor Molecular and Clinicopathologic Risk Factors Predicts Sentinel Lymph Node Metastasis in Primary Cutaneous Melanoma. JCO Precis Oncol. 2020 Apr 14;4:PO.19.00206.

170. Itoh Y, Arnold AP. Are females more variable than males in gene expression? Meta-analysis of microarray datasets. Biol Sex Differ. 2015 Oct 29;6(1):18.

171. Chrysanthou E, Sehovic E, Ostano P, Chiorino G. Comprehensive Gene Expression Analysis to Identify Differences and Similarities between Sex- and Stage-Stratified Melanoma Samples. Cells. 2022 Jan;11(7):1099.

172. Sidhu H, Capalash N. UHRF1: The key regulator of epigenetics and molecular target for cancer therapeutics. Tumor Biol. 2017 Feb 1;39(2):1010428317692205.

173. Skaug B, Chen ZJ. Emerging Role of ISG15 in Antiviral Immunity. Cell. 2010 Oct 15;143(2):187–90.

174. Fan JB, Miyauchi-Ishida S, Arimoto K ichiro, Liu D, Yan M, Liu CW, et al. Type I IFN induces protein ISGylation to enhance cytokine expression and augments colonic inflammation. Proc Natl Acad Sci U S A. 2015 Nov 17;112(46):14313–8.

175. Lin M, Li Y, Qin S, Jiao Y, Hua F. Ubiquitin-like modifier-activating enzyme 7 as a marker for the diagnosis and prognosis of breast cancer. Oncol Lett. 2020 Apr;19(4):2773–84.

176. Calvo L, Anta B, López-Benito S, Martín-Rodriguez C, Lee FS, Pérez P, et al. Bex3 Dimerization Regulates NGF-Dependent Neuronal Survival and Differentiation by Enhancing trkA Gene Transcription. J Neurosci. 2015 May 6;35(18):7190–202.

177. Gao W, Li JZH, Chen SQ, Chu CY, Chan JYW, Wong TS. BEX3 contributes to cisplatin chemoresistance in nasopharyngeal carcinoma. Cancer Med. 2017 Feb;6(2):439–51.

178. Xu T, Li X, Zhao W, Wang X, Jin L, Feng Z, et al. SF3B3-regulated mTOR alternative splicing promotes colorectal cancer progression and metastasis. J Exp Clin Cancer Res. 2024 Apr 26;43(1):126.

179. Guidi R, Wedeles C, Xu D, Kolmus K, Headland SE, Teng G, et al. Argonaute3-SF3B3 complex controls pre-mRNA splicing to restrain type 2 immunity. Cell Rep. 2023 Dec 13;42(12):113515.

180. Khorrami A, Sharif Bagheri M, Tavallaei M, Gharechahi J. The functional significance of 14-3-3 proteins in cancer: focus on lung cancer. Horm Mol Biol Clin Investig. 2017 Aug 5;32(3):/j/hmbci.2017.32.issue-3/hmbci-2017-0032/hmbci-2017-0032.xml.

181. Switzer B, Puzanov I, Skitzki JJ, Hamad L, Ernstoff MS. Managing Metastatic Melanoma in 2022: A Clinical Review. JCO Oncol Pract. 2022 May;18(5):335–51.

182. Young GJ, Bi WL, Wu WW, Johanns TM, Dunn GP, Dunn IF. Management of intracranial melanomas in the era of precision medicine. Oncotarget. 2017 Jul 13;8(51):89326–47.

183. Gordon L, Olsen C, Whiteman DC, Elliott TM, Janda M, Green A. Prevention versus early detection for long-term control of melanoma and keratinocyte carcinomas: a cost-effectiveness modelling study. BMJ Open. 2020 Feb 26;10(2):e034388.

184. Morgese F, Sampaolesi C, Torniai M, Conti A, Ranallo N, Giacchetti A, et al. Gender Differences and Outcomes in Melanoma Patients. Oncol Ther. 2020 Feb 4;8(1):103–14.

185. Radkiewicz C, Johansson ALV, Dickman PW, Lambe M, Edgren G. Sex differences in cancer risk and survival: A Swedish cohort study. Eur J Cancer Oxf Engl 1990. 2017 Oct;84:130–40.

186. Schwartz MR, Luo L, Berwick M. Sex Differences in Melanoma. Curr Epidemiol Rep. 2019 Jun;6(2):112–8.

187. Tosakoon S, Lawrence WR, Shiels MS, Jackson SS. Sex differences in cancer incidence rates by race and ethnicity: Results from the surveillance, epidemiology, and end results (SEER) Registry (2000-2019). J Clin Oncol. 2023 Jun;41(16_suppl):10547–10547.

188. Joosse A, de Vries E, Eckel R, Nijsten T, Eggermont AMM, Hölzel D, et al. Gender Differences in Melanoma Survival: Female Patients Have a Decreased Risk of Metastasis. J Invest Dermatol. 2011 Mar 1;131(3):719–26.

189. Shi F, Zhang W, Yang Y, Yang Y, Zhao J, Xie M, et al. Sex Disparities of Genomic Determinants in Response to Immune Checkpoint Inhibitors in Melanoma. Front Immunol [Internet]. 2021 [cited 2023 Dec 9];12. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2021.721409

190. Mori T, Martinez SR, O'Day SJ, Morton DL, Umetani N, Kitago M, et al. Estrogen receptor-alpha methylation predicts melanoma progression. Cancer Res. 2006 Jul 1;66(13):6692–8.

191. Singh N, Khatib J, Chiu CY, Lin J, Patel TS, Liu-Smith F. Tumor Androgen Receptor Protein Level Is Positively Associated with a Better Overall Survival in Melanoma Patients. Genes. 2023 Jan 28;14(2):345.

192. Accounting for sex and gender makes for better science. Nature. 2020 Dec 9;588(7837):196–196.

193. Dakup PP, Greer AJ, Gaddameedhi S. Let's talk about sex: A biological variable in immune response against melanoma. Pigment Cell Melanoma Res. 2022 Mar;35(2):268–79.

194. Dessinioti C, Befon A, Plaka M, Polydorou D, Kypreou K, Champsas G, et al. Independent association of sex, age and the melanoma subtype with histological regression in invasive melanomas: A retrospective study. EJC Skin Cancer. 2024 Jan 1;2:100020.

195. Enninga EAL, Moser JC, Weaver AL, Markovic SN, Brewer JD, Leontovich AA, et al. Survival of cutaneous melanoma based on sex, age, and stage in the United States, 1992–2011. Cancer Med. 2017 Sep 6;6(10):2203–12.

196. Cotton AM, Lam L, Affleck JG, Wilson IM, Peñaherrera MS, McFadden DE, et al. Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. Hum Genet. 2011;130(2):187–201.

197. Balaton BP, Brown CJ. Contribution of genetic and epigenetic changes to escape from X-chromosome inactivation. Epigenetics Chromatin. 2021 Jun 29;14(1):30.

198. Pessoa D de O, Rius FE, Papaiz DD, Ayub ALP, Morais AS, de Souza CF, et al. Transcriptional signatures underlying dynamic phenotypic switching and novel disease biomarkers in a linear cellular model of melanoma progression. Neoplasia N Y N. 2021 Apr;23(4):439–55.

199. Vandamme N, Berx G. Melanoma cells revive an embryonic transcriptional network to dictate phenotypic heterogeneity. Front Oncol. 2014;4:352.

200. Lopes-Ramos CM, Kuijjer ML, Ogino S, Fuchs CS, DeMeo DL, Glass K, et al. Gene Regulatory Network Analysis Identifies Sex-Linked Differences in Colon Cancer Drug Metabolism. Cancer Res. 2018 Oct 1;78(19):5538–47.

201. Yang W, Warrington NM, Taylor SJ, Whitmire P, Carrasco E, Singleton KW, et al. Sex differences in GBM revealed by analysis of patient imaging, transcriptome, and survival data. Sci Transl Med. 2019 Jan 2;11(473):eaao5253.

202. Tibshirani R. The Lasso Method for Variable Selection in the Cox Model. Stat Med. 1997;16(4):385–95.

203. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. BMC Med Res Methodol. 2017 Apr 7;17(1):53.

204. McLernon DJ, Giardiello D, Van Calster B, Wynants L, van Geloven N, van Smeden M, et al. Assessing Performance and Clinical Usefulness in Prediction Models With Survival Outcomes: Practical Guidance for Cox Proportional Hazards Models. Ann Intern Med. 2023 Jan 17;176(1):105–14.

205. Melanoma Research Alliance [Internet]. [cited 2023 Dec 10]. Melanoma Survival Rates. Available from: https://www.curemelanoma.org/about-melanoma/melanoma-staging/melanoma-survival-rates/

206. Smith AJ, Lambert PC, Rutherford MJ. Understanding the impact of sex and stage differences on melanoma cancer patient survival: a SEER-based study. Br J Cancer. 2021 Feb;124(3):671–7.

