Genetic Predisposition to Sporadic and Familial Multiple Myeloma

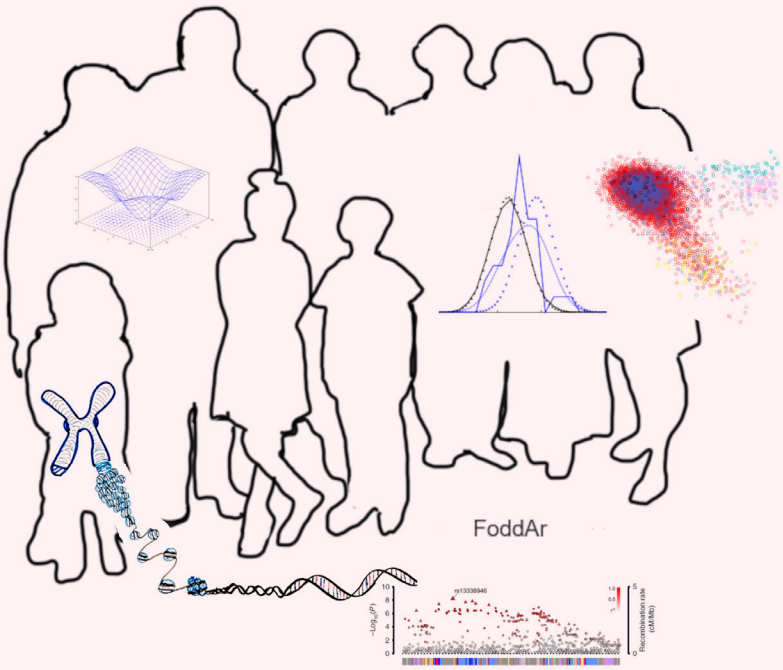
Halvarsson, Britt-Marie

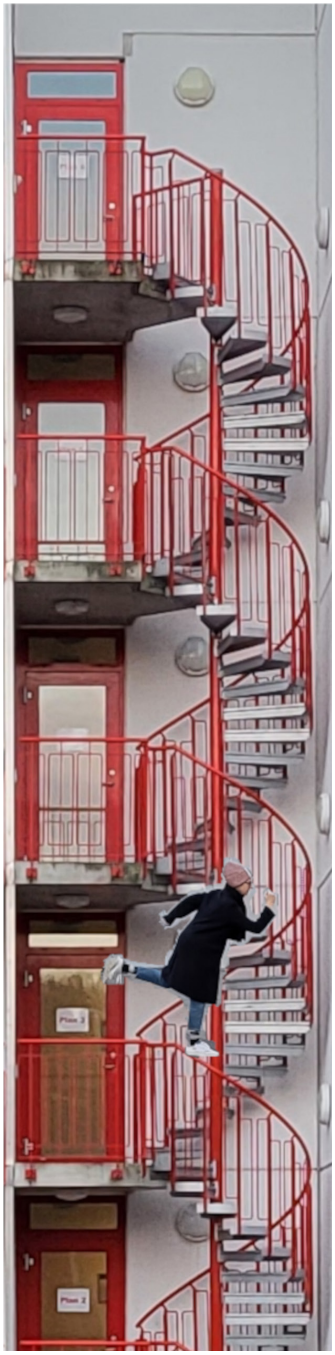
2022

Document Version:
Publisher's PDF, also known as Version of record

Genetic Predisposition to Sporadic and Familial Multiple Myeloma

BRITT-MARIE HALVARSSON
DEPARTMENT OF LABORATORY MEDICINE | FACULTY OF MEDICINE | LUND UNIVERSITY





Printed by Media-Tryck, Lund 2022 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

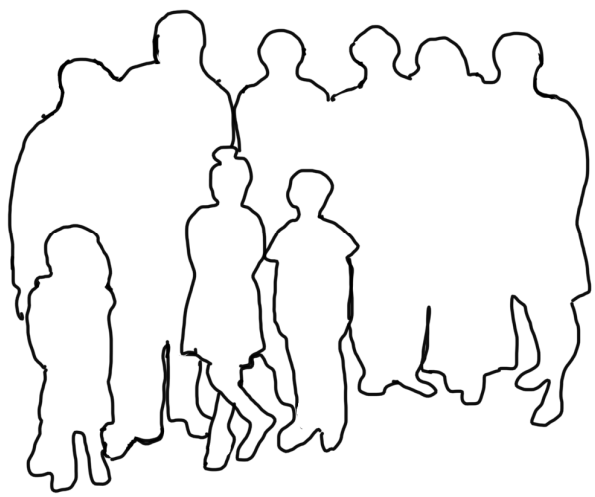
Division of Hematology and Transfusion Medicine
Department of Laboratory Medicine

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2022:1
ISBN 978-91-8021-162-8
ISSN 1652-8220

Genetic Predisposition

to

Sporadic and Familial Multiple Myeloma



Genetic Predisposition to Sporadic and Familial Multiple Myeloma

Britt-Marie Halvarsson

LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended in Segerfalksalen, Biomedicinskt Centrum, Lund

on 7 January 2022 at 13:00

Faculty opponent

Lasse Sommer Kristensen

Associate Professor, Department of Biomedicine, Aarhus University

Organization LUND UNIVERSITY Faculty of Medicine Department of Laboratory Medicine	**Document name** DOCTORAL DISSERTATION **Date of issue** 2022-01-07
Author: Britt-Marie Halvarsson	Sponsoring organization

Title and subtitle
Genetic Predisposition to Sporadic and Familial Multiple Myeloma

Abstract

Multiple Myeloma (MM) is the second most common hematological malignancy. It is defined by an uncontrolled growth of plasma cells, usually in the bone marrow. Clinically it is complicated by hypercalcemia, renal failure, anaemia, and bone pain. Although recent advances in the treatment have extended survival and quality-of-life considerably, MM remains a fatal disease.

Since the 1920's MM has been reported to aggregate in families (famililial MM). In first-degree relatives of MM patients there is a two to four-fold increased risk in developing MM, pointing at a possible inherited genetic aetiology in at least a subset of MM patients

The overall aim of this thesis is to identify germline DNA sequence variants that predispose for Multiple Myeloma (MM), and it is based on four Papers.

In Paper I, II and IV, we performed case-control genome-wide association studies (GWASs) to identify germline single nucleotide polymorphisms (SNPs) and small insertions/deletions (INDELs) that associate with MM risk. In Paper I, we identified a novel significant association with *ELL2*, and a border-line suggestive association with *TOM1*.

In Paper II and IV we collaborated internationally in GWAS meta-analyses and identified eight and six variants, respectively, in or near the genes *JARID2* (at 6p22.3), *ATG5* (6q21), *SMARCD3*, (7q36.1), *CCAT1* (8q24.21), *CDKN2A* (9p21.3), *WAC* (10p12.1), *RFWD3* (16q23.1), *PREX1* (20q13.13), *CEP120* (5q23.2), *POT1* (7q31.33), *CCDC71L* (7q22.3), *SP3* (2q31.1), *KLF2* (19p13.11) and *PRR14*/*RNF40* (16p11.2). The *TOM1* variant in Paper I replicated in these studies. The identified MM risk loci is estimated to explain a 20% of MM heritability.

In Paper III, we performed SNP microarray and whole-exome sequencing analysis on 38 cases of familial MM. Constructing polygenic risk scores, we found direct evidence for a polygenic aetiology in familial MM, and estimated that about one-third of familial MM cases were associated with an enrichment of common risk variants identified by GWAS. In Paper IV, we extended our polygenic risk scores with newly identified risk variants, and again observed an enrichment of risk variants in familial cases.

Key words
multiple myeloma, genetics, inherited predisposition, familial multiple myeloma, GWAS, polygenic risk score.

Classification system and/or index terms (if any)

Supplementary bibliographical information	**Language** English
ISSN 1652-8220	**ISBN** 978-91-8021-162-8
Recipient's notes	**Number of pages** 99 Price
	Security classification

Genetic Predisposition to Sporadic and Familial Multiple Myeloma

Britt-Marie Halvarsson

LUND
UNIVERSITY

Lund University
Faculty of Medicine
Department of Laboratory Medicine

ISBN 978-91-8021-162-8
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2022

To my offspring

Sorry about that genome thing,
but you really should blame my partner in crime too*

* The suspects in this statement contributed equally to this work

Table of Contents

Preface

“Still, life had a way of adding day by day”

Virginia Woolf, Mrs Dalloway

“Do you really want to live forever?” Alphaville asked in their 1984 song *Forever Young*. Since the answer to this question is yes, both in human beings and malignant plasma cells, this thesis aims to characterize “genetic predisposition to sporadic and familial Multiple Myeloma”. The exact meaning of this concept will be described in the following six chapters.

This thesis is meant to reflect my research during my doctoral studies put in a perspective of current knowledge in the research field.

Sigrid, August, and Selma, this is what I have been working on the past few years. I know you are disappointed that it is written in English, but yes, you are more than welcome to colour the pages as you see fit.

More adult readers will first find an introduction, an optimistic attempt to establish a basic common ground by describing some common genetic concepts. Confident readers are welcome to skip this part.

To more serious readers, I will try to hold back the jokes.

Now, without further ado, let us get to the point.

Britt-Marie Halvarsson

Lund, 22 November 2021

List of Papers

This thesis is based on the following publications or manuscripts. These papers will henceforth be referred to by their roman numerals (I-IV).

I. **Variants in ELL2 influencing immunoglobulin levels associate with multiple myeloma**

Swaminathan B*, Thorleifsson G*, Jöud M*, Ali M*, Johnsson E, Ajore R, Sulem P, **Halvarsson BM**, Eyjolfsson G, Haraldsdottir V, Hultman C, Ingelsson E, Kristinsson SY, Kähler AK, Lenhoff S, Masson G, Mellqvist UH, Månsson R, Nelander S, Olafsson I, Sigurðardottir O, Steingrimsdóttir H, Vangsted A, Vogel U, Waage A, Nahi H, Gudbjartsson DF, Rafnar T, Turesson I, Gullberg U, Stefánsson K, Hansson M, Thorsteinsdóttir U, Nilsson B.

Nature Communications, (2015). 6, 7213

II. **Genome-wide association study identifies multiple susceptibility loci for multiple myeloma**

Mitchell JS*, Li N*, Weinhold N*, Försti A*, Ali M*, van Duin M*, Thorleifsson G, Johnson DC, Chen B, **Halvarsson BM**, Gudbjartsson DF, Kuiper R, Stephens OW, Bertsch U, Broderick P, Campo C, Einsele H, Gregory WA, Gullberg U, Henrion M, Hillengass J, Hoffmann P, Jackson GH, Johnsson E, Jöud M, Kristinsson SY, Lenhoff S, Lenive O, Mellqvist UH, Migliorini G, Nahi H, Nelander S, Nickel J, Nöthen MM, Rafnar T, Ross FM, da Silva Filho MI, Swaminathan B, Thomsen H, Turesson I, Vangsted A, Vogel U, Waage A, Walker BA, Wihlborg AK, Broyl A, Davies FE, Thorsteinsdottir U, Langer C, Hansson M, Kaiser M, Sonneveld P, Stefansson K, Morgan GJ, Goldschmidt H, Hemminki K, Nilsson B, Houlston RS.

Nature Communications, (2016). 7, 12050

III. **Direct evidence for a polygenic etiology in familial multiple myeloma**

Halvarsson BM, Wihlborg AK, Ali M, Lemonakis K, Johnsson E, Niroula A, Cibulskis C, Weinhold N, Försti A, Alici E, Langer C,

Pfreundschuh M, Goldschmidt H, Mellqvist UH, Turesson I, Waage A, Hemminki K, Golub T, Nahi H, Gullberg U, Hansson M, Nilsson B.

Blood Advances, (2017). 1(10), 619-623

IV. **Identification of multiple risk loci and regulatory mechanisms influencing susceptibility to multiple myeloma**

Went M*, Sud A*, Försti A*, **Halvarsson BM***, Weinhold N, Kimber S, van Duin M, Thorleifsson G, Holroyd A, Johnson DC, Li N, Orlando G, Law PJ, Ali M, Chen B, Mitchell JS, Gudbjartsson DF, Kuiper R, Stephens OW, Bertsch U, Broderick P, Campo C, Bandapalli OR, Einsele H, Gregory WA, Gullberg U, Hillengass J, Hoffmann P, Jackson GH, Jöckel KH, Johnsson E, Kristinsson SY, Mellqvist UH, Nahi H, Easton D, Pharoah P, Dunning A, Peto J, Canzian F, Swerdlow A, Eeles RA, Kote-Jarai Z, Muir K, Pashayan N, Nickel J, Nöthen MM, Rafnar T, Ross FM, da Silva Filho MI, Thomsen H, Turesson I, Vangsted A, Andersen NF, Waage A, Walker BA, Wihlborg AK, Broyl A, Davies FE, Thorsteinsdottir U, Langer C, Hansson M, Goldschmidt H, Kaiser M, Sonneveld P, Stefansson K, Morgan GJ, Hemminki K, Nilsson B, Houlston RS; PRACTICAL consortium.

Nature Communications (2018). 9(1), 3707

Abbreviations

CNV	Copy number variation
CSR	Class-switch recombination
DSB	Double-strand break
eQTL	Expression quantitative trait locus
fMM	Familial Multiple Myeloma
GWAS	Genome-wide association study
HSC	Haematopoietic stem cell
HWE	Hardy-Weinberg Equilibrium
Ig	Immunoglobulin
LD	Linkage disequilibrium
MAF	Minor allele frequency
MDS	Multidimensional scaling
MM	Multiple Myeloma
MGUS	Monoclonal gammopathy of unknown significance
PCA	Principal component analysis
PRS	Polygenic risk score
RAF	Risk allele frequency
SHM	Somatic hypermutation
SNP	Single-nucleotide polymorphism
TAD	Topologically associating domain
TF	Transcription factor
WES	Whole-exome sequencing

Popular Introduction to Genetics

"Omnis cellula e cellula"

"All cells come from cells"
Rudolf Virchow, 1858

The human body consists of approximately 3.0 x 10^{13} human cells, predominantly red blood cells (~84%) (Sender, Fuchs, & Milo, 2016). Only ten percent of human cells are nucleated (Sender et al., 2016). These nuclei harbour our genetic material, *the genome*, organised in chromosomes (Mohanta, Mishra, & Al-Harrasi, 2021). A *chromosome* consists of a long double-stranded *deoxyribonucleic acid* (DNA) that is wound up on spools called *nucleosomes* (Fig. 1) (Mohanta et al., 2021; Widom, 1998). Nucleosomes contain subunits, *histones*, and are connected to each other like beads on a string, and densely packed (Widom, 1998).

Most of us have 46 chromosomes in each cell (diploid cells), 22 *homologous* pairs of autosomal chromosomes, containing one copy each of the same genes, and two sex chromosomes, X and Y. One chromosome in each pair originates from the biological mother and the biological father, respectively. In a two-step process called *meiosis*, one germ cell with 46 chromosomes is divided into cells with 23 chromosomes each (*haploid* cells) (Fig. 2) (Ohkura, 2015). Meiosis in females is asymmetric and produces only one oocyte and three polar bodies. However, in males, four sperm cells are produced (Coop & Przeworski, 2007). After mating, the oocyte and a sperm fuse and a *zygote* is produced, now containing 46 chromosomes, a "master" stem cell (*totipotent stem cell*) from which all human cells in the body are created (Rebuzzini, Zuccotti, & Garagna, 2021). In meiosis I, homologous chromosomes cross over and exchange genetic material (Fig. 2). This process is called *homologous recombination* and further contributes to the genetic variation in offspring genomes (Baudat, Imai, & de Massy, 2013; Coop & Przeworski, 2007).

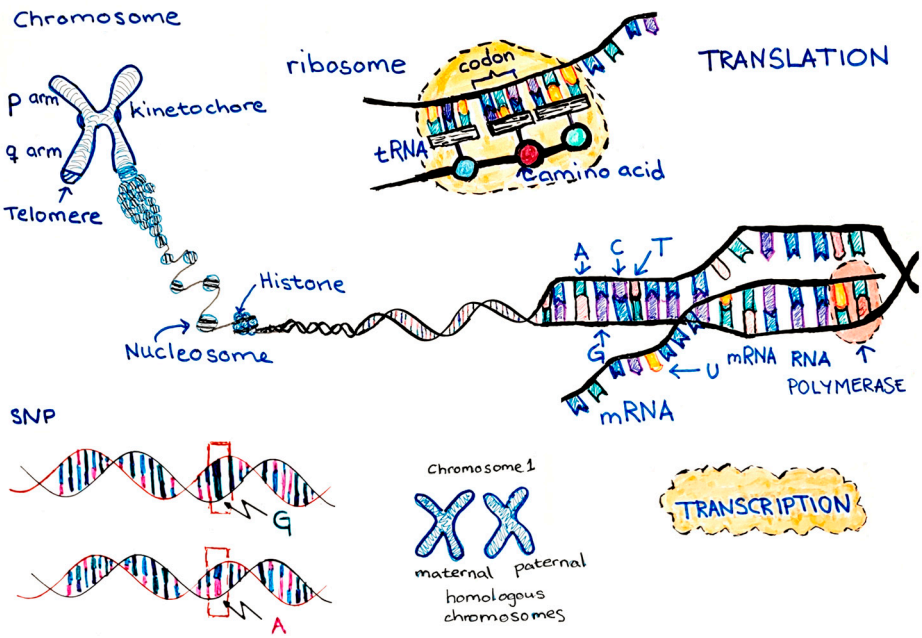


Figure 1. Schematic view over chromosomes, DNA, transcription, translation, and single-nucleotide polymorphisms.

Our genome contains approximately 20,000 protein-coding genes. Each of these encode one or several proteins necessary for tissue formation and homeostasis sustaining human life (Pertea et al., 2018). Each gene is a part of a long strand of four different bases: the purines adenine (A) and guanine (G) and the pyrimidines thymine (T) and cystine (C), that together with a sugar-phosphate backbone form nucleotides, the building blocks of DNA (Watson & Crick, 1953). The two strands of DNA are entwined and held together by the binding of bases in pairs, A to T, and G to C, respectively, forming a helix (Fig. 1) (Watson & Crick, 1953). In the protein coding parts of a gene (*exons*), triplets of nucleotides form *codons* (Crick, Barnett, Brenner, & Watts-Tobin, 1961; Gilbert, 1978). Codons encode amino acids that are identically ordered as their codons, and later assemble to form proteins. Codons can also signal, for example, the start and stop-sites in the gene. The translation of genes into protein is a process in several steps; in *transcription* a one string-copy of the gene is transcribed called messenger ribonucleic acid, *mRNA*, that subsequently is translated into amino acids through transfer RNA, *tRNA* (Hoerter & Ellis, 2021). RNAs are structured as single-stranded sequences of nucleotides, much like single strands of DNA, except thymine is substituted with the pyrimidine, uracil (U) and the backbone is a ribose instead of deoxyribose (Fig. 1) (Markham & Smith, 1951).

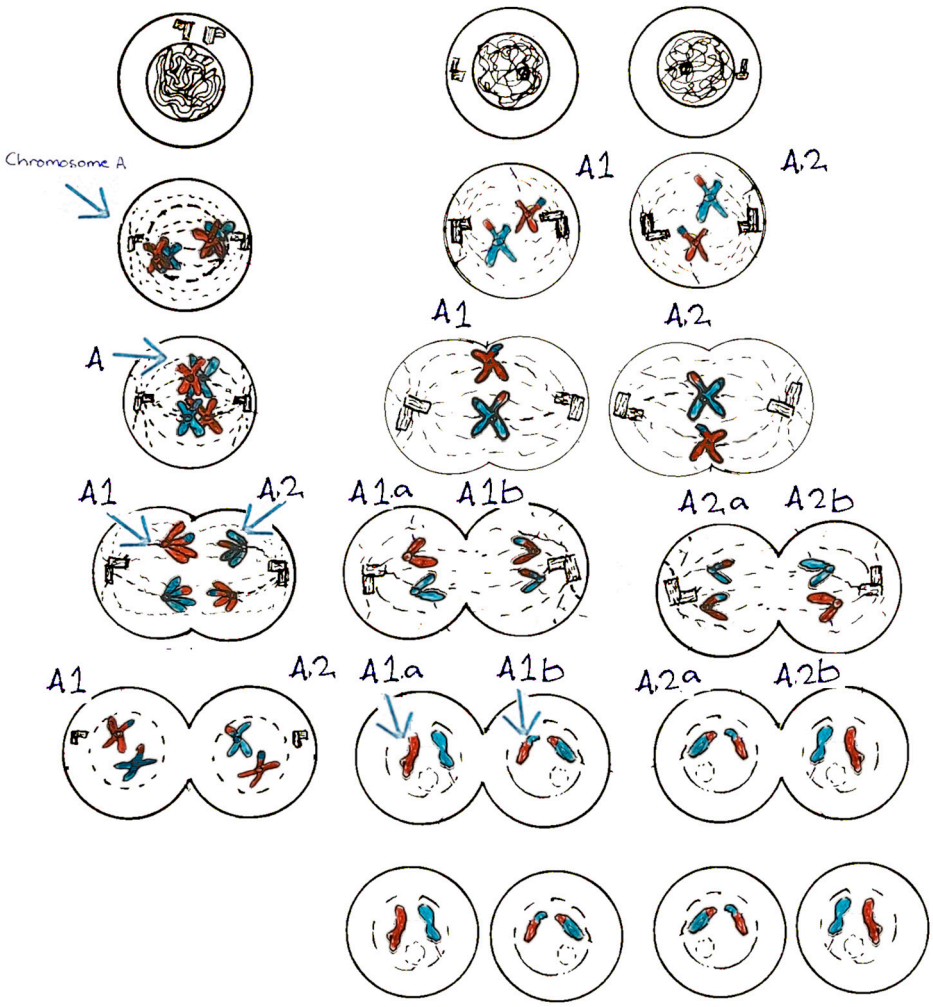


Figure 2. Schematic view over meiosis, the production of gametes. In meiosis I (left panel), homologous chromosomes first pair up and exchange fragments by crossing over. Homologous chromosomes then separate to opposite ends of the cell. The diploid cell divides to form two haploid daughter cells. In meiosis II (right panel), the two haploid cells divide to form four daughter cells, the gametes. Female meiosis is assymetric, forming one oocyte, and three polar bodies, that later undergo apoptosis.

Seldomly, there are "spelling mistakes" in the DNA sequence, "written" during DNA *replication* in meiosis. These small mutations create additional variation in our inherited DNA. In single-nucleotide polymorphisms, *SNPs,* one base is exchanged for another (Fig. 1), while an *INDEL* is either a small insertion or a deletion of one or more base pairs. These variants have emerged countless generations ago, and can therefore impact variation of common traits, including for

example hair colour. In exons, there are two main types of SNPs, *synonymous* and *non-synonymous* SNPs, respectively (Hoerter & Ellis, 2021; Shastry, 2002). At synonymous SNP sites, the base substitution results in an identical amino acid and consequently is not believed to alter the protein structure. The non-synonymous SNP is either *missense*, resulting in an amino acid change, or *non-sense* where it encodes a stop-codon (*stop-gain*) or inactivates one (*stop-loss*). If the number of nucleotides in an INDEL is not divisible by three, the genetic code will be shifted and the nucleotide triplets following the INDEL, will encode the wrong amino acids (a *frameshift* mutation). Intuitively, non-sense SNPs and frame-shift mutations will have a larger impact on the amino acid sequence than missense variants and are therefore more likely to produce proteins with altered function. There are also non-coding SNPs between (*intergenic*) or inside genes (*intronic*) (Shastry, 2002).

Understanding the functional consequences of individual variants is challenging, and their potentially damaging effect is often predicted computationally using genomic annotations (Niroula, Urolagin, & Vihinen, 2015).

The protein-coding sequences are dispersed in our DNA and are separated by DNA sequences of different lengths. While the functions of non-coding sequences are largely unknown, a large portion of intergenic DNA might have regulatory functions, regulating when and to what extent a protein-coding gene should be expressed. For example, *transcription factors* (TFs) bind to certain DNA sequences and regulate gene transcription (Spitz & Furlong, 2012). In DNA, between protein-encoding genes you can also find genes that does not encode proteins but different kinds of RNA (Esteller, 2011). These may also have regulatory functions of genes or gene products. Changes in the genome that do not alter the DNA sequence, are called *epigenetic*. Conformational changes can include making DNA accessible for transcription by unwinding it from nucleosomes (Feinberg & Tycko, 2004).

When cells become worn-out or damaged, they undergo *apoptosis*, programmed cell death, killing billions of cells every day (Newbold, Martin, Cullinane, & Bots, 2014). Consequently, these cells must be replaced, along with cells that have been shed, lysed, or killed by other cells. Similar to meiosis II, a mother cell duplicates her cellular material and divides into two identical daughter cells in a process called *mitosis* (Ohkura, 2015).

To balance the number replicating cells and check that daughter cells are properly formed, the entrance and exit into mitosis is tightly regulated in the *cell cycle* (Fig. 3) (Barnum & O'Connell, 2014). Before progressing to the next phase there is a *checkpoint*, ensuring if progression to the next phase is desired, and if so, that all processes in the current phase have been carried out satisfactorily. Likewise, mitosis must also be tightly regulated in its different phases.

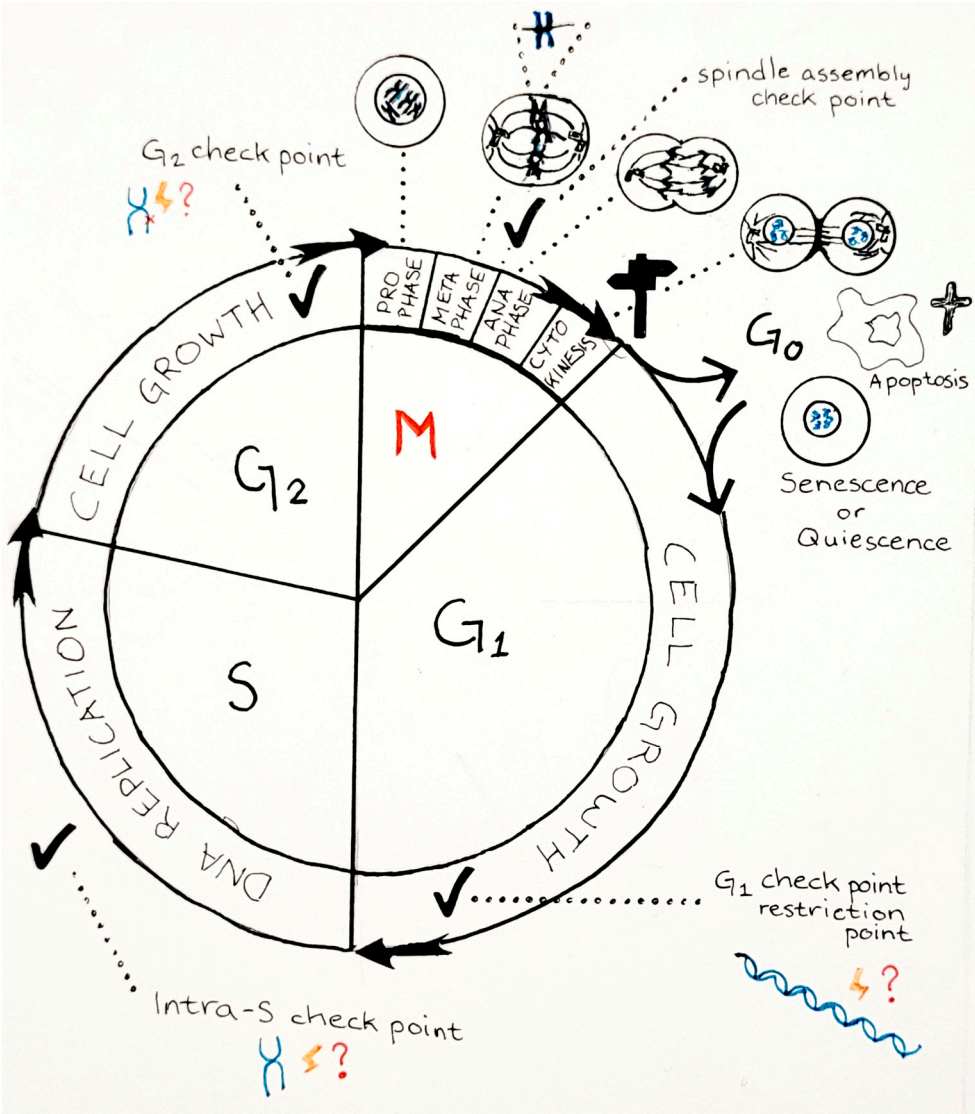


Figure 3. Schematic view over the cell cycle. M: mitosis.

DNA damage can cause mitotic dysregulation, where the cell no longer knows when to stop dividing, or skips important DNA repair checkpoints, causing further damage. Inherited genetic mutations can increase the risk of accumulating DNA damage and consequently the risk of developing tumours.

After this short introduction on genetics, we are hopefully ready to wade through this thesis with, at the very least, a sense of accomplishment in recognizing some of

the concepts. First, we will get acquainted with *Multiple Myeloma* (MM), the bone marrow cancer of plasma cells this thesis is focused on. Then, we will get technical and discuss some of the methods used to analyse inherited genomic data, and how to best compare DNA from MM cases to healthy controls. After this somewhat heavy section, we will review what is already known about inherited MM susceptibility. We will start by exploring what is known about *familial MM* (fMM), that is, when several members in a family are afflicted with MM. We will also investigate the *genome-wide association studies* (GWASs) on common genetic variation that influences the susceptibility of acquiring MM, preceding, and succeeding Paper I-IV in this thesis. We will then go through the four papers of original work this thesis is based upon. Finally, we will try to put the results from these papers into a larger perspective, including what we have learned in previous sections.

Multiple Myeloma

> "Progress in the realm of cancer research appears to be dependent upon a gradual accumulation of facts rather than upon revolutionary and outstanding discoveries."
>
> *"Recent Work on the Cancer Problem"*
> *(Geschickter 1930)*

Multiple Myeloma (MM) is a common hematologic malignancy defined by an uncontrolled expansion of plasma cells, usually in the bone marrow. It is an incurable and fatal disease that predominantly affects the elderly After non-Hodgkin's lymphoma, MM is the second most common blood cancer and constitutes approximately 15% of blood cancers (Sung et al., 2021). World-wide MM is diagnosed in about 175,000 people each year and accounts for 1% of all incident cancer cases (Bray et al., 2018; Sung et al., 2021). In Sweden, about 600 people are annually diagnosed with MM (Regionala Cancercentrum, 2021). Thanks to significant improvements in treatment over the last 15 years, the survival and quality-of-life for MM patients has improved considerably, and close to 55% are alive 5 years after diagnosis (Hemminki, Försti, & Hansson, 2021). While the causes of MM are incompletely understood, an aggregation of MM cases in families (familial MM; fMM) has consistently been reported by numerous authors since the 1920's (Table 2). Familial MM constitute about 1-2% of MM cases and studies suggest family members have a 2-4-fold increased risk of acquiring MM (Camp, Werner, & Cannon-Albright, 2008; Kristinsson et al., 2009; Morgan et al., 2014).

Table 1. Timeline over selected historical findings.