207. Jeon YJ, Yoo HM, Chung CH. ISG15 and immune diseases. Biochim Biophys Acta Mol Basis Dis. 2010 May;1802(5):485–96.

208. Kren L, Valkovsky I, Dolezel J, Capak I, Pacik D, Poprach A, et al. HLA-G and HLA-E specific mRNAs connote opposite prognostic significance in renal cell carcinoma. Diagn Pathol. 2012 May 29;7(1):58.

209. Kren L, Slaby O, Muckova K, Lzicarova E, Sova M, Vybihal V, et al. Expression of immune-modulatory molecules HLA-G and HLA-E by tumor cells in glioblastomas: an unexpected prognostic significance? Neuropathol Off J Jpn Soc Neuropathol. 2011 Apr;31(2):129–34.

210. Camilli G, Cassotta A, Battella S, Palmieri G, Santoni A, Paladini F, et al. Regulation and trafficking of the HLA-E molecules during monocyte-macrophage differentiation. J Leukoc Biol. 2016 Jan;99(1):121–30.

211. Rivera-Gonzalez G, Shook B, Horsley V. Adipocytes in Skin Health and Disease. Cold Spring Harb Perspect Med. 2014 Mar;4(3):a015271.

212. Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues – the biology of pear shape. Biol Sex Differ. 2012 May 31;3:13.

213. Rask-Andersen M, Ivansson E, Höglund J, Ek WE, Karlsson T, Johansson Å. Adiposity and sex-specific cancer risk. Cancer Cell. 2023 Jun 12;41(6):1186-1197.e4.

214. Grant EJ, Nguyen AT, Lobos CA, Szeto C, Chatzileontiadou DSM, Gras S. The unconventional role of HLA-E: The road less traveled. Mol Immunol. 2020 Apr 1;120:101–12.

215. Selitsky SR, Mose LE, Smith CC, Chai S, Hoadley KA, Dittmer DP, et al. Prognostic value of B cells in cutaneous melanoma. Genome Med. 2019 May 28;11(1):36.

216. Chistiakov DA, Sobenin IA, Orekhov AN, Bobryshev YV. Myeloid dendritic cells: Development, functions, and role in atherosclerotic inflammation. Immunobiology. 2015 Jun;220(6):833–44.

217. Wardlaw CP, Petrini JHJ. ISG15 conjugation to proteins on nascent DNA mitigates DNA replication stress. Nat Commun. 2022 Oct 10;13(1):5971.

218. Gerard C, Shum B, Nathan P, Turajlic S. Immuno-oncology approaches in uveal melanoma: tebentafusp and beyond. Immuno-Oncol Technol [Internet]. 2023 Sep [cited 2023 Dec 11];19. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10362360/

219. Veland N, Chen T. Chapter 2 - Mechanisms of DNA Methylation and Demethylation During Mammalian Development. In: Tollefsbol TO, editor. Handbook of Epigenetics (Second Edition) [Internet]. Academic Press; 2017 [cited 2023 Dec 12]. p. 11–24. Available from: https://www.sciencedirect.com/science/article/pii/B978012805388100002X

220. Kostyrko K, Román M, Lee AG, Simpson DR, Dinh PT, Leung SG, et al. UHRF1 is a mediator of KRAS driven oncogenesis in lung adenocarcinoma. Nat Commun. 2023 Jul 5;14(1):3966.

221. Mancini M, Magnani E, Macchi F, Bonapace IM. The multi-functionality of UHRF1: epigenome maintenance and preservation of genome integrity. Nucleic Acids Res. 2021 Jun 21;49(11):6053–68.

222. Wu SC, Kim A, Gu Y, Martinez DI, Zocchi L, Chen CC, et al. UHRF1 overexpression promotes osteosarcoma metastasis through altered exosome production and AMPK/SEMA3E suppression. Oncogenesis. 2022 Sep 6;11(1):1–12.

223. Li D, Chen B, Zeng Y, Wang H. UHRF1 Could Be a Prognostic Biomarker and Correlated with Immune Cell Infiltration in Hepatocellular Carcinoma. Int J Gen Med. 2021 Oct 13;14:6769–76.

224. Constant SL, Bottomly K. Induction of Th1 and Th2 CD4+ T cell responses: the alternative approaches. Annu Rev Immunol. 1997;15:297–322.

225. Nonaka K, Saio M, Umemura N, Kikuchi A, Takahashi T, Osada S, et al. Th1 polarization in the tumor microenvironment upregulates the myeloid-derived suppressor-like function of macrophages. Cell Immunol. 2021 Nov 1;369:104437.

226. Andreu-Sanz D, Kobold S. Role and Potential of Different T Helper Cell Subsets in Adoptive Cell Therapy. Cancers. 2023 Mar 8;15(6):1650.

227. Alqathama A. BRAF in malignant melanoma progression and metastasis: potentials and challenges. Am J Cancer Res. 2020 Apr 1;10(4):1103–14.

228. Khodursky S, Jiang CS, Zheng EB, Vaughan R, Schrider DR, Zhao L. Sex differences in interindividual gene expression variability across human tissues. PNAS Nexus. 2022 Oct 26;1(5):pgac243.

229. Feingold A. Sex Differences in Variability in Intellectual Abilities: A New Look at an Old Controversy. Rev Educ Res. 1992;62(1):61–84.

230. Shields SA. The Variability Hypothesis: The History of a Biological Model of Sex Differences in Intelligence. Signs J Women Cult Soc. 1982 Jul;7(4):769–97.

231. Kaluve AM, Le JT, Graham BM. Female rodents are not more variable than male rodents: A meta-analysis of preclinical studies of fear and anxiety. Neurosci Biobehav Rev. 2022 Dec 1;143:104962.

232. Graham BM. Battle of the sexes: who is more variable, and does it really matter? Lab Anim. 2023 May;52(5):107–8.

233. Beery AK. Inclusion of females does not increase variability in rodent research studies. Curr Opin Behav Sci. 2018 Oct;23:143–9.

234. Levy DR, Hunter N, Lin S, Robinson EM, Gillis W, Conlin EB, et al. Mouse spontaneous behavior reflects individual variation rather than estrous state. Curr Biol CB. 2023 Apr 10;33(7):1358-1364.e4.

235. Kazi JU, Kabir NN, Rönnstrand L. Brain-Expressed X-linked (BEX) proteins in human cancers. Biochim Biophys Acta BBA - Rev Cancer. 2015 Dec 1;1856(2):226–33.

236. Aisa A, Tan Y, Li X, Zhang D, Shi Y, Yuan Y. Comprehensive Analysis of the Brain-Expressed X-Link Protein Family in Glioblastoma Multiforme. Front Oncol. 2022 Jul 4;12:911942.

237. Sun C. The SF3b complex: splicing and beyond. Cell Mol Life Sci. 2020 Sep 1;77(18):3583–95.

238. Chen K, Xiao H, Zeng J, Yu G, Zhou H, Huang C, et al. Alternative Splicing of EZH2 pre-mRNA by SF3B3 Contributes to the Tumorigenic Potential of Renal Cancer. Clin Cancer Res. 2017 Jul 1;23(13):3428–41.

239. Takayama K ichi, Suzuki T, Fujimura T, Yamada Y, Takahashi S, Homma Y, et al. Dysregulation of spliceosome gene expression in advanced prostate cancer by RNA-binding protein PSF. Proc Natl Acad Sci. 2017 Sep 26;114(39):10461–6.

240. Hakobyan S, Loeffler-Wirth H, Arakelyan A, Binder H, Kunz M. A Transcriptome-Wide Isoform Landscape of Melanocytic Nevi and Primary Melanomas Identifies Gene Isoforms Associated with Malignancy. Int J Mol Sci [Internet]. 2021 Jul [cited 2023 Dec 12];22(13). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8268681/

241. Mantovani A, Sozzani S, Locati M, Allavena P, Sica A. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. Trends Immunol. 2002 Nov;23(11):549–55.

242. Zhou Q, Fang T, Wei S, Chai S, Yang H, Tao M, et al. Macrophages in melanoma: A double-edged sword and targeted therapy strategies (Review). Exp Ther Med. 2022 Aug 26;24(4):640.

243. Zappasodi R, Merghoub T, Wolchok JD. Emerging Concepts for Immune Checkpoint Blockade-Based Combination Therapies. Cancer Cell. 2018 Oct 8;34(4):690.

244. Ma J, Yao Y, Tian Y, Chen K, Liu B. Advances in sex disparities for cancer immunotherapy: unveiling the dilemma of Yin and Yang. Biol Sex Differ. 2022 Oct 22;13(1):58.