1844	First well-documented case of MM, “mollities ossium”
1847	Henry Bence Jones reports on urinary protein later to be called Bence Jones Protein in a patient with MM
1850	Macintyre reports on the clinical course of above patient, first well-known report of MM
1854	Rudolf Virchow describes amyloids
1867	Herman Weber found amyloid in kidneys and spleen of multiple myeloma patient
1873	Term “Multiple Myeloma” coined by J von Rustizky
1875	The term “plasma cell” is first used by Heinrich Wilhelm Gottfried von Waldeyer-Hartz (but probably a description of a tissue mast cell)
1880	The term “Bence Jones Protein” is coined by Fleischer
1889	Kahler detect cardinal findings (1) bone pain (2) pathologic fractures (3) cachexia (4) BJ proteinuria
1890	Santiago Ramón y Cajal described plasma cells accurately
1890	Emil Adolf von Behring and Shibasaburō Kisato describes antibodies as antitoxins
1891	Paul Ehrlich describes the antitoxin as an antibody
1895	T. von Marschalkó further describes plasma cells
1900	James H Wright describes myeloma tumour cells as essentially plasma cells
1922	Stanhope Bayne-Jones and D. W. Wilson discovered that there were two types of Bence Jones proteins
1929	Bone marrow aspiration introduced by Arinkin
1937	“The Tiselius apparatus” is described (electrophoresis)
1939	IgG is discovered by Tiselius and Kabat
1944	IgM is discovered by Waldenström and Pedersen, and Kunkel independently
1948	Ouchterlony double immunodiffusion invented by Örjan Ouchterlony
1953	Grabar technique immunoelectrophoresis
1956	Leonard Korngold and Rose Lipari discovered relation between serum proteins in MM and Bence Jones protein
1958	First HLA antigen is detected by Jean Dausset
1959	IgA discovered by Heremans et al
1961	MGUS discovered by Jan Waldenström
1964	IgD discovered by Rowe and Fahey
1965	Cooper, Peterson and Good delineate T and B lymphocytes in chicken
1966	IgE discovered independently by Ishizaka, Ishizaka and Hornbrook, and Johansson and Bennich

Clinical Characteristics

MM is characterized by an uncontrolled growth of plasma cells in the bone marrow, clinically resulting in hypercalcemia, renal failure, anaemia, and bone pain. Plasma cells are responsible for producing antibodies, and normally, the type of antibody produced varies from plasma cell to plasma cell, leading to a functional, polyclonal antibody pool. In MM, however, the malignant plasma cell clone produces an antibody with a single specificity, i.e., *monoclonal* antibodies (also known as a *paraprotein* or "*M protein*"). The monoclonal antibodies can be detected in peripheral blood or in urine through protein electrophoresis. In the bone marrow, malignant plasma cells outcompete normal blood cell formation, resulting in anaemia and suppressed polyclonal antibody production leading to an elevated risk of infections. Additionally, the stimulation of osteoclasts by tumour-secreted factors causes lytic bone lesions, leading to pain, pathologic fractures, and hypercalcemia. Finally, the pathological antibody production often produces free immunoglobulin light chains, which can cause kidney damage and amyloidosis.

MM is defined by M protein (IgG or IgA) ≥30g/l in blood serum, or monoclonal light chains in urine ≥0.5g in 24h, and/or ≥10% clonal plasma cells in the bone marrow; AND related organ tissue or tissue impairment (hypercalcemia, kidney failure, anaemia or bone lesions) (Regionala Cancercentrum, 2021).

MM is preceded by monoclonal gammopathy of uncertain significance (MGUS), a common, clonal condition that is present in approximately 1.6-3.2% of Caucasians older than 50 years (Axelsson, Bachmann, & Hallen, 1966; Robert A. Kyle et al., 2006). Patients with MGUS do not yet satisfy the criteria for MM, but progress to MM at an annual rate of approximately 1% (R. A. Kyle et al., 2002).

Epidemiology and Risk Factors

MM, like MGUS, is most prevalent in the elderly. The median age at diagnosis is 72 (Regionala Cancercentrum, 2021). In addition to MGUS, other risk factors include increasing age, male sex 54% and African American or African genetic descent (Gebregziabher, Bernstein, Wang, & Cozen, 2006; Samy, Ross, Bolton, Morris, & Oliver, 2015; Thorsteinsdottir et al., 2018). More inherited risk factors will be discussed in the next chapter. Obesity is another MM risk factor (Parikh, Tariq, Marinac, & Shah, 2021) and has historically been suspected to increase MM risk (Lauby-Secretan et al., 2016; Parikh et al., 2021). However, a recent genetic study that used Mendelian randomisation, could not find an association between obesity an MM risk (Went et al., 2017). Neither alcohol consumption nor smoking confers a higher risk of developing MM (Andreotti et al., 2013; Andreotti et al., 2015). Alcohol consumption even confers a lower MM risk (Andreotti et al., 2013; Santo, Liao, Andreotti, Purdue, & Hofmann, 2019). Another suggested risk factors

is exposure to pesticides, and farmers have been reported to have an increased MM risk (R. A. Kyle & Rajkumar, 2007). Other occupations, for example, firefighters have also been suggested to associate with MM (Hemminki, Forsti, Houlston, & Sud, 2021).

Treatment and Prognosis

A detailed review of current treatment protocols is beyond the scope of this thesis. In short, the anti-neoplastic treatment in MM is based on various combinations of corticosteroids, proteasome inhibitors (e.g., bortezomib), immunomodulatory drugs (e.g., thalidomide or lenalidomide), therapeutic antibodies (e.g., the anti-CD38 antibody daratumumab), conventional chemotherapy (e.g., melphalan), and high-dose chemotherapy with stem cell transplant. The choice of treatment depends on age, comorbidities, genetic markers, and response to therapy. Prognostic factors include age, co-morbidity, kidney function, stage of disease, treatment-tolerance, cytogenic aberrations, and blood levels of S-beta2-mikroglobulin and lactate dehydrogenase (LDH) (Regionala Cancercentrum, 2021). The introduction of proteasome inhibitors, immunomodulatory drugs, and therapeutic antibodies has significantly improved outcome. However, MM remains an incurable and ultimately fatal disease.

Short Overview of B, Plasma and MM Cell Development

Haematopoiesis is the continuous formation of blood cells from haemopoietic stem cells (HSCs) and involves several steps of stem cells and progenitor cells developing into more mature cells (Fig. 4). The plasma cell is a mature, fully differentiated B cell that secretes large amounts of antibodies (immunoglobulins, Ig) (Male, 2014). Antibodies are Y-shaped proteins that contain heavy chains and light chains, with variable regions at their antigen-binding sites (Male, 2014). During B cell development, several processes at different stages create diversity in the antibody repertoire, fitting them to their targets (antigens) in a Darwinistic “survival of the fittest”-like manner.

B cell development begins in the bone marrow where haematopoietic stem cells (HSCs) mature stepwise to produce common lymphoid progenitors (CLPs), from which pro-B cells develop (Meffre, Casellas, & Nussenzweig, 2000). In the Ig Heavy Chain genes (*IgH*) of these cells, DNA is rearranged in a process called VDJ recombination (Male, 2014; Meffre et al., 2000). In the *IgH* gene, there are several almost identical copies of the three segments encoding the variable (V), the diversity (D) and the joining (J) regions of the Ig heavy chain.

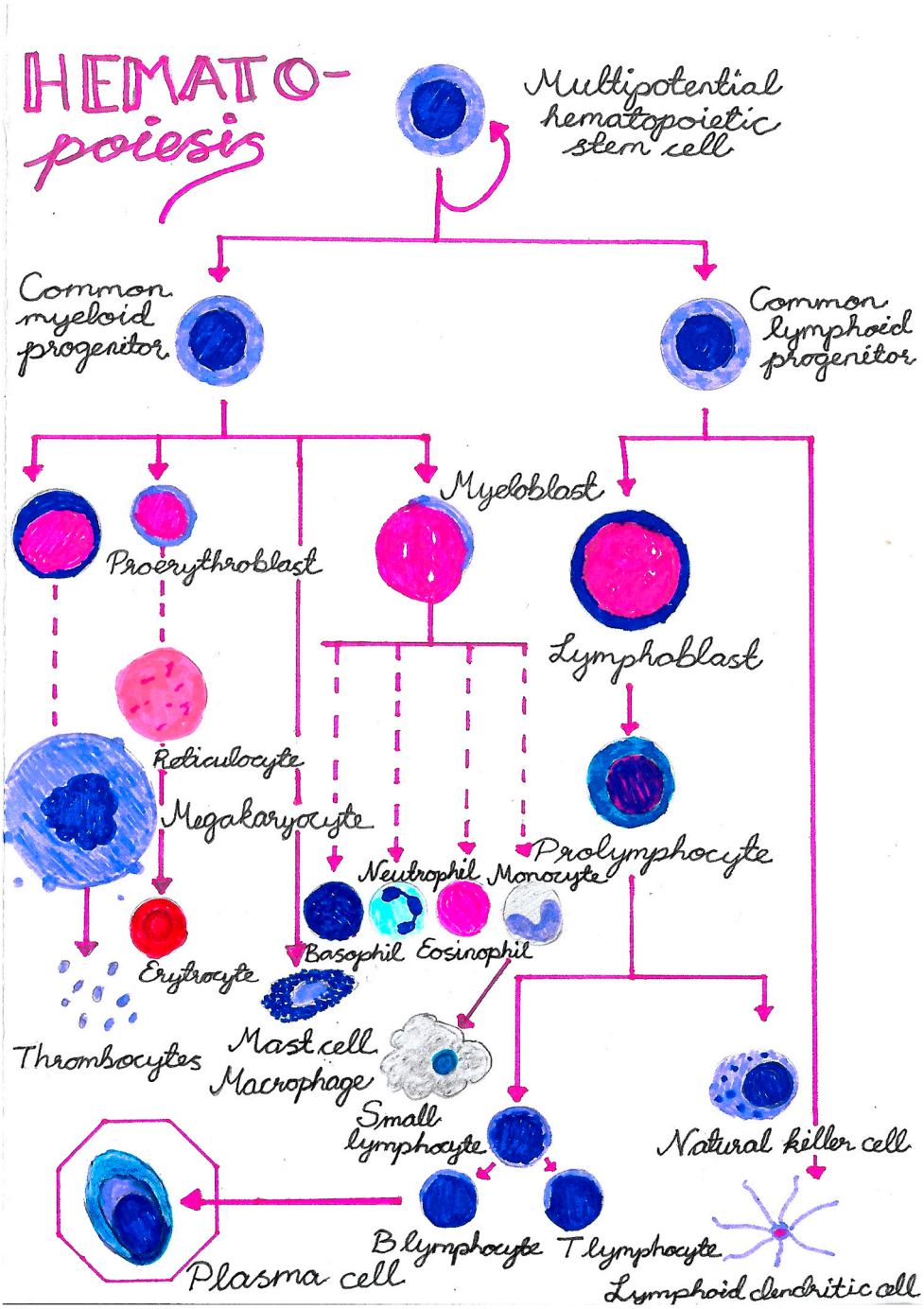


Figure 4. Schematic view over haematopoiesis.

Through DNA sequence rearrangement, one copy of V, D and J, respectively, combine and VDJ recombination is complete. In pre-B cells, a similar recombination occurs at the Ig light chain locus, but with only two segments, V and J (Meffre et al., 2000). At this point, immature B cells express IgM antibodies on their surfaces as B cell receptors (BCRs) (Meffre et al., 2000). They are now ready to leave the bone marrow and transition to secondary lymphatic tissue, such as the spleen or lymph nodes (Male, 2014; Meffre et al., 2000). Upon exposure to antigen, the naïve B cells that bind with high affinity, survive at the expense of low-affinity cells that undergo apoptosis (Barwick, Gupta, Vertino, & Boise, 2019; Male, 2014). The positively selected activated B cells either differentiate into short-lived plasma cells secreting IgM, or undergoes further maturation. Following T cell-dependent B cell activation, B cells rapidly proliferate in germinal centres. During this proliferation, somatic hypermutation (SHM) and class-switch recombination (CSR) introduces further variation, and B cell clones with high-affinity antibodies are positively selected (Barwick et al., 2019). Somatic hypermutation is a process where point-mutations are introduced in the Ig heavy and light chain genes' variable segments, while CSR is similar to V(D)J-recombination but takes place in gene segments that encode the constant region of the Ig protein that determines Ig class (Male, 2014). High-affinity B cells subsequently differentiate either to become memory B cells or plasmablasts (Barwick et al., 2019; Tellier & Nutt, 2019). Plasmablasts then migrate to the bone marrow and further matures into non-proliferating long-lived plasma cells (Tellier & Nutt, 2019).

The cell of origin in MM has been proposed to be a post-germinal centre clonotypic B cell (Barwick et al., 2019; Johnsen et al., 2016). It has then been suggested that this B cell differentiates to a premalignant plasma cell in the bone marrow, with persistent proliferation abilities (Johnsen et al., 2016). Transformation into malignant plasma cells would then be a stepwise process of accumulating DNA damage, first with a premalignant state where a clone of plasma cells with acquired somatic mutations that promote malignant transformation, secrete identical, monoclonal antibodies (MGUS) (Johnsen et al., 2016).

The existence of MM stem cells (MMSCs) is debated (W. Guo et al., 2021). The cancer stem cell theory is based on the notion that a few cancer cells in solid or liquid tumours have stem cell-like properties (Dalerba, Cho, & Clarke, 2007). Stem cells, such as haematopoietic stem cells and embryonic stem cells, mainly have three properties (Dalerba et al., 2007). Firstly, they have the ability to differentiate into different types of cells in a hierarchal process that continuously replenishes tissue or blood cells. Secondly, they can self-renew, that is, form new identical stem cells to maintain the stem cell pool. Lastly, stem cells can maintain a tissue homeostasis by regulating the number of differentiating cells in different environments by responding to different stimuli. The observations that support the cancer stem cell theory are: (1) only a few cells have the ability to expand when transplanted into

immunodeficient mice, (2) the CSCs have a distinct expression of surface markers and separates when using for example flow cytometry, and (3) the tumours are heterogenic with a mixed population of tumorigenic and nontumorigenic cancer cells (Dalerba et al., 2007). The earliest putative MMSC is the suggested cell of origin, a post-germinal centre pre-PC B cell (Johnsen et al., 2016).

A detailed review of MM development through different stages of acquired mutations is beyond the scope of this thesis. In short, in early oncogenesis two main types of genomic aberrations occur, and two different types of MM cells can be identified: (1) the nonhyperdiploid cells, which have chromosomal translocations, and (2) the hyperdiploid cells, which have an elevated number of chromosomes (Ninkovic & Quach, 2020). Nonhyperdiploid MM translocations (the transfer of a chromosome part to a non-homologous chromosome) predominantly involve the *IgH* locus (Ninkovic & Quach, 2020). Hyperdiploid MM is characterised by gain of primarily the odd chromosomes, 3, 5, 7, 9, 11, 15, 19, and 21, respectively. This distinction in MM cells is used clinically, where approximately 50% of MM cases belong to the hyperdiploid subgroup (Ninkovic & Quach, 2020). Clinically, the hyperdiploid MM subgroup is generally associated with a more favourable prognosis (Prideaux, Conway O'Brien, & Chevassut, 2014).

Plasma Cells and Autophagy

Autophagy is an evolutionarily conserved degradation process where organelles and macromolecules are sequestered into double-membrane vesicles and transported in autophagosomes, which fuse with lysosomes for degradation or recycling (Hansen, Rubinsztein, & Walker, 2018; Oliva & Cenci, 2014). In nutritional starvation, autophagy is induced, and is important for sustaining energy homeostasis in the cell (Hansen et al., 2018; Oliva & Cenci, 2014).

In plasma cells, autophagy is important for differentiation and maintenance of long-lived plasma cells in the bone marrow, but also for immunoglobulin (Ig) production (Oliva & Cenci, 2014; Pengo et al., 2013). During plasma cell differentiation, the endoplasmic reticulum (ER) expands rapidly to accommodate Ig synthesis and secretion. This intense process triggers ER and proteasome stress through degradation of Ig by-products (Oliva & Cenci, 2014). In chimeric mice, where autophagy is inactivated, plasma cells exhibit an expanded ER and secrete more antibodies. However, the autophagy-deficient plasma cells die prematurely, and there is a loss of long-lived plasma cells in the bone marrow (Oliva & Cenci, 2014; Pengo et al., 2013).