245. Perng YC, Lenschow DJ. ISG15 in antiviral immunity and beyond. Nat Rev Microbiol. 2018 Jul;16(7):423–39.

246. Dulberger CL, McMurtrey CP, Hölzemer A, Neu KE, Liu V, Steinbach AM, et al. Human leukocyte antigen F (HLA-F) presents peptides and regulates immunity through interactions with NK-cell receptors. Immunity. 2017 Jun 20;46(6):1018-1029.e7.

247. Lin A, Yan WH. The Emerging Roles of Human Leukocyte Antigen-F in Immune Modulation and Viral Infection. Front Immunol [Internet]. 2019 [cited 2023 Dec 22];10. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2019.00964

248. Ritz U, Seliger B. The Transporter Associated With Antigen Processing (TAP): Structural Integrity, Expression, Function, and Its Clinical Relevance. Mol Med. 2001 Mar;7(3):149–58.

249. Du Y, Zuo L, Xiong Y, Wang X, Zou J, Xu H. CD8A is a Promising Biomarker Associated with Immunocytes Infiltration in Hyperoxia-Induced Bronchopulmonary Dysplasia. J Inflamm Res. 2023 Apr 17;16:1653–69.

250. Gassenmaier M, Rentschler M, Fehrenbacher B, Eigentler TK, Ikenberg K, Kosnopfel C, et al. Expression of DNA Methyltransferase 1 Is a Hallmark of Melanoma, Correlating with Proliferation and Response to B-Raf and Mitogen-Activated Protein Kinase Inhibition in Melanocytic Tumors. Am J Pathol. 2020 Oct;190(10):2155–64.

251. Zhang X, Bustos MA, Shoji Y, Ramos RI, Iida Y, Gentry R, et al. Acetylated DNMT1 Downregulation and Related Regulatory Factors Influence Metastatic Melanoma Patients Survival. Cancers. 2021 Sep 18;13(18):4691.

252. Bjeije H, Soltani BM, Behmanesh M, Zali MR. YWHAE long non-coding RNA competes with miR-323a-3p and miR-532-5p through activating K-Ras/Erk1/2 and

PI3K/Akt signaling pathways in HCT116 cells. Hum Mol Genet. 2019 Oct 1;28(19):3219–31.

253. Liu Q, Dai Y, Yu H, Shen Y, Deng J, Lu W, et al. [NKD1 promotes glucose uptake in colon cancer cells by activating YWHAE transcription]. Nan Fang Yi Ke Da Xue Xue Bao. 2023 Apr 20;43(4):585–9.

254. Qi W, Liu X, Qiao D, Martinez JD. Isoform-specific expression of 14-3-3 proteins in human lung cancer tissues. Int J Cancer. 2005 Jan 20;113(3):359–63.

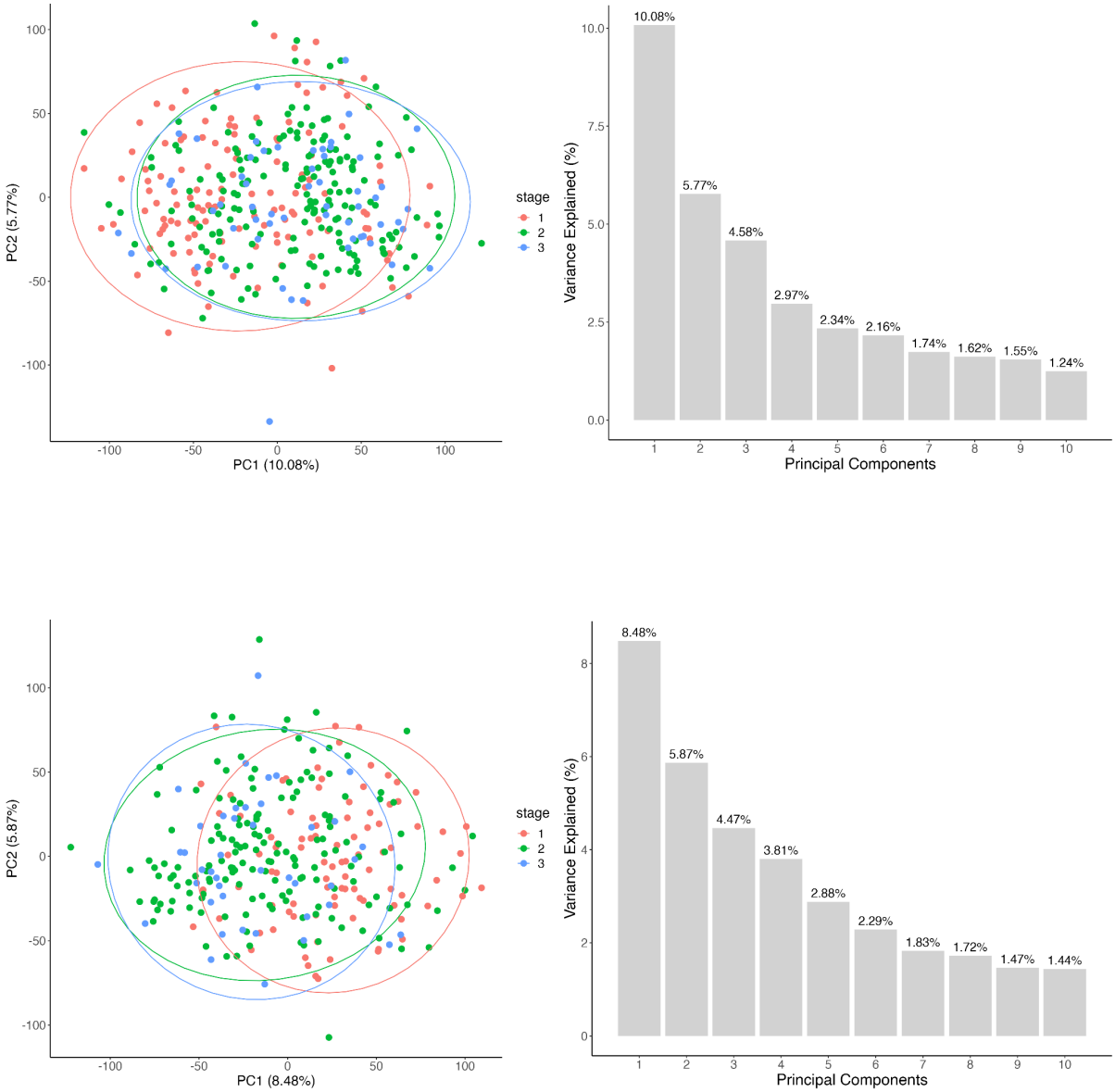
255. Ko BS, Chang TC, Hsu C, Chen YC, Shen TL, Chen SC, et al. Overexpression of 14-3-3ε predicts tumour metastasis and poor survival in hepatocellular carcinoma. Histopathology. 2011 Apr;58(5):705–11.

Appendix

Appendix Table Gerami et al: this table shows the coefficients and raw p-values of the 26 genes that were evaluated in a univariate cox model in females and males together and for each sex separately.

Univariate Cox regression analysis on stage 1,2

		Females/Males		Females		Males	
OGS	ilmn_Ids	coeff	raw_pval	coeff	raw_pval	coeff	raw_pval
BAP1	ilmn_1768363	-0.1131697	0.43277328	-0.146431	0.50372485	-0.0966316	0.61025844
SPP1	ilmn_1651354	0.10450678	0.04935817	0.16341204	0.03918129	0.05759916	0.42601717
CXCL14	ilmn_1748323	-0.1502832	0.00090831	-0.2086808	0.00069023	-0.0860885	0.1967502
CLCA2	ilmn_1803236	-0.0339538	0.33540285	-0.0981829	0.04214433	0.03271842	0.51852772
S100A8	ilmn_1729801	-0.0183731	0.63382732	-0.0673874	0.19122534	0.04737669	0.40894427
BTG1	ilmn_1775743	-0.2113652	0.26879816	-0.3899736	0.1336957	-0.0176018	0.94954463
SAP130	ilmn_1700044	-0.1633633	0.34739257	-0.0716464	0.79452327	-0.1844859	0.42804482
ARG1	ilmn_1812281	0.00392211	0.89638619	-0.0092855	0.81321829	0.02307746	0.62188866
KRT6B	ilmn_1721354	-0.0196437	0.58924704	-0.1039019	0.03850444	0.05522539	0.28370222
GJA1	ilmn_1727087	-0.1228872	0.05368132	-0.170808	0.05493377	-0.0736919	0.42343941
ID2	ilmn_1793990	-0.4116541	0.00005914	-0.6362492	0.00082493	-0.297173	0.02659153
EIF1B	ilmn_1679324	0.12691441	0.39430279	0.15312395	0.49542963	0.10016981	0.61617681
S100A9	ilmn_1750974	-0.0451422	0.34527162	-0.1069737	0.09038744	0.03073612	0.67073223
CRABP2	ilmn_1690170	-0.0565653	0.39986432	-0.0782221	0.41708663	-0.0297083	0.75085116
KRT14	ilmn_1665035	-0.049807	0.20195582	-0.1573893	0.00532756	0.02156164	0.68569674
ROBO1	ilmn_1666468	0.14200412	0.2702264	0.03989639	0.80752827	0.23945774	0.1802506
RBM23	ilmn_1780756	0.57892216	0.02314599	0.18248365	0.62894905	0.87046302	0.01137622
TACSTD2	ilmn_1739001	-0.0533238	0.15725571	-0.1208729	0.02477668	-0.0011647	0.98286038
DSC1	ilmn_1730284	-0.001025	0.97102613	-0.040874	0.26975766	0.04981211	0.25736745
SPRR1B	ilmn_1711174	-0.0015455	0.96074106	-0.0576074	0.18157584	0.05888035	0.19846998
TRIM29	ilmn_1741755	-0.0070153	0.8403605	-0.069153	0.16391841	0.04930675	0.31384245
AQP3	ilmn_1651574	-0.0780323	0.26228918	-0.2179107	0.03278404	0.05111126	0.58978509
TYRP1	ilmn_2054652	0.00623041	0.87617637	0.00504103	0.93461613	0.00515562	0.92281546
PPL	ilmn_1806030	-0.0819602	0.15990836	-0.1694	0.04389314	0.00010735	0.99895114
LTA4H	ilmn_1690342	0.18680919	0.16270322	0.22891438	0.25860116	0.12711426	0.47478989
CST6	ilmn_1698666	-0.0091998	0.75189674	-0.0492875	0.21313614	0.03510487	0.42482525