Methodology

DNA Sequencing Techniques

In 1977, Frederick Sanger et al introduced the first DNA sequencing technique. It would remain the most commonly used for many years, and an improved version of the technique is still used today (Heather & Chain, 2016; Sanger, Nicklen, & Coulson, 1977). Sanger used a dideoxynucleotide or "chain-termination" technique, currently referred to simply as Sanger sequencing. In short, Sanger sequencing is based on a "synthetic" DNA replication, incorporating the four nucleotide bases each labelled with a fluorescent tag that terminates the replicate nucleotide chain prematurely (dideoxynucleotides or ddNTP) (Heather & Chain, 2016). A carefully titrated mix of ddNTPs and "regular" nucleotides (dNTPs) produces DNA sequences of different lengths each ending with a fluorescent dideoxynucleotide (Heather & Chain, 2016). Using gel electrophoresis, these are then ordered by size allowing the sequence to be identified by their fluorescent tags (Heather & Chain, 2016). With this technique, single genes or small genomes up to approximately 5,000 bases can be sequenced (McCombie, McPherson, & Mardis, 2019).

Sequence-by-synthesis

The first "next-generation" sequencing technique was introduced by Margulies et al in 2005 (Margulies et al., 2005). This technique enabled highly parallelised sequencing by using micro-beads and then pyrosequencing (454, Roche). Illumina followed using flow cells instead of beads (Heather & Chain, 2016). On beads or clustered on a flow cell, the target DNA was then sequenced using a luminescent pyrosequencing technique measuring pyrophosphate production, or a fluorescent technique similar to Sanger sequencing, respectively (Heather & Chain, 2016). Both methods run in cycles in a massively parallel manner. While 454 was the first major commercially available next-generation sequencing method, the Illumina technique is the most widely used today (Heather & Chain, 2016). Because sequencing in these two methods is performed by a process similar to replication, they are called sequence-by-synthesis techniques, or SBS for short.

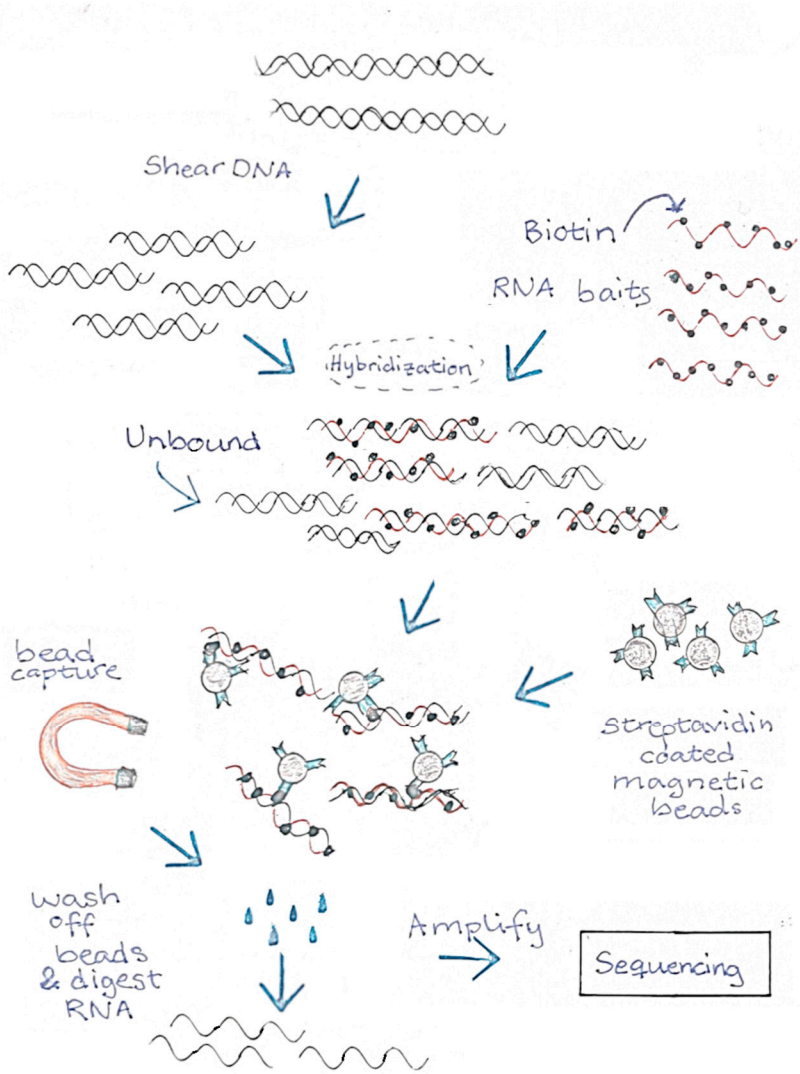


Figure 5. Whole-exome sequencing. Agilent SureSelect Target Capture. Adapted from fig. 1 (Giuffre et al., 2012).

SNP microarrays

While Sanger Sequencing, 454 and Illumina are SBS techniques, genotyping single-nucleotide polymorphisms (SNPs) focuses on single bases that are interrogated simultaneously. In for example Illumina SNP microarrays, beads arranged in microarrays are coated with complementary oligonucleotides, short sequences of nucleotides, that stop one base before the SNP genotyped (LaFramboise, 2009). The beads are distributed in wells in an array. The sample DNA is then allowed to

hybridize with (bind to) the probe DNA. Labelled nucleotides are added that emit fluorescent light to bind the last spot, i.e., at the SNP site. Illumina's Infinium Global Screening Array, for example, include approximately 650,000 SNPs, genotyped simultaneously (Illumina, 2021).

Whole-exome sequencing

In whole-exome sequencing only the exons of genes are selected for sequencing. After DNA is sheared into fragments, biotin-coated RNA probes are hybridised to the sample DNA "capturing" only the exons (Fig. 5) (Bamshad et al., 2011). Streptavidin-coated magnetic beads then bind the biotin on the probes and they are pulled down using a magnet (Bamshad et al., 2011). The beads are washed off and the exons are ready for next-generation sequencing.

Statistical analyses and study design

Hardy-Weinberg Equilibrium

At a biallelic site with the alleles A and a, the four possible genotypes, are AA, Aa, aA and aa. Genotypes can be homozygous (AA, aa), or heterozygous (Aa, aA). The allele frequency in a population is the proportion of a specific allele (A) among all added alleles in the population (2 A:s per homozygous individuals, 1 A per heterozygous). Over time, the genotypes reach an equilibrium called Hardy-Weinberg equilibrium (HWE) (Laird, 2011). If the two allele frequencies are denoted p and q, and since $pq = qp$ and $p + q = 1$, the Hardy-Weinberg Equilibrium, derived from Mendel's first law of segregation, can be stated:

$$p^2 + 2pq + q^2 = 1$$

Mendel's first law assumes that every individual inherits one allele randomly from the mother's two alleles, and likewise from the father. Deviations from HWE can be caused by evolutionary events such as mutation, selection, gene flow or non-random mating, for example inbreeding (Laird, 2011). In genetic statistical analyses, deviation can result from genotyping and genotype calling errors, and population stratification (Wittke-Thompson, Pluzhnikov, & Cox, 2005).

Linkage Disequilibrium

Linkage Disequilibrium (LD) is defined as the non-random association between alleles at different loci in the genome in a given population. If there is no statistical

dependency between alleles at two biallelic loci with alleles A/a and B/b, then the probability of inheriting both A and B equals the product of their allele frequencies, $p_{AB} = p_A \times p_B$. In this case, the alleles are in linkage equilibrium. For example, this situation often occurs if the loci are located on different chromosomes and there are no functional connections between the two genes. However, if the loci are located closely together on the same chromosome, the two alleles tend to be inherited together as there is physical, and as a result also statistical, dependency between the two loci. In this case, $p_{AB} \neq p_A \times p_B$, and the two loci are in linkage disequilibrium.

The linkage disequilibrium constant (D) is the deviation from equilibrium defined:

$$D = p_{AB} - p_A p_B$$

However, not knowing which allele (e.g., A or a) is the most frequent, D could be negative, and D is also sensitive towards very rare variants. Correcting for this (calculations not shown) leads to the D' metric, which attains values between 0 and 1, where 1 denotes total LD and 0 denotes linkage equilibrium. D' and a correlation coefficient r^2, are usually used as a metrics of how likely it is that two SNPs are co-inherited in the same haplotype block (Laird, 2011).

Genome-wide association studies (GWASs)

The first genome-wide association study (GWAS) was published in 2006, scanning the genome for variants predisposing to age-related macular degeneration (Dewan et al., 2006). The first study with rigorous methodology, however, was performed by the Wellcome Trust Case Control Consortium (WTCCC) on seven major diseases and was published in 2007 (Wellcome Trust Case Control, 2007).

In GWASs, the aim is to find DNA sequence variants that associate with a given disease phenotype or quantitative trait (Uffelmann et al., 2021). Variants that are overrepresented among cases relative to controls are identified. The variants that are represented on SNP microarrays are chosen with the intention to capture common variation in the whole genome and serve as tag SNPs for the regions they are located in, often tagging a haplotype block. The genome is "scanned" by testing associations between all SNPs and the trait simultaneously using an additive model calculating the per allele risk. Because all SNPs are tested, appropriate multiple-testing corrections need to be applied, usually Bonferroni correction (Johnson et al., 2010; Uffelmann et al., 2021).

The quality of data is ensured by applying different filters. Variants and samples with a high level of missing data are usually excluded, as are variants that show significant deviation from HWE (Laurie et al., 2010; Weale, 2010).

In a case-control study, it is crucial that the cases and controls are comparable apart from the trait studied. If not, their variant distribution may differ for reasons other than trait-associated variation, and, as a result, there might be a higher rate of false positive findings. Particularly, it is crucial that the cases and controls have the same geographic ancestry. For example, if cases of Scandinavian ancestry are compared with controls of Southern European ancestry, one could see a false association for the variant responsible for lactase persistence, as this trait is more common in Scandinavia (Swallow, 2003).

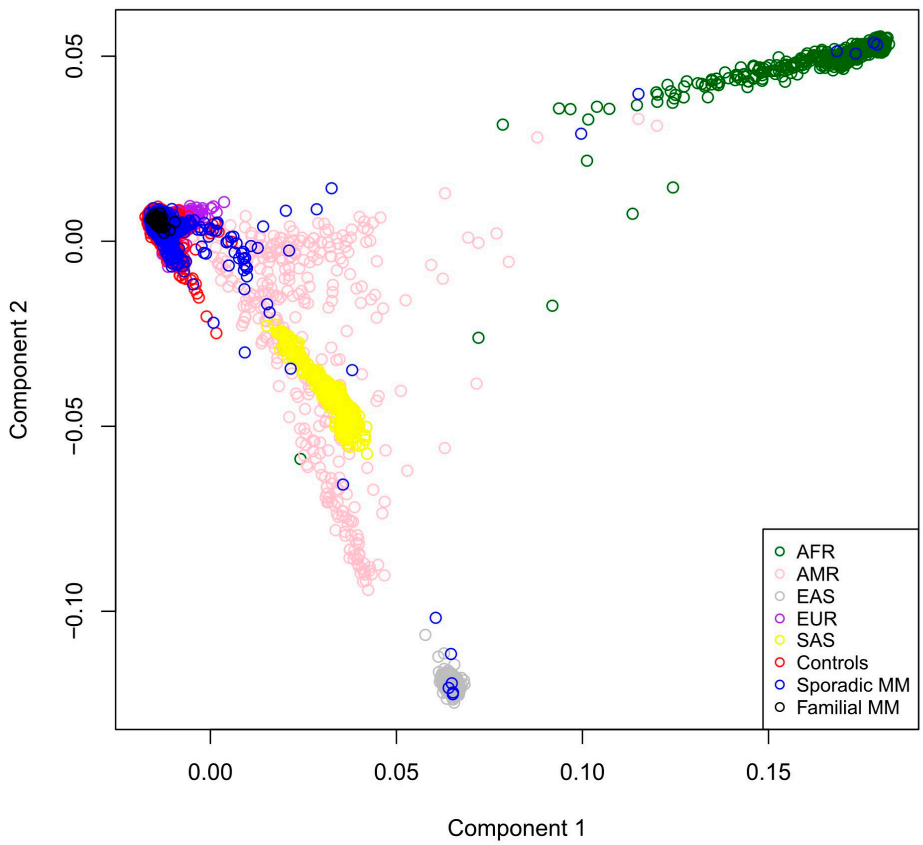


Figure 6. MDS-plot from Paper III. Populations from 1000G: AFR: African, AMR: American, EAS: East Asian, EUR: European, SAS: South Asian.

Principal Component Analysis (PCA) is frequently employed to either match cases and controls based on geographic ancestry, or to control for this genetic variation by including the first principal components as covariates in the statistical analysis (Price et al., 2006; Weale, 2010). The first principal components tell us, on a population level (both cases and controls), where in our high-dimensional we can find the most variation. This is a type of dimension-reduction technique where high-dimensional data is viewed in fewer dimensions. A large proportion of common genetic variation can be attributed to ancestry. By spiking in data from individuals of known geographic ancestry, individuals will co-cluster corresponding to ancestry when plotting the first principal components. Multidimensional scaling (MDS) is a similar analysis that can correct for cryptic relatedness seen in data sets when for example, two related individuals participate in the same study (Weale, 2010). This could lead to false positive (or negative) results by inflating the importance of certain alleles. MDS analysis can be used for detecting individuals with regions of DNA that are shared due to a common ancestor. $\hat{\pi}$ is defined as the proportion of alleles shared Identity By Descent (IBD), i.e, from a common ancestor (Purcell et al., 2007). Two siblings, for example, can at a locus either share identical alleles, share one of the alleles or share none of the alleles, rendering an IBD score of 2, 1 and 0 at that locus, respectively. Samples with a pair-wise $\hat{\pi} > 0.20$ is usually considered closely related and excluded. Identity By State (IBS) is a closely related metric using the same score system estimating the “genetic distance" between unrelated individuals and can be used for plotting clusters similar to PCA (Fig. 6) (Purcell et al., 2007).

Imputation of unobserved SNPs and INDELs not probed on SNP microarray, is commonly used in GWASs to increase genomic resolution (Uffelmann et al., 2021). The SNP microarray data only inform about the combined genotype of SNP for both chromosomes, and not which in the pair of homologous chromosomes an allele is located. Therefore, before imputation, the SNP data first needs to be haplotype-phased (Uffelmann et al., 2021). In haplotype phasing, each of the two alleles in a genotype is assigned to the predicted underlying chromosome using a mathematical reconstruction of homologous chromosomes. This prediction would not be possible if LD structures were not preserved over generations (Y. Li, Willer, Sanna, & Abecasis, 2009). To predict the SNP alleles between two genotyped SNPs, a reference panel of haplotypes is used to calculate the probability that, for example, the alleles of two SNPs are in LD. As with matching cases and controls, a panel of haplotypes from people of similar genetic descent improves imputation quality. The HapMap project was launched in 2002 and investigated LD structure in 269 samples (International HapMap et al., 2007). Since then, larger whole-genome data sets have been produced for example 1000 genomes (1000G) (T. G. P. Consortium, 2012; Genomes Project et al., 2010; Genomes Project et al., 2015). 1000G phase 3, contains whole-genome sequencing data from 2,054 participants world-wide. In

addition, the UK10K project has provided a reference data set on a cohort of 4,000 participants (U. K. Consortium et al., 2015).

GWAS Meta-Analysis

Combining data sets from for example international collaborators increases the sample size, improving statistical power to detect additional variants. Using summary data has been reported to be as efficient as using individual participant data, keeping data protected (Lin & Zeng, 2010). The statistical models used are often random effects or fixed effects. The limitation of a fixed-effects model is that it does not consider that the direction of association might be different in the populations included. Using a random effects model allows for this but is more stringent and has reduced statistical power (Thompson, Attia, & Minelli, 2011).

When combining data sets from different populations that could be genotyped on different sequencing platforms, perhaps imputed with different reference populations, and/or using different bioinformatic tools; there is a chance that these differences would be captured in analysis rather than true candidate variants (Evangelou & Ioannidis, 2013). Other sources of heterogeneity include, differences in phenotype definition, differences in ancestry mix, differences in quality control, differences in genotyping technique, population stratification and gene-environment effects specific to one or several of the studies included (Evangelou & Ioannidis, 2013). I^2 is a metric that attempts to measure the heterogeneity between data sets. If Q equals the Cochran's Q statistic and k is the number of studies included in meta-analysis I^2 is defined:

$$I^2 = \frac{100(Q - (k - 1))}{Q}$$

I^2 estimates the percentage of variance that is not due to chance (West, Gartlehner, & Mansfield, 2010). Zero to 25% is considered low, 25 to 50% moderate and more than 50% high heterogeneity. Cochran's Q statistic is based on the potential difference between studies being larger than differences between individual participants, and follows a χ^2 distribution (West et al., 2010).

Polygenic Risk Scores

Polygenic risk scores (PRSs) use GWAS data to produce a score based on the sum of risk alleles weighted by their effect sizes (Ni et al., 2021; Torkamani, Wineinger, & Topol, 2018). This assumes an additive model of genetic risk, where having two risk alleles confers a two-fold risk increase compared with having one. In the simple

version, the effect size (odds ratio) estimated in the GWAS analysis is used as weight (Ni et al., 2021). If we let S be the PRS, k the index of the SNP, β the log-transformed odds ratio (OR) for SNP k, and n the number of risk alleles, the polygenic risk score can be defined:

$$S = \sum_{k=1}^{k} \beta_k n_k$$

In more advanced models, the allele weights are estimated using other statistical models, though the additive model is still common (Chatterjee, Shi, & Garcia-Closas, 2016; Ni et al., 2021; Torkamani et al., 2018) . Further increasing predictive power, the statistical models often include variants with *p*-values above the genome-wide significance threshold (Slunecka et al., 2021). This introduces considerations on LD structure since there can be many variants under a chosen threshold in the same haploblock (Slunecka et al., 2021).

Heritability

Heritability is an estimate on how much of the phenotypic variation that can be explained by genetic variation, and is based on the following assumption:

$$Var(Y) = Var(G) + Var(\epsilon) + 2Cov(G, \epsilon)$$

where *Var(Y)* is the phenotypic variation, *Var(G)* the genetically driven variation in the phenotype, *Var(ϵ)* the environmental driven variation in the phenotype, and *Cov(G, ϵ)* is the covariation between genetic and the environmental components (Laird, 2011). *Cov(G, ϵ)* is often assumed to be zero, in which the *broad-sense heritability* becomes:

$$Var(G)/Var(Y)$$

Var(G) can be further divided into additive genetic variance V_A and dominant genetic variance V_D (Laird, 2011). If only considering the additive genetic variance the *narrow-sense heritability* is:

$$h^2 = V_A/Var(Y)$$

Using parent-child trios or mono-and dizygotic twins, h^2 can be estimated. In GWASs, heritability attributed to common variation by using all SNPs, both typed and imputed, can be estimated using genetic relationship matrices (GRMs) representing genetic similarity between subjects (Hall & Bush, 2016). The proportion of heritability that can be attributed to the identified risk variants is often reported.

The estimated proportion of heritability that is explained through common variation represented by tag SNPs, however, is low, and the "missing heritability" problem is widely discussed (Genin, 2020; Manolio et al., 2009). Potential genetic causes that have been suggested include large effects from rare variants, and structural variants (Genin, 2020). The statistical models used have also been criticized for only considering additive effects and not more complex effects, including dominant effects, gene-environment or gene-gene effects (Aschard et al., 2012). Increasing sample sizes improves power to detect significant variation, and small sample sizes could explain part of the missing heritability. Additionally, the use of heritability has been questioned, since it concerns common variation with small to modest effects, that has little clinical applications.

Analysing Exome Data

Since most Mendelian disease genes have been found in the protein-coding parts of the genome, whole-exome sequencing (WES) is commonly used when searching for rare disease-causing variants in familial disease (Bamshad et al., 2011). Rare protein-altering variants are more frequently predicted to have functional consequences or to be deleterious, than common variants. This is consistent with the assumption that rare disease-associated variants confer large effects compared with the moderate effects usually attributed to common GWAS variants (Bamshad et al., 2011).

Limitations of WES analyses include inadequate definition of truly protein-coding sequences, poorly defined traits and not considering non-coding parts of the genome (Bamshad et al., 2011). Technical issues include differences in probe efficiency, low coverage in certain regions, sequencing errors due to chance, or for example high locus heterogeneity, and difficulties in sequencing small insertions or deletions (INDELs).

Several different strategies can be used to analyse WES data (Bamshad et al., 2011). In parent-child trios, WES data can be filtered to identify *de novo* mutations in the affected child. Several affected members in a pedigree can be investigated for segregating variants. Another strategy is to filter out variants from unrelated affected cases that are present in reference databases such as 1000G, and then apply filters based on predicted pathogenic effect and evolutionary conservation. Other association approaches include mutation burden testing, either in single genes or genes encoding proteins in the same functional pathway, and more direct analyses testing for mutations in known candidate genes. These study designs all have their limitations. Recessive diseases could be caused by low-frequency rather than rare or ultra-rare variants, and be filtered out in analysis. Association studies are often underpowered due to small sample sizes, and *de novo* mutations detected in parent-child are not necessarily pathogenic.

Germline Genetics of MM

When a trait is observed more commonly in a family than would be expected by chance, one can suspect that the trait is caused by inherited variants in the germline genome. Other causes for familial aggregation can be common environmental exposure within a family, or a combination of common exposure and genetic factors. If the trait is caused by a variant in a single gene, the mode of inheritance is monogenic; if multiple variants in several genes cause the trait, the mode of inheritance is polygenic (Laird, 2011).

Case Reports and Genetic Epidemiology

From the first reports of familial multiple myeloma (fMM) in the 1920's, until the introduction of next-generation sequencing in the 2000s, the literature consists mainly of case studies (Geschickter & Copeland, 1928; Meyerding, 1925). Most publications report pairs of siblings with MM, but parent and child, and second-degree relatives were also reported (Table 2). Seldomly, families with more than two affected family members have been reported. In the 1960's, when all isotypes of immunoglobulins (Ig) had been discovered and immuno-electrophoresis was widely used, reports started to include the Ig isotype of the M protein. IgG-κ was, as in sporadic MM, the most frequent isotype. IgG-κ is also the most commonly concordant isotype in MM families (Table 2).

A number of studies have reported a two- to four-fold elevated MM risk in first-degree relatives to MM probands (Camp et al., 2008; Kristinsson et al., 2009; Morgan et al., 2014). In a recent epidemiological study, using data obtained through linking the Swedish Cancer Registry to the Multi-generation Registry provided by Statistics Sweden, Frank et al observed an increased risk in MM relatives by a factor of 1.79 to 2.89, dependent on the type of relation (Frank et al., 2016).

Table 2. MM families described in the literature. d.: died at. BJ: Bence Jones protein

Reference	No. of cases	Relation	Ig	Comment
(Meyerding, 1925)	2	Nephew (57), aunt (bone disease and fractured leg, grandfather tumor of the neck)	-	
(Geschickter & Copeland, 1928)	2	Brothers	-	
(Mandema & Wildervanck, 1954)	2	Sisters (63,65)	-	
(Nadeau, Magalini, & Stefanini, 1956)	3	Father (58), son (67) and daughter (66)	-	
(Herrell, Ruff, & Bayrd, 1958)	2	Brothers (53, 55)	-	
(Hirsch & Schwarz, 1959)	2	Sisters (59, 57)	-	
(Castleman, 1959)	2	Brothers (70)	-	Possible plasmacytoma
(Manson, 1961)	2	Sisters (76, 67)	-	
(Grossman, Ownby, Grossman, Kaplan, & Wolfe, 1963)	2	Brothers (71, 63)	-	
(Leoncini, 1964)	2	Sisters (60, 67)	IgA BJ	BJ both in serum and urine
(Thomas, 1964)	2	Brother (56) and sister (60)	-	
(Alexander & Benninghoff, 1965, 1967)	3	2 brothers (58, 67) and 1 sister (72)	-	
(Spengler et al., 1966)	2	Niece and aunt	IgG-λ IgA-λ	
(Riccardi, Di Girolamo, & Parziale, 1966)	2	First cousins	-	
(Robbins, 1967)	2	Brothers (57, 59)	-	
Williams 1967 ??				
(Rigby, Pratt, Rosenlof, & Lemon, 1968)	2	Uncle (d. 63) and one of twin nieces (d. 38)	-	
(Takekuma et al., 1968)	2	Brother and sister	IgD IgG	
(Berlin, Odeberg, & Weingart, 1968)	2	Mother (69) and son (49)	IgG-κ, IgG-κ +IgA-κ	
(Manigand et al., 1970)	2	Mother (71) and daughter (47)	IgG κ	
(Barbieri & Grampa, 1972)	2	Sisters (47, 52)	IgG -	
(Meijers, De Leeu, & Voormolen-Kalova, 1972)	3 (+3)	Son (45), mother (69), maternal uncle (59)	-	5 siblings in 1st gen, 1 brother with MGUS (IgG1-λ) 4 siblings in 2nd gen 1 brother with IgA1-κ + IgG1-λ MGUS, 1 sister IgG1-κ MGUS
(Goldstone, Wood, & Cook, 1973)	2	Mother (84) and daughter (60)	IgG IgA	
(Boga, Jako, Doman, Magyar, & Konyar, 1973)	2	Mother (80 or 78) and son (60)	IgA-κ IgG-κ	
(Maldonado & Kyle, 1974)	2	Brothers (74, 65)	-	10 siblings in total
(Maldonado & Kyle, 1974)	2+1	Brother (67) and sister (55) + brother (67)	- /- IgG-κ	11 siblings in total, one MGUS at first, then MM

Reference	No. of cases	Relation	Ig	Comment
(Maldonado & Kyle, 1974)	2	Brother (64) and sister (56)	IgG-κ IgG-κ	5 siblings in total
(Maldonado & Kyle, 1974)	2	Brother (70) and sister (58)	IgG-κ IgA-κ	8 siblings in total
(Maldonado & Kyle, 1974)	2	Brother (56) and sister (59)	IgG-κ IgG	11 siblings in total
(Maldonado & Kyle, 1974)	2	Brothers (55, 77)	LC IgG-κ	10 siblings in total
(Maldonado & Kyle, 1974)	3	Brother (70), sister (77), nephew (46)	IgA IgG-λ	Daughter (34) and brother MGUS (79) IgA-λ + IgG
(Maldonado & Kyle, 1974)	2	Sisters (72, 54)	IgA-κ IgG	2 of 3 daughters
(Law, 1976)	2	Brothers (57, 75)	IgG-κ IgG-κ	
(Wiedermann, Urban, Wiedermann, & Cidl, 1976)	2	Brothers (~75, 70)	IgG-κ IgA-κ	
(Zawadzki, Aizawa, Kraj, Haradin, & Fisher, 1977)	2	Mother (62) and son (39)	IgG-λ IgG-κ	
	2	Female and male, first cousins (both 73)	IgA-κ IgG-λ	
	2	Brother (74) and sister (63)	IgG-κ, -	4 siblings in total
	2	First cousins (both male and 65)	IgA, -	
	2	Brothers (58, 69)	IgG-λ, IgG-κ +IgA-λ	
(Isobe, Shaura, Itoh, & Fujita, 1981)	2	Brothers	IgA and IgG-κ	
(Shoenfeld, Berliner, Shaklai, Gallant, & Pinkhas, 1982)	2	Mother (78) and son (63)	IgG-λ IgG-λ	
(Bourguet, Grufferman, Delzell, DeLong, & Cohen, 1985)	2	Brothers (75, d. 50s)	IgA-λ, -	Interview based
	2	Brother and sister (47, d. 53)	IgA-κ, -	
	2	Sisters (70,-)	IgG-λ, -	
(Gobbi et al., 1985)	2	Siblings	-	
	2	Siblings	-	
(Horwitz, Levy, & Rosner, 1985)	3	Sister (53) and 2 brothers (51, 56)	IgG-κ, IgG-κ IgG-κ	All three siblings
(Hubert et al., 1985)	2	Father (67) and daughter (48)	IgG-κ IgG-κ	
(Judson, Wiltshaw, & Newland, 1985)	2	Male monozygotic twins (61, 62)	IgG-κ IgG-κ	
(Klingler & Daweke, 1985)	2	Mother (63) and son (46 or 56)	-, IgG-λ	
(Grosbois et al., 1986)	2	Brothers (58, 57)	IgG-κ IgG-κ	Identical HLA haplotype, 2 healthy brothers
(McCrea & Morris, 1986)	2	Father (82) and son (53)	IgG-κ, IgG	

Reference	No. of cases	Relation	Ig	Comment
(McCrea & Morris, 1986)	2	Mother (83) and daughter (52)	IgA-κ BJP-κ	
(McCrea & Morris, 1986)	2	Male monozygotic twins (58, 61)	IgA-λ, BJP-κ	
(Rostoker, Uzzan, Baumelou, & Chapman, 1986)	2	Mother and daughter	IgG-κ IgG-κ	
(Comotti, Bassan, Buzzetti, Finazzi, & Barbui, 1987)	2	Female possibly monozygotic twins (60, 68)	BJP IgG-κ	
(Unakami et al., 1987; Watanabe, Suzuki, & Unakami, 1988) ??	2	2 brothers	-	1 brother with MGUS 6 siblings in total
(Kanoh, Funahashi, Okada, & Horii, 1988)	2	Father (72) and son (53)	IgG-λ IgG-κ	
(Loth, Perrotta, Lima, Whiteaker, & Robinson, 1991)	2	Sisters (58, 56)	LC-λ IgA-λ	Identical HLA 8 siblings in total, 3 and 2 offspring, respectively, healthy
(Abreu, Goncalves, da Silva, & Viseu, 1992)	2	Sisters (52, 46)	IgG-λ IgG-λ	
(Willems, Kuypers, Meijerink, Holdrinet, & Mensink, 1993)	2	Father (70) and son (41)	-	
(Willems et al., 1993)	2	First male cousins (43,55)	-	
(Willems et al., 1993)	2	Father and daughter (56)	-	
(Crozes-Bony, Palazzo, Meyer, De Bandt, & Kahn, 1995)	2	Father and daughter	LC LC	
(Snowden & Greaves, 1995)	2	Male monozygotic twins (62, 63)	IgA-λ IgA-λ	
(Roddie, Dang, & Parker, 1998)	3	2 sisters (66, 77) and 1 brother (75)	LC-κ IgG-κ IgG-κ	
(Grosbois et al., 1999)	2	Father (89) and son (61)	IgG-κ, IgA-κ +LC-κ	
(Grosbois et al., 1999)	2	Brother (63) and sister (68)	IgG-κ, IgG-κ,	MGUS IgG-κ in 1 brother
(Grosbois et al., 1999)	2	Sisters (60, 80)	IgG-κ, IgG-κ	
(Grosbois et al., 1999)	2	Mother (77) and son (67)	IgG-κ IgG-λ	MGUS in sister(/aunt)
(Grosbois et al., 1999)	2	Mother (57) and daughter (46)	IgG	
(Grosbois et al., 1999)	2	Aunt (81) and nephew (33)	IgG-κ, IgG-κ	MGUS IgM in relative
(Grosbois et al., 1999)	2	Brother (49) and sister (65)	IgG-κ IgG-κ	
(Grosbois et al., 1999)	2	Sisters (82, 88)	IgG-λ, IgG-κ	
(Grosbois et al., 1999)	2	Sisters (72, 62)	IgG-κ IgG-κ	
(Grosbois et al., 1999)	2	Mother (81) and son (58)	IgG-κ IgG-κ	
(Grosbois et al., 1999)	2	Brothers (65, 71)	IgG-κ IgG-κ	
(Grosbois et al., 1999)	2	Twin sisters (68, 69)	IgG-λ IgG-κ	