Appendix Figure 1: Stage stratified PCA on females only (top plot) and males only (bottom plot).

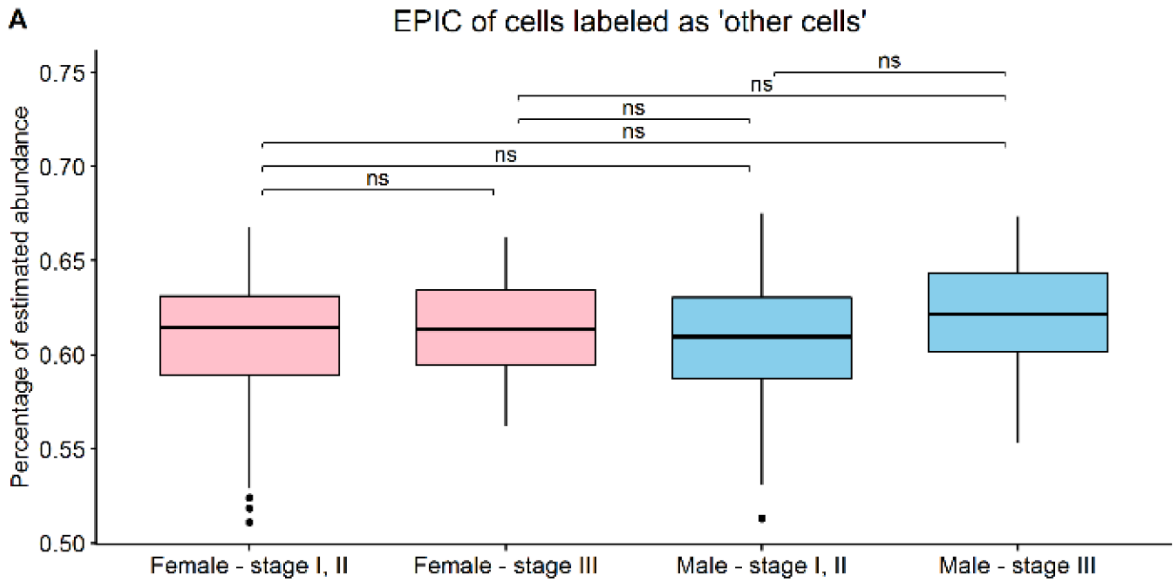
Table 2: Multivariable Cox regression analysis on each gene in TCGA adjusting for age and stage (top table females only, bottom table males only)

ilmn_Ids	OGS	coef	exp(coef)	se(coef)	z	Pr(>|z|)	BH
ilmn_2370091	**BEX3**	1.03876	2.82570	0.22299	4.65834	0.00000	0.09357
ilmn_1747052	**ITGA4**	0.82024	2.27105	0.19784	4.14607	0.00003	0.40216
ilmn_1677997	**MAZ**	1.25086	3.49334	0.34337	3.64291	0.00027	0.69318
ilmn_2411282	**QSOX1**	0.92862	2.53102	0.24774	3.74838	0.00018	0.69318
ilmn_1803110	**SF3B3**	0.88132	2.41410	0.22822	3.86177	0.00011	0.69318
ilmn_1790577	SLC35F2	-0.28940	0.74871	0.07887	-3.66954	0.00024	0.69318
ilmn_1740045	ESR2	-0.33382	0.71618	0.09189	-3.63285	0.00028	0.69318
ilmn_1710899	TSPAN18	-0.50978	0.60063	0.13863	-3.67723	0.00024	0.69318
ilmn_1745021	SLC30A1	-0.58788	0.55550	0.16195	-3.63004	0.00028	0.69318
ilmn_1678260	BCORL2	-0.66689	0.51330	0.18366	-3.63108	0.00028	0.69318
ilmn_1699489	TUBB6	-0.75146	0.47168	0.19752	-3.80450	0.00014	0.69318
ilmn_1769409	C9ORF123	0.50438	1.65596	0.14126	3.57059	0.00036	0.71692
ilmn_3251217	PDXDC2	-0.24550	0.78232	0.06890	-3.56321	0.00037	0.71692
ilmn_1702383	CNGB1	-0.91491	0.40055	0.25622	-3.57078	0.00036	0.71692
ilmn_1731175	XKR6	0.75699	2.13186	0.21401	3.53716	0.00040	0.74152
ilmn_1792314	ACTR1A	-0.79778	0.45033	0.22656	-3.52130	0.00043	0.74152
ilmn_3235404	SNORA57	-1.63146	0.19564	0.46586	-3.50206	0.00046	0.75290
ilmn_3166332	ERCC-00131-02	0.70349	2.02079	0.20546	3.42389	0.00062	0.81121
ilmn_1805995	MGC46496	0.44874	1.56634	0.13137	3.41594	0.00064	0.81121
ilmn_1727567	OLIG2	0.30662	1.35882	0.08976	3.41601	0.00064	0.81121
ilmn_1765299	PRIMA1	0.28042	1.32368	0.08117	3.45462	0.00055	0.81121
ilmn_2370976	FER1L3	-0.44206	0.64271	0.12883	-3.43127	0.00060	0.81121
ilmn_2229170	CRSP9	0.77718	2.17533	0.23093	3.36540	0.00076	0.81833
ilmn_1763883	KRTAP19-3	0.70002	2.01378	0.21191	3.30330	0.00096	0.81833
ilmn_3308642	MIR597	0.47880	1.61414	0.14129	3.38876	0.00070	0.81833
ilmn_2377185	TCEB2	0.43932	1.55166	0.12997	3.38013	0.00072	0.81833
ilmn_1722680	C18ORF4	-0.41995	0.65708	0.12576	-3.33929	0.00084	0.81833
ilmn_2360710	TPM1	-0.50608	0.60285	0.15310	-3.30566	0.00095	0.81833
ilmn_3310196	MIR1302-5	-0.79403	0.45202	0.24013	-3.30660	0.00094	0.81833

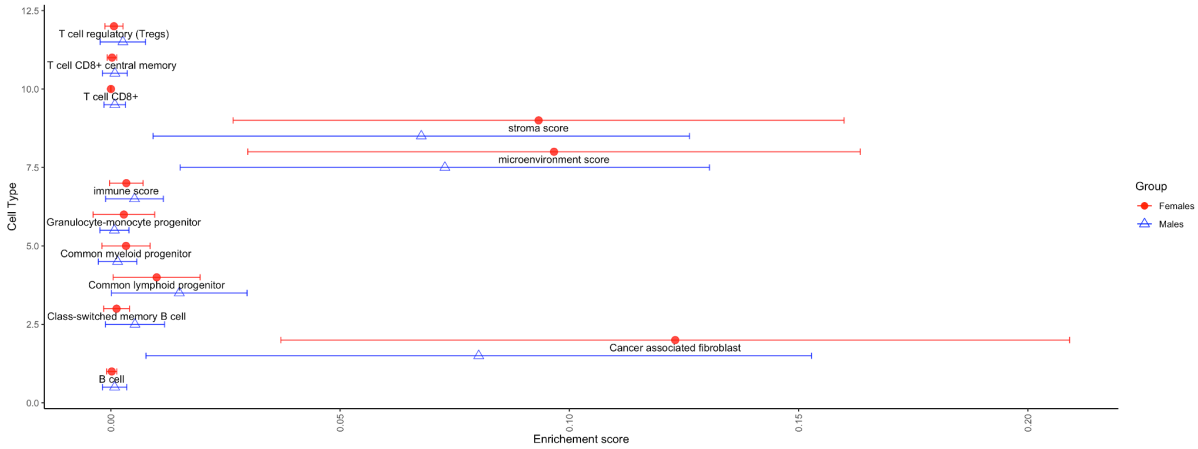
Appendix Table 1: Multivariable Cox regression analysis of raw p-value < 0.001 significant genes on stage I-II males, adjusting for age and stage.