Reference	No. of cases	Relation	Ig	Comment
(Grosbois et al., 1999)	2	Sisters (53, 51)	IgA-λ, IgG-κ	
(Grosbois et al., 1999)	2	Brother (66) and sister (69)	IgG-κ IgG-κ	
(Grosbois et al., 1999)	2	Brothers (81, 61)	IgG-κ IgG-λ	
(Watanabe, Suzuki, Murakami, & Komatsu, 1999)	2	Sisters (79, 68)	IgG-λ IgA-κ + BJ-κ	11 siblings, 1 sister with leukemia
(Sobol et al., 2002)	4	Mother (76), daughter (65), granddaughter (56) and niece (56)	-, -, IgA-λ, -	Mother and daughter also had breast cancer. Granddaughter with *BRCA1* mutation
(H. T. Lynch et al., 2005)	4	Brother (60) and daughter (40), sister (72) and daughter (59)	-	4 siblings in total
	4	2 sisters (d. 78, d. 60s), brother (65) and first sister's son (55)	-	8 siblings in total
	2	Sisters (60, 51)	-	8 siblings in total, considerable family history of different cancer types
	3	Mother (79) and both offspring; daughter (66) and son (71)	-	5 siblings, leukemia in 2 family members
	3	Mother (63) and 2 daughters (76, 79)	-	3 children in total (all daughters)
	3	Father (64), daughter (45) and father's aunt (68)	-	
	4	Mother (76), daughter (65) another daughter's daughter (56), niece (56)	-	First two cases also had breast cancer (73, 63)
	2	Father (d. 75) and son (61)	-	Extensive family history of cancer especially in breast
	2	Grandson (53) and maternal grandfather (68)	-	2 siblings in total (grandson)
	2	Father (78) and son (65)	-	4 siblings in total (son)
	2	Father (74) and son (51)	-	4 siblings in total (son)
	2	Mother (65) and daughter (60)	-	4 siblings in total (daughter), 5 siblings in total (mother)
	2	Mother (80) and daughter (61)	-	Only child
	2	Father (65) and son (45)	-	3 siblings in total
	2	Father (84) and son (62)	-	3 siblings in total
	2	Woman (d. 56) and mother's aunt (d. 60s)	-	Spouse also affected (38), 2 siblings in total
	3	Brother (70s) and sister (d. 66) + nephew (49)	-	6 and 3 siblings in total, respectively
	3	2 brothers (58, 53) and 1 sister (62)	-	9 siblings in total, 2 brothers with MGUS
	2	Brother (55) and sister (67)	-	7 siblings in total
	2	Brothers (63, 62)	-	3 brothers in total

Reference	No. of cases	Relation	Ig	Comment
	2	Sisters (-, 60)	-	No other siblings
	2	Sisters (-, 76)	-	No other siblings
	2	Sisters (53, 49)	-	No other siblings
	2	Brother and sister (60, 47)	-	3 siblings in total
	2	Male cousins (67, 58)	-	3 siblings in total in both cousins
(Olujohungbe, Gledhill, & Satchithananathan, 2006)	2	Male monozygotic twins (78, 81)	IgG-λ IgG-κ	
(Gerkes, de Jong, Sijmons, & Vellenga, 2007)	3	Father (71), son (45) and daghter (65)	IgG x3	2 siblings in total
	2	Father (60) and son (38)	IgA-λ IgG	5 and 6 siblings in total, respectively
(Zepeda & Dominguez, 2007)	3	Mother (56), daughter (56) and son (53)	-	
(Ozet, Guran, & Beksac, 2008)	3	Man (50), paternal uncle (76) with daughter (48)	IgA, ns, IgG	ns:non-secretory
(Henry T. Lynch et al., 2008)	5	Sister (80) with daughter (45), 2 brothers (50, 72) and 1 paternal half-sister (85)	-	Totally 3 familiy members with MGUS, 5 with prostate cancer
(Coleman et al., 2009)	2	Mother (68) and son (51)	-	3 and 2 siblings in total respectively
	2	Mother (75) and son (53)	-	7 and 2 siblings in total respectively, leukemia in familiy history
(Jain, Ascensao, & Schechter, 2009)	2	Brother (51) and sister (d. 53)	κ, -	5 siblings in total, MGUS in two brothers (IgM-λ+IgA-λ, 60 and IgG-κ, 59)
	2	Brother (SMM, 50) and sister (59)	IgG-λ -	8 siblings in total IgG-λ+IgA-κ MGUS in mother (80), IgG-λ+IgG-κ in brother's son, brother concurrent Polecytaemia Vera
	2	Brother (68) and sister (d. 58)	IgA-κ, -	
	2	Mother (d. 72) and son (63)	-, IgG-κ + IgG-λ	
(Várkonyi et al., 2009)	2	Brother (66) and sister (68)	IgG-κ IgG-κ	Identical HLA haplotype
	2	Brothers (54, 48)	LC-κ IgG-κ	
(Wiernik, Wickramasinghe, & Dutcher, 2015)	2	Father (61) and son (42)	-	
	2	Father (59) and son (40)	-	
	2	Mother (66) and son (58)	-	
(Bolli et al., 2017)	2	Brother (d. 74) and sister (d. 66)	-	+ MGUS in sister's daughter and son (niece and nephew) MGUS

Reference	No. of cases	Relation	Ig	Comment
(Y. Y. Li, Fan, Wang, Bai, & Liu, 2017)	2	Brothers (70, 77)	IgG-λ IgA-κ	
(Pertesi et al., 2019)	2	Brothers (60, -)	IgG-κ, -	2 brothers (IgG-κ, -) and 1 sister (IgG-κ+l) MGUS, 11 siblings in total
	2	Mother and son (39)	-, IgG-κ	Another son (IgM-κ) MGUS
	2	Brother (78) and sister	-	5 siblings in total
	3	Brothers and niece (47)	-	6 siblings in total
(Catalano et al., 2021)	2	Father (90) and son (62)	-	Father: 2 siblings in total, son only child
	2	Male cousins (67, 70)	-	1 father with MGUS
	2	Male cousin (74) and son (47) of female cousin	-	
	2	Brothers (69, 57)	-	5 siblings in total
	2	Brothers (66, 55)	-	2 sisters, 1 brother and mother had other cancer types, 9 siblings in total
	2	Sisters (48, 69)	-	Second sister SMM, another sister plasmacytoma (70) 3 siblings in total
	2	Father (76) and son (49)	-	Son only child
	2	Uncle (71) and nephew (43)	-	4 siblings in total for both relatives
	2	Father (79) and son (45)	-	Son only child
	2	Father (60) and son (38)	-	6 siblings in total in both generations
	3	Brother (80) and sister (76) with son (56)	-	4 siblings in both generations

Germline Molecular Genetics in Familial MM

The first germline genetic study to my knowledge, is when Willems et al. sequenced three exons from the ***TP53*** gene in three families, each with two cases of MM (Willems et al., 1993). The gene *TP53* encodes a tumour suppressor protein (p53) that is crucial in cell cycle control and cell death (J. Chen, 2016). Acquired mutations in *TP53* are common in several cancer types, and rare germline *TP53* variants cause the Li-Fraumeni cancer-predisposition syndrome (Varley, 2003).

Willems et al noticed somatic *TP53* mutations in eight out of ten myeloma cell lines and were curious if germline mutations of *TP53* could explain the multiple occurrences of MM cases in the families studied (Willems et al., 1993). They focused on exon 5, 7 and 8 of *TP53*, but no germline variants were detected.

Roddie et al studied *TP53* in three siblings with MM (Roddie et al., 1998). Their father had died from gastric cancer and one son had lung cancer. Searching in the exon 5-8, only one mutation was found in exon 5 in one of three siblings, and it was suspected to be acquired rather than inherited.

Another gene that has been studied in familial cases of MM is the cyclin dependent kinase inhibitor 2A gene (***CDKN2A***). This tumour suppressor gene is important in cell cycle control, regulating two cell cycle pathways that involve p53 and the retinoblastoma-associated protein (pRb), respectively (Chan, Chiang, & Ngeow, 2021; Huschtscha & Reddel, 1999). Its effects are mediated by two separate proteins encoded by *CDKN2A*, but with different reading frames ($p14^{ARF}$ and $p16^{INK4}$). The isoform $p14^{ARF}$ binds MDM2, sequestering p53 in the nucleolus and preventing its breakdown (Basu & Murphy, 2016). The $p16^{INK4}$ isoform binds to cyclin-dependent kinase 4 (CDK4) and cyclin-dependent kinase 6 (CDK6) inhibiting their association with cyclins that phosphorylates rRb, a key regulator of the G1/S transition in the cell cycle (Huschtscha & Reddel, 1999). When rRb is hyperphosphorylated, it dissociates from the transcription factor-complexes involving E2Fs, whose target genes subsequently induce S-phase entry (Huschtscha & Reddel, 1999).

At least two cancer syndromes are associated with mutations in *CDKN2A* (Henry T. Lynch et al., 2002) The familial atypical multiple mole-malignant melanoma syndrome (FAMMM) is an autosomal dominant disease in which affected family members present with multiple dysplastic nevi and/or malignant melanoma. In a similar syndrome with autosomal dominant inheritance, the familial atypical multiple mole melanoma-pancreatic carcinoma syndrome (FAMMPC), family members are affected by either both or one of (multiple) melanoma and pancreatic cancer.

Interestingly, in a family of northern European descent reported by Dilworth et al, four out of nine family members developed melanoma and one developed MM (Dilworth et al., 2000). Another family member developed an unknown brain malignancy. When sequencing *CDKN2A* in family members afflicted by cancers, they found that they all carried a germline mutation resulting from a 24-nucleotide duplication in the start of exon 1α. None of three non-carriers developed a malignant disease. Additionally, a loss of heterozygosity (LOH) was detected in malignant plasma cells from the family member with MM.

In another case study by Shah et al, a 36-year-old woman developed IgG-λ MM and several other cancer types (Shah, Boyd, Houlston, & Kaiser, 2017). After successful treatment with melphalan, prednisolone followed by high dose melphalan with autologous stem cell rescue, the M protein could no longer be detected. A few months before the diagnosis, she had been treated for malignant melanoma, additional melanomas were treated at age 53, 58 and 62. At the age of 34 in situ breast cancer was discovered and treated with mastectomy. Still in remission she

developed lung adenocarcinoma at the age of 66. Sanger sequencing of *CDKN2A* in peripheral blood detected a heterozygous germline mutation c. 213.C>A. Additionally LOH of this mutation was detected in the lung cancer DNA. Because of the patient's complete remission from MM, the authors suggest that this mutation is associated with a favourable prognosis, possibly by conferring increased sensitivity to chemotherapy.

The *CDKN2A* region is also a susceptibility locus detected in genome-wide association studies in many cancer types including MM. This will be discussed in the Paper II-section of this thesis.

Grass et al investigated two paratargs, paratarg-7 and paratarg-8, that are often targeted by paraproteins in MM and MGUS (Grass et al., 2009; Grass et al., 2011). While no mutations were detected in the encoding genes, *STOML2* and *ATG13*, respectively, they found that when these paratargs are hyperphosphorylated (pP-7 and pP-8), they confer an increased risk of developing MM or MGUS. This feature showed a dominant inheritance, but the genetic basis of this endophenotype remains unknown, and it has not been individually replicated.

The stomatin-like protein 2, mitochondrial protein (SLP-2) encoded by ***STOML2***, is an inner membrane mitochondrial protein that has been reported to be essential for respiratory chain supercomplex formation, important for optimal mitochondrial function (Mitsopoulos et al., 2015). SLP-2 deficient T cells show a defective mitochondrial respiration, which results in reduced proliferation and an impaired $CD4^+$ T cell response (Christie et al., 2012; Mitsopoulos et al., 2015). Preuss et al reported that hyperphosphorylation of SLP-2 is a consequence of inactivated PP2A, a ubiquitously expressed phosphatase (Preuss et al., 2011). Its consequences, however, remain unknown. SLP-2 is overexpressed in several cancer types including oesophageal, lung and laryngeal cancer, and is generally associated with poor survival (D. Chang et al., 2010; D. Liu et al., 2013; F. Sun et al., 2015; L. Zhang et al., 2006).

In a nutrient-rich environment, **ATG13** is hyperphosphorylated, and formation of the autophagy-inducing complex, including ULK1 is inhibited (Puente, Hendrickson, & Jiang, 2016). Recently, ULK1-ATG13-induced autophagy has been linked to cell cycle control (Z. Li et al., 2020). Li et al show that CDK1-induced ULK1-ATG13 phosphorylation promotes mitotic autophagy and cell cycle progression, and that a double-knockout of *ULK1* and *ATG13* blocks cell cycle progression decreasing cancer cell proliferation (Z. Li et al., 2020). Whether or not autophagy is active throughout the whole cell cycle has been debated (Z. Li, Ji, Wang, Liu, & Zhang, 2016). Li et al found that autophagic flux i.e., the rate of autophagic degradation, is high in early phases of mitosis and in the S-phase, although still present at different levels throughout the whole cell cycle (Z. Li et al.,

2016). They speculate that a high autophagic flux could help degrade damaged organelles and provide energy during S-phase and mitosis.

Pertesi et al identified the DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease (***DIS3***) as a familial MM gene, detecting putative loss-of-function variants in four unrelated families (Pertesi et al., 2019). This gene, located at 13q21.33, is recurrently somatically mutated in MM (Lohr et al., 2014). Pertesi et al used whole-exome sequencing to investigate germline variants in 154 families with two or more cases of MM and/or MGUS.

DIS3 encodes a catalytically active protein that contains two domains, PIN and RNB, with endonuclealytic and exonuclealytic activity, respectively, and is a subunit of the RNA exosome complex (Tomecki et al., 2010). The RNA exosome complex has important roles in mRNA processing, and the maintenance of genome stability (Laffleur et al., 2021). The complex consists of 11 subunits and is located both in the nucleus and the cytoplasm (Laffleur & Basu, 2019; Laffleur et al., 2021). Other subunits of the complex include DICER and EXOSC10, and different subunit compositions affect ribonuclease activity (Laffleur & Basu, 2019). The RNA exosome complex can degrade RNA that is unbound from DNA by RNA helicases, thereby regulating the accumulation of R-loops (DNA:RNA hybrids) (Laffleur & Basu, 2019). R-loops are crucial for DNA repair, DNA replication, RNA splicing and class-switch recombination in maturing B cells (Laffleur & Basu, 2019).

During class-switch recombination and somatic hypermutation in differentiating B cells, the RNA exosome is involved in RNA transcription and processing (Laffleur & Basu, 2019). Without the RNA exosome complex the variation in the immunoglobulin repertoire is altered (Laffleur & Basu, 2019; Laffleur et al., 2021). Recently, it has been reported that an accumulation of RNAs exposes B cells to chromosomal translocations, by disturbing CTCF/cohesin activity, and thereby the topological structure of DNA (Laffleur et al., 2021). If DNA:RNA hybrids are accumulated, sense single-stranded DNA are overexposed to AID-induced deamination. This affects the distribution of mutations in somatic hypermutation (Laffleur et al., 2021). In CSR, it causes a DNA overhang at the DNA break points, and the DNA repair is shifted from non-homology end-joining (NHEJ) of blunt DNA-overhang ends to microhomology repair of considerable end overhang (Laffleur et al., 2021).

In two male siblings with MM and MGUS, Pertesi et al found a splicing variant (rs769194741) at a slice donor site in exon 13, predicted to cause skipping of this exon and a premature termination codon, resulting in a truncated protein lacking parts of its two active domains. The remaining two unaffected siblings were non-carriers. Using lymphoblastoid cell lines from the carriers of this variant, they showed that the variant allele was transcribed, but then subsequently degraded by

non-sense mediated decay (NMD). Consistent with this, a reduction of expression by 50% could be seen compared to non-carriers.

In another family, containing a female proband with MM, her mother affected with amyloidosis and two of her uncles affected with MM, they found a second splicing variant also at a splice donor site, but in exon 14 that encodes parts of the RNB domain. Sequencing data was only available from the proband, her unaffected sister and her mother. The variant was, however, only found in the proband making it unlikely that it segregated with her uncles.

The third variant (rs141067458) resulted in a stop-loss and was found in two unrelated families. In the first family, WES data was available from seven out of eleven siblings of whom three were affected with MGUS, and two with MM. Only one of the siblings, a man affected with MGUS, did not carry the variant. The second family consisted of five siblings, where data was available from three siblings, two unaffected and one affected with MM. Another sibling had died from MM in the past. Both unaffected sisters were non-carriers and all five available (unaffected) offspring also had wildtype genotypes. No alteration in mRNA expression could be found in the carrier, but DIS3 protein levels were reduced by approximately 50%.

Rare, deleterious variant burden tests were also performed, showing an excess of such variants in MM cases compared to unaffected controls (OR = 1.92, 95% CI: 1.25-2.96). The stop-loss variant had a higher frequency in cases compared with controls (OR = 3.07, 95% CI: 1.38-6.87). In an independent data set, the variant was very rare, and although non-significantly, it was consistently more common in MM cases than controls with a similar OR (3.15).

Bolli et al analysed WES data on six family members in a family with five cases of either MM or MGUS (Bolli et al., 2017). They filtered out variants with a minor allele frequency (MAF) above 1%, as well as non-coding or silent variants, and ranked the remaining variants by their Common Annotation Dependent Depletion (CADD) scores, predicting their damaging effect. The variant with the highest score was a missense mutation in *EP300*. To shortlist potential common or rare variants associated with MM, Bolli et al used a non-filtered version of the data set and then investigated these variants (1,642) in a case-control study with unselected MM cases and healthy controls. They found three significant variants in their analysis; one variant in *MKL1* and two variants in *NDUFA8*.

The E1A binding protein p300 gene (***EP300***) encodes a lysine acetyltransferase, the transcriptional coactivator p300, which is involved in the regulation of transcription, proliferation, the cell cycle, cell differentiation and the DNA damage response (Attar & Kurdistani, 2017; Dutta, Tiu, & Sakamoto, 2016; Dutto, Scalera, & Prosperi, 2018). p300 participates in the regulation of G1/S transition, S phase progression by acetylating DNA replication and chromatin assembly proteins, and

by inducing *CDKN1A* (Abbas & Dutta, 2009; Attar & Kurdistani, 2017; Dutta et al., 2016). p300 is required in haematopoiesis and perturbs HSC differentiation in embryonic stem cells and chimeric mice (Dutta et al., 2016; Rebel et al., 2002).

EP300 translocations have been reported in myelodysplastic syndrome and acute myeloid leukaemia (Attar & Kurdistani, 2017). *EP300* is recurrently somatically mutated in several cancer types, including skin squamous cell, bladder, small cell lung cancer and marginal zone B-cell lymphoma (Attar & Kurdistani, 2017).

Recently, a p300-inhibitor combined with a BET-inhibitor (a related transcriptional regulator) showed anti-tumour effects in 16 different MM cell lines, both alone and in combination, through downregulation of c-MYC via suppression of IRF4 (Ryan, Giles, & Morgan, 2021).

Wei et al suggested Lysine demethylase 1A (***KDM1A***, also called *LSD1*) as a new MM susceptibility gene (Wei et al., 2018). The authors used WES data on fMM cases from 50 kindreds, and an additional 400 cases with either fMM or early-onset MM. They found a predicted deleterious, truncating germline mutation (c.805_806delAG) in *KDM1A* segregating in one family, and another such mutation in one early-onset case (c.707delA). A higher rate of *KDM1A* mutations in unselected patients with MM than in controls were detected using a mutation burden test. *KDM1A* knockdown mice resulted in increased immunoglobulin production, presence of a paraprotein, and plasma cell expansion. Introducing an additional *KDM1A* mutation in lymphoblastoid B cells from carriers of germline *KDM1A* mutations resulted in a lower *KDM1A* expression relative to wild-type plasma cells. Additionally, upregulation of the oncogene *CCND2* (Cyclin D2), involved in the MYC pathway, was observed in MM cells in *KDM1A* mutation carriers. Based on these findings, Wei et al suggested that *KDM1A* acts as a tumour suppressor gene.

KDM1A targets methylated H3 lysine 4 and 9 (H3K4 and H3K9, respectively) and demethylation of active and repressive histone marks results in transcriptional repression and activation, respectively (Kim, Kim, & Baek, 2021). In addition, KDM1A has non-histone functions, for instance protection of proteins from degradation, sustaining cellular metabolic balance, maintenance of self-renewal potential in stem cells, and regulation of differentiation, including in haematopoiesis (Kim et al., 2021). How this is consistent with its tumour suppressor role in MM remains to be elucidated.

Waller et al investigated SNP data from MM cases in eleven high-risk pedigrees (HRPs) in Utah, and highlighted *ARID1A* and *USP45* as candidate risk genes in MM (Waller et al., 2018). By developing a new method, they mapped shared genomic segments within eleven high-risk families, and identified one statistically significant 1.8 Mb region at 6q16 in one family. This region encompasses nine genes: *POU3F2*, *FBXL4*, *FAXC*, *COQ3*, *PNISR*, *USP45*, *TSTD3*, *CCNC*, and *PRDM13*. In addition,

Waller et al, identified one borderline-significant 8.9 Mb region at 1p36.11, where segments from two families overlapped. The overlapping segment contains 30 genes including *ARID1A*.

Waller et al then expanded their sample set and performed a directed analysis of the two segregating genomic regions using WES in an additional sample set with 44 more families with either fMM/MGUS or early-onset MM. Prioritizing rare and potentially deleterious variants, they were unable to find a segregating variant in the family with a shared segment at 6q16. However, in two of the additional MM/MGUS families, two separate variants were detected, located in the same domain of *USP45*; one stop-gain and one missense mutation segregating in three and two siblings, respectively. Two rare variants were discovered in the promising and overlapping segment at 1p36.11. In one of the original HRPs, a potentially deleterious missense variant was found in *ARID1A*, segregating in three out of four cases. The other variant was a missense mutation, also detected in *ARID1A*, and carried by two cousins. In the second HRP with the overlapping segment, no segregating variant was identified.

The AT-rich interaction domain-containing protein 1A, encoded by ***ARID1A***, is a key protein in a SWI/SNF chromatin remodelling complex (BAF). The main function of this complex is making chromatin accessible for transcription factors (Pagliaroli & Trizzino, 2021). ARID1A binds nucleosome DNA and disrupts histone-DNA contacts, initiates DNA translocation and forms DNA loops, making chromatin more accessible for binding (Xu & Tang, 2021). ARID1A is crucial for embryonic development through maintaining embryonic stem cell function (Gao et al., 2008).

While *ARID1A* is the most frequently somatically mutated gene in the SWI/SNF complex, several genes encoding different other subunits are also recurrently mutated in cancer cells, collectively in about 20% of all cancers (Pagliaroli & Trizzino, 2021; J. N. Wu & Roberts, 2013; Xu & Tang, 2021). ARID1A has been reported to have both oncogenic and tumour suppressive functions dependent on context and shows signs of cancer type-specific functions. (Pagliaroli & Trizzino, 2021; Xu & Tang, 2021). In hepatocellular cancer (HCC), *ARID1A* is upregulated in approximately 83% of tumours and has tumour-promoting functions in early phases of malignant transformation (Xu & Tang, 2021). However, ARID1A also exerts tumour suppression functions in several other cancer types and has been reported to inhibit common cancer-promoting pathways, including the PI3K/AKT and YAP/TAZ pathway, respectively (J. N. Wu & Roberts, 2013; Xu & Tang, 2021). As a DNA damage response, ARID1A has also been reported to initiate and maintain cell cycle arrest promoting DNA repair, through upregulation of *CDKN1A* expression (p21) (Xu & Tang, 2021). Additionally, ARID1A has been found to interact directly with p53 (Guan, Wang, & Shih Ie, 2011).

The ubiquitin specific peptidase 45 gene (***USP45***) encodes a deubiquitylase that interacts with ERCC1, through a motif in its N-terminal that is highly conserved in USP45 orthologues (Perez-Oliva et al., 2015). In humans, *USP45* is not well-studied, and its functions remain largely unknown. Perez-Oliva et al reported that UPS45 regulates DNA repair through deubiquination of ERCC1. ERCC1, together with XPF, form a DNA repair nuclease that targets interstrand DNA cross-linkings and repairs (in particular UV-light-induced) DNA damage through nucleotide excision repair (Perez-Oliva et al., 2015).

USP45 has been suggested to also form a complex with Spindly influencing cell migration (Conte, Griffis, Hickson, & Perez-Oliva, 2018). The Spindly protein, encoded by *SPDL1*, recruits dynein/dynactin to kinetochores during mitosis, affecting for instance kinetochore-microtubule attachment and spindle orientation (Conte, Baird, Davidson, & Griffis, 2018). While Spindly plays an important role in mitosis, aberrant mitosis is not observed in the absence of USP45, indicating that Spindly's role in its complex with USP45 is separate and possibly restricted to the interphase (Conte, Baird, et al., 2018).

Biallelic mutations in *USP45* are believed to have caused a severe form of inherited retinal dystrophies, Leber congenital amaurosis, in two Chinese families (Yi et al., 2019). Additionally, *USP45* has been associated with risk of early-onset coronary disease, and somatic mutations in *USP45* have been detected in breast cancer with an *APOBEC*-mutational signature (Trevino, 2019; Yamada et al., 2018).

Further expanding their sample set, with 186 familial or early onset, as well as 733 sporadic MM, Waller et al widened their search to include fellow functional complex members to the two candidate genes' proteins, in total 22 and 14 genes, respectively. In the 22 *USP45*-related genes, comparing to un-affected controls, the authors found a genome-wide suggestive segment including *ERCC1* and *ERCC2,* that segregated in 3 MM distantly related cases. They also searched in 919 MM/MGUS exomes for the 23 genes and found a predicted pathogenic missense variant in *ERCC4* in one early-onset and one sporadic MM case. Additionally, they found a stop-gain variant in *ERCC3* in the same domain, in a second early-onset MM case. Performing burden testing in all MM cases and controls, they also found two significant genes *GTF2H1* and *DDB1*. In the 14 genes connected to *ARID1A* they identified a suggestive borderline genome-wide significant segment in four distantly related MM cases containing the gene *PBRM1* and 31 other genes. No coding variants could be detected, but burden testing proved significant in seven out of the 15 genes encoding the BAF complex: *ARID1A*, *ARID1B*, *SMARCA4*, *ACTL6A*, *SMARCD3*, *SMARCC2*, and *SMARCE1*. Using a binomial distribution, the authors argue that occurrence in MM cases of two and seven burdened genes, respectively, in these complexes, is an unlikely finding.

Catalano et al searched for rare germline variants through whole-genome and exome sequencing in 21 MM families containing 46 affected (MM/MGUS) and 20 unaffected family members in total (Catalano et al., 2021). They used an in-house prioritization pipeline for rare variants (MAF < 0.001), that segregated in family members and had a CADD score > 20. Missense and loss-of-function (LoF) variants were then further filtered by different prediction scores based on evolutionary conservation, and several tools to predict if it is a deleterious mutation. They also searched for copy number variants (CNVs).

The authors detected 109 missense, 36 LoF and seven CNVs fitting their criteria. In four families, there were none. Two of the genes detected were found in two unrelated families (*KIF1B* and *DCHS1*). A few of the genes have been connected to MM. *SAMHD1* is a somatically mutated driver gene in MM and is involved in somatic hypermutation and class-switch recombination in B cell development (Azhar, Begum, & Husain, 2021). Variants in several of the genes are involved in IgG N-glycosylation: *B4GALT1,* ABCF2, *TAB1* (Shadrina et al., 2021). Common variants in or near genes involved in IgG glycosylation including *ABCF2* have associated with MM susceptibility in genome-wide association studies (GWASs) *TNFRSF13B* by Chubb et al and *ELL2* from Paper I (Chubb et al., 2013; Shadrina et al., 2021).

In conclusion, the authors highlight *DAB2IP* (tumour suppressor); *ABL2* (an oncogene); *SAMHD1*; *KMT2A and USP28* (has functional relationship to MM predisposition genes); *FOXO1*, *B4GALT1* and *NKX3-2* (immune-related); *FGFBP1* and *FGFBP2* (possible enhances MM proliferation), and finally *PROM1* (stem cell-related), as strong MM candidate genes.

Genetic Predisposition in Sporadic Multiple Myeloma

In 2012, Broderick et al performed the first genome-wide association study on MM susceptibility using SNP data from 1,675 MM cases and 5,903 controls, recruited in the UK and Germany (Broderick et al., 2011). The authors found significant associations with variants at two loci, 3p22.1 and 7p15. At 3p22.1, the lead variant rs1052501 maps to *ULK4* exon 17. While the variant encodes an amino acid change, it was predicted as non-pathogenic by SIFT and PolyPhen-2 (Adzhubei et al., 2010; Sim et al., 2012). The 7p15 variant tagged by rs4487645 maps to *DNAH11* intron 80 (82 exons), in a region of linkage disequilibrium also containing *CDCA7L*. Additionally, Broderick et al detected a promising association at 2p23.3, tagged by the intronic variant rs6746082 in *DTNB*.

Table 3. MM risk loci (excluding Paper !, II and IV). Genome-wide significant results ($p < 5 \times 10^{-8}$) in bold.
Ref: (1) (Broderick et al., 2011) (2) (Weinhold et al., 2013) (3) (Chubb et al., 2013) (4) (Duran-Lozano et al., 2021)

Location	Gene(s)	SNP	Risk allele	Risk allele frequency	OR P	Ref
2p23.3	*DTNB*	rs6746082	A	0.82 UK 0.84 German 0.82 UK repl	1.29 (1.17-1.42) 4.02×10^{-7}	1
3p22.1	***ULK4***	**rs1052501**	**G**	**0.20 UK 0.19 German 0.21 UK repl**	**1.32 (1.20-1.45) 1.81×10^{-8}**	**1**
7p15.3	***CDCA7L DNAH11***	**rs4487645**	**C**	**0.71 UK 0.76 German 0.73 UK repl**	**1.38 (1.28-1.50) 2.06×10^{-14}**	**1**
11q13.3	***CCND1***	**rs603965**	**G**	**0.72 UK 0.67 German 0.78 German repl**	**1.82 (1.52-2.19) 2.92×10^{-10}**	**2**
3q26.2	***LRRC34 ACTRT3 MYNN TERC***	**rs10936599**	**G**	**0.80 UK 0.79 German 0.80 UK repl 1 0.79 UK repl 2**	**1.26 (1.18-1.33) 1.74×10^{-13}**	**3**
6p21.33	***HLA PSORS1C2***	**rs2285803**	**A**	**0.32 UK 0.36 German 0.32 UK repl 1 0.29 UK repl 2 0.33 German repl**	**1.19 (1.13-1.26) 1.18×10^{-10}**	**3**
17p11.2	*TNFRSF13B*	rs4273077	G	0.12 UK 0.14 German 0.12 UK repl 1 0.11 UK repl 2 0.12 German repl	1.26 (1.16-1.36) 1.41×10^{-7}	3
22q13.1	***CBX7***	**rs877529**	**A**	**0.51 UK 0.45 German 0.49 UK repl 1 0.47 UK repl 2 0.46 German repl**	**1.23 (1.17-1.29) 2.29×10^{-16}**	**3**
13q.13.3	***SOHLH2***	**rs75712673**	**G**	**0.03 European**	**1.35 2.2×10^{-14}**	**4**

In a follow-up study, expanding the data set with an additional 698 MM cases recruited in Germany, Weinhold et al conducted a subgroup analysis on 1,661 of the MM cases where karyotype data was available. The authors found a polymorphism in *CCND1* that associated with risk of developing t(11;14)(q13;q32) MM (Weinhold et al., 2013). Further expanding the data sets from UK and Germany, Chubb et al could detect four additional risk loci in a subsequent GWAS totalling 4,692 MM cases and 10,990 controls (Chubb et al., 2013). The novel risk loci mapped to 3q26.2, 6p21.33, 17p11.2 and 22q13.1, containing the genes *MYNN*/*TERC*/*LRRC34*, *PSORS1C2*, *TNFRSF13B* and *CBX7*, respectively.

Duran-Lozano et al performed a GWAS selectively in Nordic populations totalling 5,320 cases and 422,289 controls (Duran-Lozano et al., 2021). The authors used SNP-data from a more homogenous population and imputed this data with a more

closely matched reference population. They were able to identify a low frequency intronic variant in *SOHLH2* at 13q13.3 (lead variant rs200203825). One of the variants within the LD block, rs75712673, maps to a region with open chromatin in plasma cells, and likely upregulates *SOHLH2* aberrantly in this cell type.