	TCGA Females stage I-II - adjusted for age and stage	
Genes	beta coefficient	raw p-value
UHRF1	0.68265	0.29215
HLA-E	-0.51014	0.35057
UBA7	0.12021	0.61689
NGFRAP1 (BEX3)	0.23570	0.64898
SF3B3	0.15262	0.89501

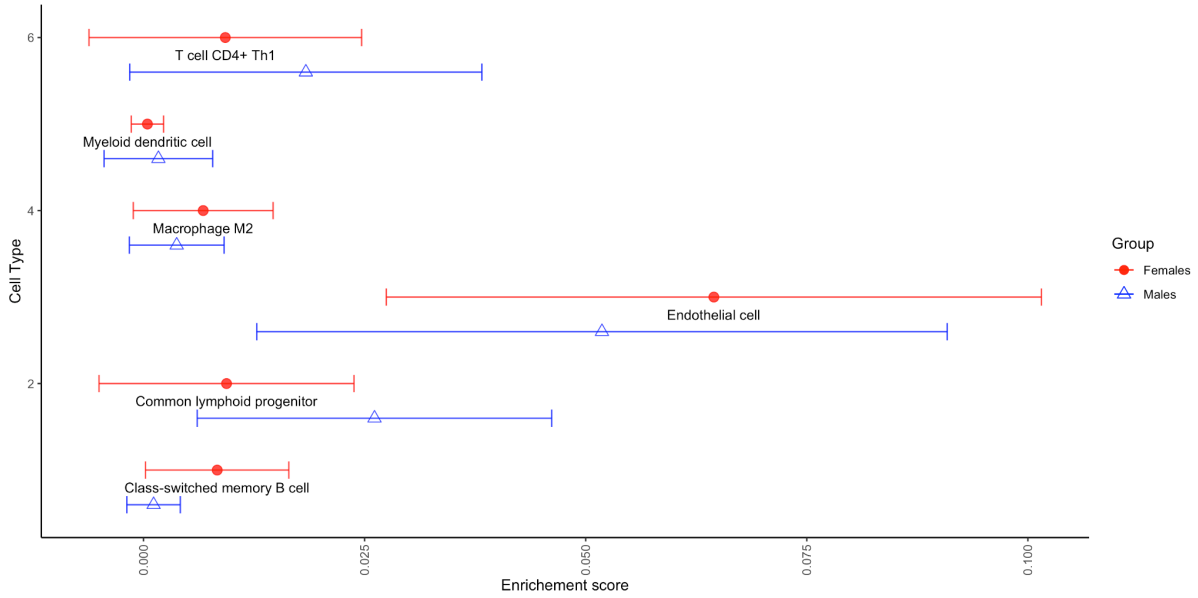
	TCGA Males stage I-II - adjusted for age and stage	
Genes	beta coefficient	raw p-value
UHRF1	0.08332	0.89885
HLA-E	-0.31236	0.56068
UBA7	-0.20154	0.46811
NGFRAP1 (BEX3)	-0.12116	0.59936
SF3B3	-1.38346	0.26275



Appendix Figure 2: estimate proportion of cancer cells performed on sex and stage subgroups. ns = not significant



Appendix Figure 3: Significantly differently enriched immune cell subtypes between females and males of GTEx Sun Exposed normal skin .



Appendix Figure 4: Significantly differently enriched immune cell subtypes between females and males of GTEx Not Sun Exposed normal skin .

Appendix Table 3: Functions obtained from GeneMANIA network analysis based on female-specific genes UHRF1, HLA-E and UBA7. The reported network functions have FRD < 0.05 and refer to more than 4 genes in the network.

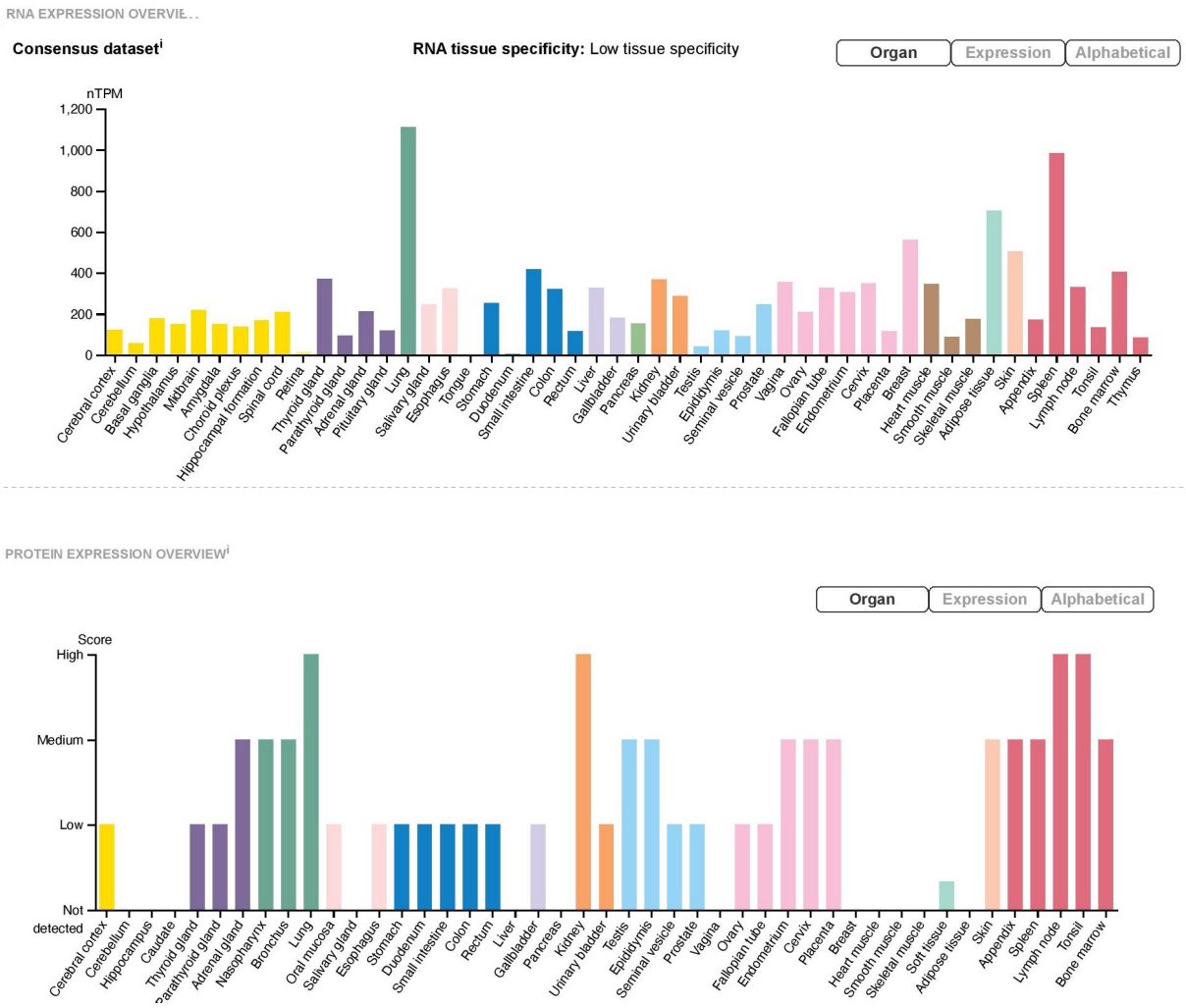
Network function (UHRF1, HLA-E, UBA7)	FDR	Genes in network	Genes in genome
endoplasmic reticulum membrane	2.7294E-05	7	268
antigen binding	4.67E-08	6	36
integral component of endoplasmic reticulum membrane	2.41E-07	6	52
phagocytic vesicle	4.82E-07	6	62
antigen processing and presentation of exogenous peptide antigen via MHC class I	1.0169E-06	6	76
endocytic vesicle membrane	1.2534E-06	6	81
antigen processing and presentation of peptide antigen via MHC class I	1.5506E-06	6	86
intrinsic component of endoplasmic reticulum membrane	2.7294E-05	6	151
lymphocyte mediated immunity	5.4522E-05	6	176
antigen processing and presentation of exogenous peptide antigen	5.4522E-05	6	176
antigen processing and presentation of exogenous antigen	0.00006039	6	181
antigen processing and presentation of peptide antigen	0.00007163	6	190
antigen processing and presentation	0.00011026	6	208
endocytic vesicle	0.00011429	6	211
negative regulation of cytokine production	0.00013615	6	221
MHC protein binding	5.32E-07	5	27
regulation of type I interferon production	0.00011433	5	106
type I interferon production	0.00015132	5	115
ubiquitin-like protein ligase binding	0.00711129	5	266

Appendix Table 4: Functions obtained from GeneMANIA network analysis based on female-specific genes BEX3 and SF3B3. The reported network functions have FRD < 0.05 and refer to more than 4 genes in the network

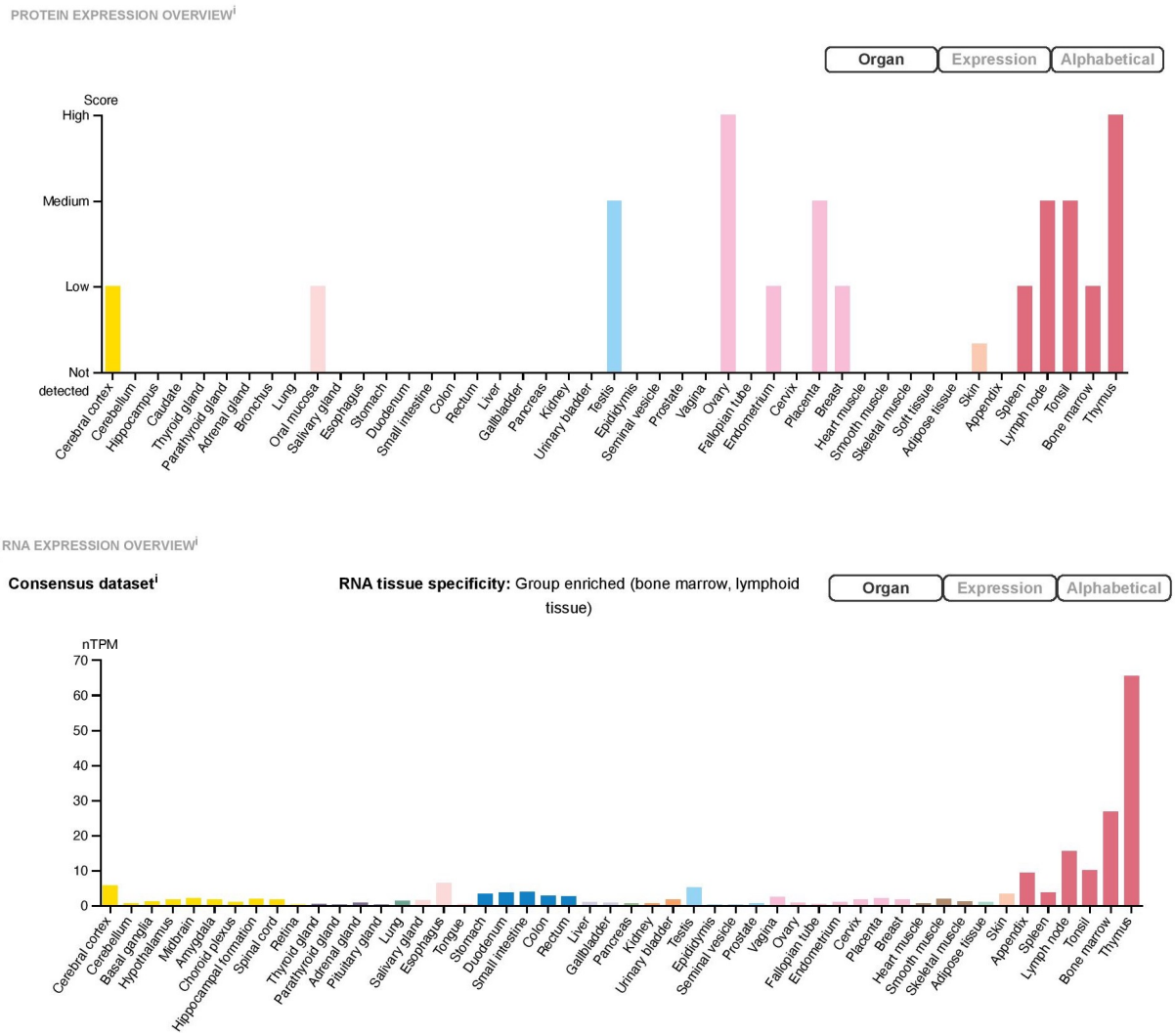
Network Function (SF3B3, BEX3)	FDR	Genes in network	Genes in genome
U2-type spliceosomal complex	3.75E-08	7	86
spliceosomal complex	3.75E-08	8	145
protein export from nucleus	9.137E-05	6	154
regulation of cysteine-type endopeptidase activity involved in apoptotic process	9.5595E-05	6	169
nuclear export	0.00011338	6	184
positive regulation of cysteine-type endopeptidase activity involved in apoptotic process	0.00011338	5	86
regulation of cysteine-type endopeptidase activity	0.00014013	6	195
positive regulation of cysteine-type endopeptidase activity	0.00044006	5	127
ribonucleoprotein complex export from nucleus	0.00044006	5	129
ribonucleoprotein complex localization	0.00044006	5	129
RNA export from nucleus	0.00053323	5	137
negative regulation of hydrolase activity	0.00053323	6	272
regulation of endopeptidase activity	0.00053323	6	270
positive regulation of endopeptidase activity	0.0005751	5	146
extrinsic apoptotic signalling pathway	0.0005751	5	146
nucleocytoplasmic transport	0.00060856	6	287
nuclear transport	0.00061379	6	291
regulation of peptidase activity	0.00063263	6	299
RNA transport	0.00063263	5	156
positive regulation of peptidase activity	0.00064877	5	159
nucleic acid transport	0.00068198	5	162
establishment of RNA localization	0.00071687	5	165
nucleobase-containing compound transport	0.00151175	5	195
RNA localization	0.00157123	5	198
protein-containing complex localization	0.00421087	5	244
positive regulation of proteolysis	0.00974265	5	298

Appendix Table 5: This table shows the -ΔCT values obtained from the rt-qPCR of the five sex specific prognostic biomarkers.

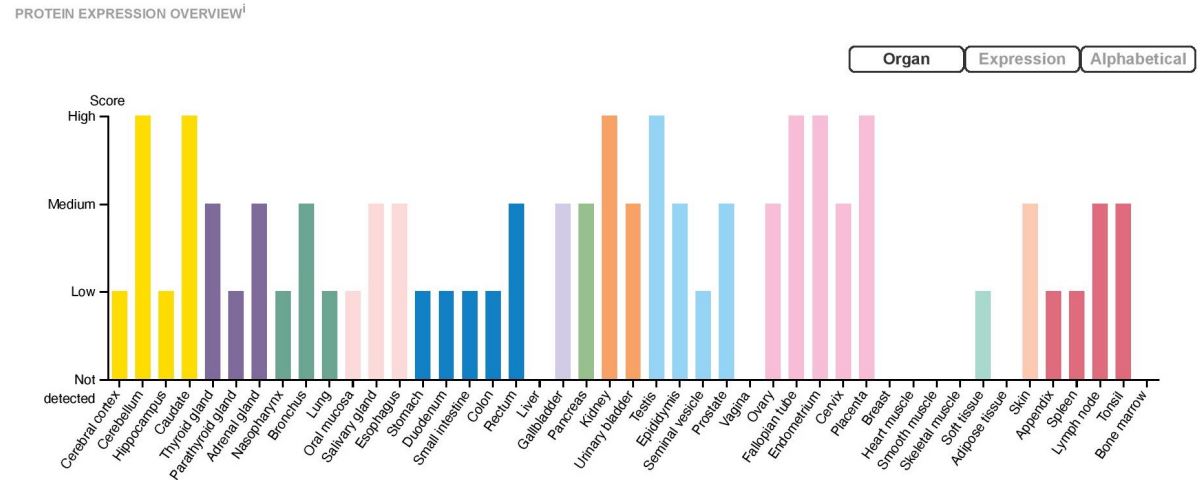
						BEX3	**HLA-E**	SF3B3	UBA7	**UHRF1**
Cq Mean	stage	age	status	time	sex	ilmn_2370091	**ilmn_1765258**	ilmn_1803110	ilmn_1794612	**ilmn_1786065**
SAMPLE_13	2	79	0	8	1	-3.45	-4.03	-4.66	-6.40	-6.69
SAMPLE_14	2	74	1	6	1	-2.87	-5.50	-3.43	-5.92	-6.50
SAMPLE_18	1	72	0	8	1	-4.40	-3.61	-4.52	-6.02	-7.11
SAMPLE_21	2	72	0	9	1	-3.35	-5.26	-3.89	-5.83	-7.31
SAMPLE_22	2	46	0	9	1	-3.27	-9.40	-3.24	-5.70	-5.52
SAMPLE_33	2	80	1	1	1	-3.35	-5.26	-4.52	-8.29	-6.23
SAMPLE_35	1	84	0	8	1	-3.76	-4.63	-3.77	-6.13	-6.09
SAMPLE_34	1	36	0	6	1	-3.49	-8.97	-3.30	-5.66	-6.15
SAMPLE_40	2	74	1	3	1	-2.50	-8.30	-3.01	-7.51	-5.38
SAMPLE_43	2	55	1	8	1	-4.00	-7.20	-4.15	-6.13	-6.55
SAMPLE_16	1	40	0	7	1	-3.73	-6.26	-2.70	-5.37	-9.28
SAMPLE_20	2	63	0	8	1	-4.32	-5.76	-4.11	-6.53	-7.19
SAMPLE_23	2	52	0	7	1	-2.98	-2.46	-3.64	-6.12	-7.80
SAMPLE_24	1	77	0	9	1	-5.04	-4.39	-3.33	-6.09	-7.05
SAMPLE_32	1	49	0	6	1	-2.52	-5.03	-2.51	-6.75	-6.92
SAMPLE_48	2	52	0	8	1	-3.03	-7.48	-3.18	-6.35	-5.31
SAMPLE_6	2	65	0	10	1	-2.20	-4.71	-3.79	-6.70	-6.00
SAMPLE_10	1	67	0	9	1	-3.98	-4.10	-4.28	-4.71	-6.79
SAMPLE_31	1	45	0	10	1	-4.78	-7.70	-3.53	-5.15	-6.03
SAMPLE_47	2	28	0	8	1	-2.91	-4.96	-3.65	-6.22	-6.61
SAMPLE_7	2	95	1	1	1	-3.00	-4.77	-4.11	-8.49	-6.33
SAMPLE_8	2	74	0	8	1	-2.39	-3.37	-3.82	-6.95	-5.81
SAMPLE_1	2	76	1	4	0	-3.26	-3.77	-4.49	-9.01	-6.94
SAMPLE_12	3	84	1	3	0	-3.31	-7.57	-2.81	-5.61	-5.92
SAMPLE_2	3	71	0	11	0	-2.98	-6.14	-3.52	-7.04	-5.29
SAMPLE_9	2	48	0	10	0	-4.10	-4.69	-4.27	-6.34	-8.99
SAMPLE_26	1	64	0	9	0	-3.48	-2.89	-4.08	-4.41	-7.37
SAMPLE_30	2	74	0	10	0	-2.93	-8.38	-3.14	-5.70	-6.80
SAMPLE_37	2	65	0	8	0	-4.54	-3.83	-4.62	-6.36	-7.12
SAMPLE_44	1	52	0	9	0	-3.83	-7.03	-3.06	-6.59	-6.22
SAMPLE_3	1	51	0	10	0	-5.60	-9.51	-4.52	-3.55	-6.13
SAMPLE_15	2	58	0	9	0	-3.49	-9.00	-2.51	-6.09	-6.82
SAMPLE_17	1	70	0	9	0	-3.37	-7.94	-2.03	-5.09	-6.05
SAMPLE_19	1	60	0	8	0	-4.46	-4.25	-4.20	-5.31	-8.22
SAMPLE_25	1	62	0	10	0	-2.76	-5.61	-2.81	-5.70	-5.79
SAMPLE_27	1	51	0	10	0	-4.57	-9.45	-3.60	-5.55	-7.05
SAMPLE_29	1	58	0	7	0	-4.57	-4.16	-3.68	-5.69	-7.26
SAMPLE_36	1	62	0	9	0	-2.64	-4.60	-3.20	-5.65	-8.14
SAMPLE_41	1	45	0	10	0	-4.00	-8.90	-2.66	-6.29	-6.25
SAMPLE_4	2	77	1	5	0	-3.19	-3.37	-2.90	-6.03	-5.41
SAMPLE_42	1	80	1	6	0	-4.06	-6.16	-2.99	-5.08	-7.29
SAMPLE_5	2	65	0	7	0	-3.62	-3.28	-4.58	-6.04	-7.68
SAMPLE_38	2	62	0	11	0	-3.58	-5.02	-4.46	-6.31	-6.94

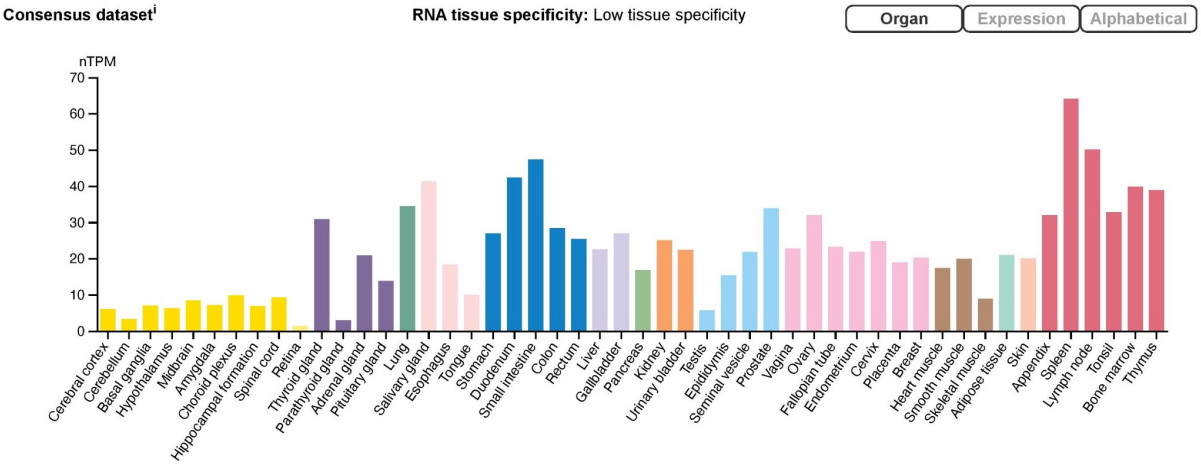


Appendix Figure 5: Protein and RNA expression of HLA-E across different tissues, obtained from Protein Atlas.

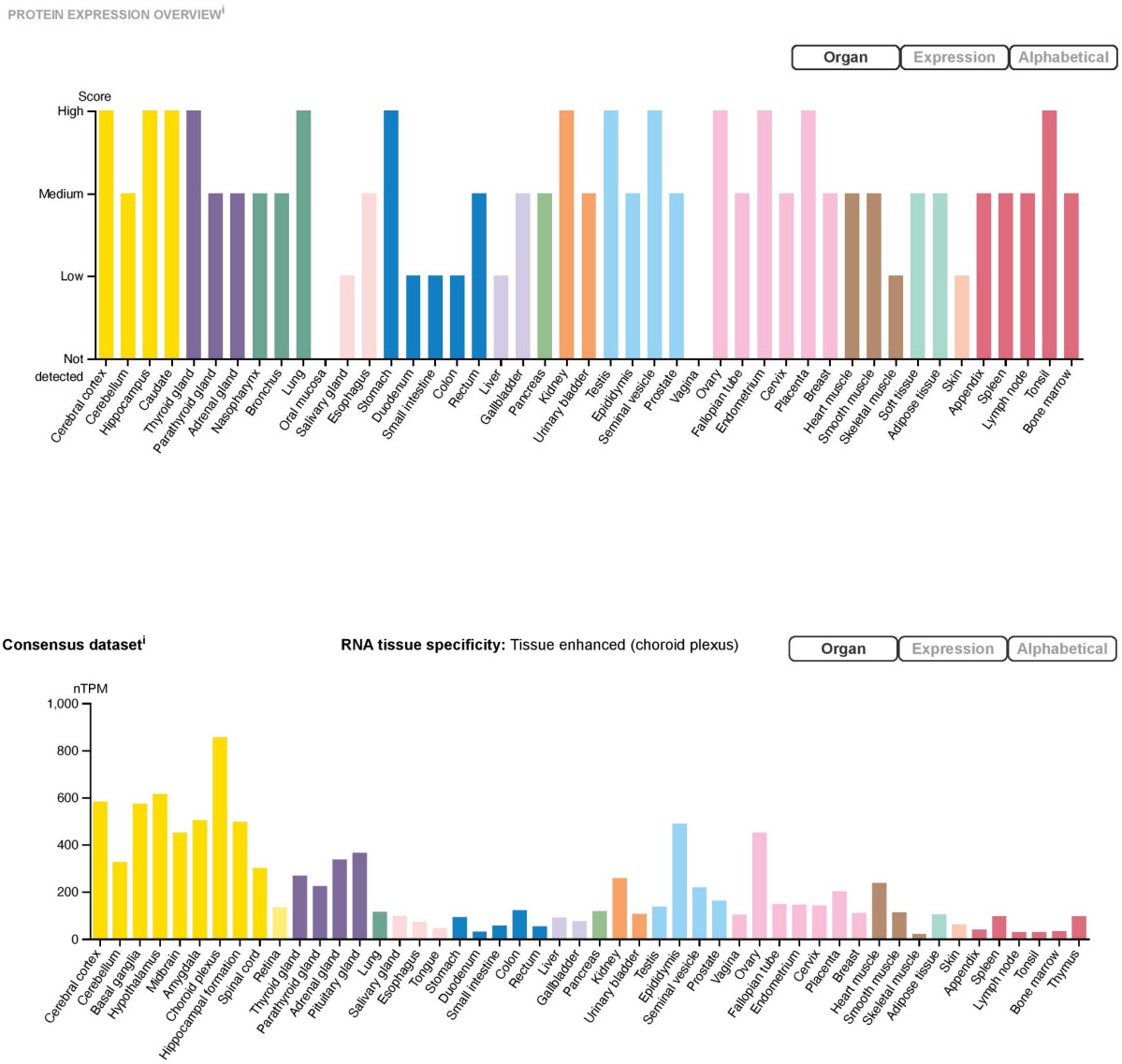


Appendix Figure 6: Protein and RNA expression of UHRF1 across different tissues, obtained from Protein Atlas.

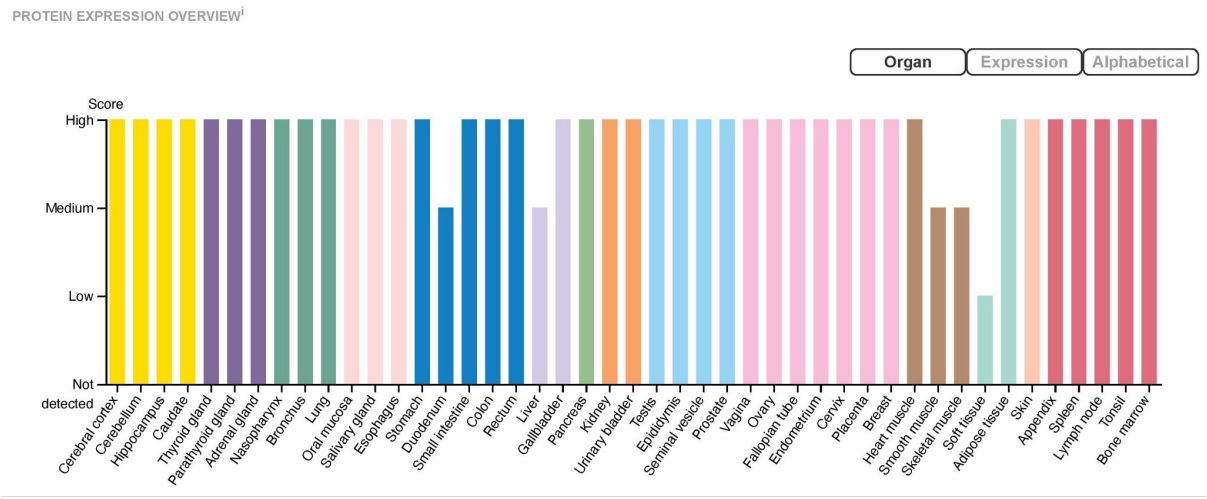


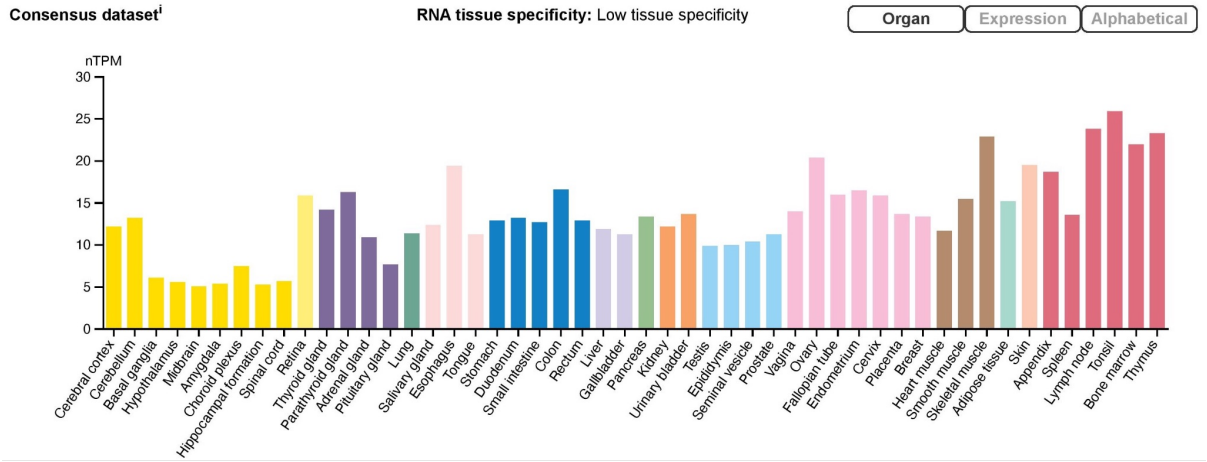


Appendix Figure 7: Protein and RNA expression of UBA7 across different tissues, obtained from Protein Atlas.



Appendix Figure 8: Protein and RNA expression of BEX3 across different tissues, obtained from Protein Atlas.





Appendix Figure 9: Protein and RNA expression of SF3B3 across different tissues, obtained from Protein Atlas.

Appendix Table 6: This table presents sex-specific multivariable Cox regression results separately in female and male stage III melanomas, including beta coefficients (Coef), hazard ratios (HR), and raw p-values. The third column, labelled pro-tumour(+) and anti-tumour(-), denotes the reported role of each biomarker in existing literature. The 'OGS' column corresponds to the official gene symbol, while the 'Function' column specifies the identified functional involvement of each gene.

			Sex specific multivariable Cox regression analysis adjusting for age and stage					
			Females stage III			**Males stage III**		
OGS	Function	pro-tumour (+), anti-tumour(-)	Coef	HR	raw_p-val	Coef	HR	raw_pval
CKS2	cell cycle	+	0.2704	1.3105	0.1351	0.8391	2.3143	0.0108
CDC2	cell cycle	+	0.1931	1.2130	0.3570	0.3256	1.3848	0.1954
CCNB1	cell cycle	+	0.2664	1.3052	0.1467	0.2374	1.2680	0.4192
CENPF	cell cycle	+	0.2412	1.2727	0.1293	0.5802	1.7865	0.0104
DHFR	cell cycle	+	0.3529	1.4231	0.2625	1.1914	3.2918	0.0221
HCAP-G	mitosis	+	0.3497	1.4187	0.0734	0.2649	1.3033	0.2966
STK6	mitosis	+	0.2666	1.3055	0.0962	0.5764	1.7796	0.0227
BUB1	mitotic spindle checkpoint inhibition	+	0.3980	1.4889	0.0179	-0.0307	0.9698	0.8334
BIRC5	apoptosis	+	0.2742	1.3155	0.1054	0.2625	1.3002	0.3434
P2RY14	stimulation of apoptosis	-	-0.2125	0.8086	0.2725	-0.0314	0.9691	0.8711
TOP2A	DNA replication	+	0.3175	1.3737	0.0802	0.4022	1.4951	0.1402
RRM2	DNA replication	+	0.7076	2.0292	0.0042	0.0439	1.0449	0.8457
TYMS	DNA replication	+	0.5208	1.6834	0.0183	0.0236	1.0239	0.9506
PCNA	DNA replication	+	0.5358	1.7088	0.0227	0.1143	1.1211	0.7375
MCM4	DNA replication	+	0.3424	1.4083	0.1539	0.3876	1.4734	0.1136

MCM6	DNA replication	+	0.7758	2.1723	0.1400	1.4162	4.1215	0.0251
GLRX2	stress response	+	0.3457	1.4130	0.4086	-0.0775	0.9254	0.8376
DNAJA1	stress response	+	0.2292	1.2576	0.5113	0.3263	1.3859	0.5529
HSPA4	stress response	+	-0.1208	0.8862	0.3790	0.6965	2.0068	0.0936
HSPA5	stress response	+	-1.2135	0.2972	0.0731	-0.0337	0.9669	0.9413
HSPD1	stress response	+	-0.1193	0.8876	0.7638	0.8269	2.2862	0.2305
TXNIP	stress response	-	-0.4128	0.6618	0.2282	-0.5455	0.5796	0.1868
CACYBP	ubiquitin cycle	+	0.3931	1.4816	0.1819	-0.1477	0.8626	0.6866
CNN3	Aactin and calmodulin binding	+	-0.0574	0.9442	0.7720	0.0746	1.0775	0.8443
STMN2	intracellular signalling	-	-0.1726	0.8414	0.1072	-0.2458	0.7821	0.0201