The spermatogenesis and oogenesis specific basic helix-loop-helix 2 gene (***SOHLH2***) encodes a transcription factor that regulates male and female germ cell differentiation, and has also been connected to cancer.

The Present Investigation

The overall aim of this thesis is to identify germline DNA sequence variants that predispose for Multiple Myeloma (MM).

In Paper I, II and IV, we performed case-control genome-wide association studies (GWASs) to identify germline single-nucleotide polymorphisms (SNPs) and small insertions/deletions (INDELs) that associate with MM risk.

In Paper I, we identified a novel significant association with *ELL2*, and a borderline suggestive association with *TOM1*.

In Paper II and IV we collaborated internationally in GWAS meta-analyses, and identified eight and six variants, respectively. The *TOM1* variant in Paper I replicated in these studies.

In Paper III, we performed SNP microarray and whole-exome sequencing analysis on 38 cases of familial MM. Constructing polygenic risk scores, we found direct evidence for a polygenic aetiology in familial MM, and estimated that about one-third of familial MM cases were associated with an enrichment of common risk variants identified by GWAS. In Paper IV, we extended our polygenic risk scores with newly identified risk variants, and again observed an enrichment of risk variants in familial cases.

Registries

To identify patients with familial MM, we cross-linked the Swedish Cancer Registry and The Swedish Multi-generation Register using Swedish personal identification numbers. Samples were then obtained from the Swedish National MM Biobank, which banks samples from patients with newly diagnosed MM.

In this study, a combination of these registries yielded registry data (1958-2011) from approximately 18,000 MM cases, out of which roughly 200 familial MM cases (1.1%). We calculated the relative risk of cancers recorded in the Swedish Cancer Registry, in first-degree relatives of MM patients. We constructed a model using Cox proportional hazards and a marginal survival hazards model, based on the model presented in Pfeiffer et al, that we implemented in C++ (Pfeiffer et al., 2004). The results obtained with this model were largely comparable to a conventional relative risk model based of a 2x2 contingency table (unpublished data; Table 4). Only MM, chronic lymphocytic leukaemia (CLL), and prostate cancer showed significant relative risk increases after corrected for multiple testing (Bonferroni limit $p < 0.0015$).

Table 4. Tumour types with nominally significant relative risk increase in first-degree relatives of MM patients, as determined by cross-linking the Swedish Cancer Registry and the Swedish Multi-generation Registry.

Tumour Type	Relative Risk Model			Cox Model
	RR	(95% CI)	*p*	Est. RR
Multiple Myeloma	**2.21**	(1.92-2.54)	1.78e-29	2.19
Bladder	**1.14**	(1.03-1.27)	0.00615	1.14
Chronic Lymphocytic Leukemia	**1.42**	(1.16-1.73)	2.97e-4	1.41
Colon	**1.10**	(1.01-1.20)	0.0139	1.11
Hepatocellular	**1.27**	(1.04-1.56)	0.0104	1.28
Ovarian	**1.15**	(1.00-1.31)	0.0245	1.12
Prostate	**1.10**	(1.05-1.17)	1.79e-4	1.09

Personal Identification Numbers

In 1947, personal identification numbers (PINs) were introduced in Sweden (SCB, 2017). At first, the Church of Sweden (Svenska Kyrkan) kept the records in parish registers (Ekbom, 2011; Skatteverket, 2021). Since 1991, however, The Swedish Tax Agency (Skatteverket) keep the records (SCB, 2017; Skatteverket, 2021). The IDs earlier consisted of a six-digit birthdate and a three-digit birth number. In 1967, one digit was added to the birth number, the "control digit", that could be calculated from the other three (SCB, 2017). Even though there are only 300 new-borns each day, on some dates the birth numbers have run out, especially on 1 January and 1 July in the 1950s and 1960s. This is due to a large number of immigrants having been assigned these dates owing to less stringent homeland records. Because of this,

some PINs have been reused. In 2016, approximately 23,000 PINs were estimated to have been reused (SCB, 2017). Since 2009 there is a possibility to receive a PIN with a birth date close to the accurate one. More than 11,000 Swedish inhabitants have a PIN with an incorrect birth date. Roughly 500 PINs are changed every year, and this is mainly due to incorrectly registered birth dates (SCB, 2017).

The Swedish Multi-generation Register

The Swedish Multi-generation Register is managed by Statistics Sweden (SCB) and comprises individuals born 1932 or later, who have been registered inhabitants in Sweden at some point in time since 1961 (Ekbom, 2011; SCB, 2021). These individuals, called "index persons", are linked to their biological parents. The overall coverage is high, with information on 97% of biological mothers and 95% of fathers available as of December 2005 (Ekbom, 2011). Additional information includes sex, country of birth, parents' PINs and country of birth (SCB, 2021).

The Swedish Cancer Registry

The Swedish Cancer Registry (SCR) was founded in 1958 and is managed by the National Board of Health and Welfare (Socialstyrelsen) (Socialstyrelsen, 2019; Turesson et al., 2007). In Sweden, health care providers are obligated to report newly diagnosed malignancies to the registry and approximately 60,000 malignant tumours are registered every year (Socialstyrelsen, 2019). Cancer reports are sent to one of the six regional cancer centres, which register tumours and send yearly updates to the national registry (Socialstyrelsen, 2019). The completeness and diagnostic accuracy for MM is high in SCR. Between 1964 and 2003, the completeness for MM was 95%, and accuracy 93% (Turesson et al., 2007).

Results and Discussion

PAPER I

In Paper I, we performed a GWAS and identified one new MM risk locus at 5q31 (lead variant rs56219066), in the elongation factor for RNA polymerase II 2 gen*e (ELL2*). We also found an association between the identified *ELL2* MM risk allele and lower IgA and IgG levels in peripheral blood. We also identified a promising borderline-significant locus at 22q13 which harbours a variant in intron 1 of *TOM1*. We noted that *ELL2* and *TOM1* are preferentially expressed in normal and malignant plasma cells

Four out of seven previously reported risk loci replicated at $p < 5 \times 10^{-8}$ (*ULK4*, *CDCA7L*, *TNFRSF13B* and *HLA*).

ELL2 encodes a key protein of the Super Elongation Complex (SEC), that enhances RNA polymerase II activity and is involved in transcription elongation (Z. Luo, Lin, & Shilatifard, 2012). ELL2 plays a key role in the differentiation of mature and memory B cells to plasma cells, and Ig synthesis (Santos, Arumemi, Park, Borghesi, & Milcarek, 2011). Briefly, ELL2 contributes to shifting the production of membrane-bound Ig to secretory Ig through enhancing *IGH*-mRNA translation (Martincic, Alkan, Cheatle, Borghesi, & Milcarek, 2009; K. S. Park et al., 2014; Shell, Martincic, Tran, & Milcarek, 2007).

The haploblock tagged by rs56219066 contains many variants in high LD, making it hard to pinpoint the causal variant(s). Among the variants in the haploblock, we noted a missense variant (rs3815768) encoding an amino acid substitution in a functional domain of ELL2. Additionally, in a follow-up paper, Ali et al identified an eQTL effect that could be responsible for lower IgA and IgG levels in risk allele carriers, as well as a putative causal variant rs3777189 that disrupts a MAFF/G/K transcription factor binding site, consistent with the decreased *ELL2* expression in plasma cells associated with the *ELL2* MM risk allele (Ali et al., 2018). Interestingly, conditional B cell-lineage *ELL2* knockout mice show reduced numbers of plasma cells in the spleen as well as reduced numbers of antibody-producing cells in the bone marrow (Ali et al., 2018). In mouse plasmacytoma cells, *ELL2* silencing decreases the ratio of secreted versus membrane-encoding immunoglobulin heavy chain transcripts.

Table 5.
MM risk loci as reported in Paper I, II and IV.
Genome-wide significant results ($p < 5 \times 10^{-8}$) in bold. RAF: risk allele frequency

Location	Gene(s)	SNP	Risk allele	RAF	OR *p*	Paper
5q31	***ELL2***	**rs56219066**	**T**	**0.73**	**1.25 (1.16-1.34)** **9.6×10^{-10}**	**I**
22q12.3	*TOM1*	rs138740	C	0.36 (Iceland: 0.42	5.7×10^{-8} 1.18 (1.11-1.25)	I
6p22.3	***JARID2***	**rs34229995**	**G**	**0.029**	**1.37** **1.31×10^{-8}**	**II**
6q21	***ATG5***	**rs9372120**	**G**	**0.218**	**1.18** **9.09×10^{-15}**	**II**
7q36.1	***SMARCD3***	**rs7781265**	**T**	**0.125**	**1.19** **9.71×10^{-9}**	**II**
8q24.21	***CCAT1***	**rs1948915**	**C**	**0.345**	**1.13** **4.20×10^{-11}**	**II**
9p21.3	***CDKN2A***	**rs2811710**	**G**	**0.657**	**1.15** **1.72 x 10-13**	**II**
10p12.1	***WAC***	**rs2790457**	**G**	**0.739**	**1.12** **1.77×10^{-8}**	**II**
16q23.1	***RFWD3***	**rs7193451**	**T**	**0.585**	**1.13** **5.00×10^{-12}**	**II**
20q13.13	***PREX1***	**rs6066835**	**C**	**0.083**	**1.26** **1.36×10^{-13}**	**II**
2q31.1	***SP3***	**rs4325816**	**T**	**0.77**	**1.12** **7.37×10^{-9}**	**IV**
5q23.2	***CEP120***	**rs6595443**	**T**	**0.43**	**1.11** **1.20×10^{-8}**	**IV**
6p25.3	*IRF4*	rs1050976	T	0.45 (Sweden)	1.10 5.93×10^{-8}	IV
7q22.3	***CCDC71L***	**rs17507636**	**C**	**0.74**	**1.12** **9.20×10^{-9}**	**IV**
7q31.33	***POT1***	**rs58618031**	**T**	**0.72**	**1.12** **2.73×10^{-8}**	**IV**
16p11.2	***PRR14***	**rs13338946**	**C**	**0.26**	**1.15** **1.02×10^{-13}**	**IV**
19p13.11	***KLF2***	**rs11086029**	**T**	**0.24**	**1.14** **6.79×10^{-11}**	**IV**

The Target of myb1 membrane trafficking protein (***TOM1***) is highly expressed in the bone marrow and its encoded protein is thought to be associated with vesicular trafficking at endosomes. TOM1 is an Endosomal Sorting Complex Required for Transport (ESCRT) protein and part of a complex at the endosomal membrane surface (Roach, Lang, Xiong, Ryhanen, & Capelluto, 2021). When cell-surface membrane proteins are ready to be degraded or recycled they are transported to the cytosol in vesicles through endocytosis. The vesicles are sent to endosomes where they bind to ESCRT proteins that invaginate the membrane to form intraluminal vesicles containing the plasma membrane protein or the so-called *cargo*. In the

endosome the vesicles are either proteolyzed in what is now a lysosome or recycled and sent to the plasma membrane (Roach et al., 2021).

TOM1 is also involved in autophagy and is thought to be a part of forming an amphisome, that fuses with a lysosome, via an alternative pathway, to become an autolysosome (Roach et al., 2021).

TOM1 has also been implicated in regulation of immunological pathways. Recently, the first pathological mutation in *TOM1* was reported, a missense mutation causing an early onset autoimmunity syndrome in a mother and son. Apart from autoimmune symptoms, they both had combined immunodeficiency which included antibody deficiency and lack of switched memory B cells (Keskitalo et al., 2019).

The borderline *TOM1*-association was later replicated in Paper II.

PAPER II

In Paper II, we participated in a multicentre collaboration conducting a joint meta-analysis containing 9,866 cases and 239,188 controls. We identified eight new MM risk loci in or near *JARID2* (at 6p22.3), *ATG5* (6q21), *SMARCD3*, (7q36.1), *CCAT1* (8q24.21), *CDKN2A* (9p21.3), *WAC* (10p12.1), *RFWD3* (16q23.1) and *PREX1* (20q13.13). All previously reported risk loci in or near *DTNB*, *ULK4*, *MYNN/TERC*, *ELL2*, *PSORS1C2*, *CDCA7L/DNAH11*, *CCND1*, *TNFRSF13B*, *TOM1* and *CBX7*, were confirmed (Broderick et al., 2011; Chubb et al., 2013; Weinhold et al., 2013).. We found decreased gene expression of *WAC* and *PREX1* in risk variant carriers Chromatin looping interactions and TADs were also detected at four of the risk loci. None of the risk variants associated with either age at diagnosis or sex. Furthermore, no associations were found with cytogenic subtype or MM-specific overall survival.

The Jumonji and AT-rich interaction domain containing 2 gene (***JARID2***) encodes a protein that binds directly to DNA and functions as a transcriptional regulator (Kouznetsova, Tchekanov, Li, Yan, & Tsigelny, 2019; Su, Deng, Shang, & Xiao, 2015). JARID2 acts as an accessory protein to the Polycomb repressive complex-2 (PRC2) that mediates histone H3K27 methylation associated with decreased gene expression (Martin & Moorehead, 2020). JARID2 stabilises PRC2 and facilitates its recruitment to target loci (Kouznetsova et al., 2019). JARID2 also regulates PRC2 methyltransferase activity and interacts with Polycomb repressive complex-1 (PRC1) mediating PRC1-PRC2 crosstalk and co-recruitment (Kouznetsova et al., 2019; Martin & Moorehead, 2020).

As a subunit of PRC2, JARID2 regulates haematopoietic stem cell (HSC) function where a depletion of JARID2 enhances HSC expansion and reconstitution capacity (Kinkel et al., 2015). Furthermore, *Jarid2* knockout mice show reduced abundance of B cell progenitors in the bone marrow, and a shift towards T cell differentiation

in peripheral blood, consistent with reduced production of B cells (Celik et al., 2018).

JARID2 is downregulated in B-CLL and acute monocytic leukaemia (AMOL) and is commonly mutated in T-cell acute leukaemia (Su et al., 2015). Su et al showed that JARID2 influences G1/S transition in leukemic cells through increasing H3K27 trimethylation on the *CCND1* promoter and thereby negatively regulating its expression. Consistent with this, the authors observed that *JARID2* overexpression inhibited leukaemia cell proliferation, and conversely that *JARID2* knockdown promoted proliferation. Taken together, this advocates *JARID2* as a potential tumour suppressor gene in leukaemia (Su et al., 2015).

The lead variant rs34229995 is located 2.2-kB upstream of *JARID2*. *JARID2* maps to 6p22.3-p21.31, a chromosomal region that is commonly somatically amplified in MM cells. Additionally, *JARID2* is crucial in developing embryonic cells and is required for proper neurogenesis and morphogenesis (Corley & Kroll, 2015). A clinically distinct neurodevelopmental syndrome has been reported in 16 individuals with *JARID2* haploinsufficiency (Baroy et al., 2013; Verberne et al., 2021).

The autophagy related 5 gene (***ATG5***) encodes a protein important in autophagy, lymphocyte development and proliferation, and apoptosis among other cellular processes (Ye, Zhou, & Zhang, 2018). In autophagy, ATG5 is crucial for forming autophagic vesicles. Mouse studies where *Atg5* was knocked out at different stages of B cell differentiation, support that ATG5, via autophagy, is required for the survival of pro-B cells and progression to pre-B cells (Miller, Zhao, Stephenson, Cadwell, & Pua, 2008). Although *Atg5*-deficient B cells do not fail to mature, Ig production is impaired in plasma cells through diminished CD138 expression. This implies that ATG5 is important in late stages of B cell differentiation (Conway et al., 2013).

The 6q21 lead variant rs9372120 maps to *ATG5* intron 6. Using Hi-C-data, we detected looping interactions and a TAD at rs9372120 in the lymphoblastic cell line GM12878. The *ATG5* region also showed evidence for intra-chromosome looping with the transcriptional repressor gene *PRDM1*, encoding the protein Blimp-1 that is crucial for irreversible plasma cell differentiation and maintenance of long-lived plasma cells (Calame, Lin, & Tunyaplin, 2003; Martins & Calame, 2008). Analysing MM-RNA sequencing data we could show that *ATG5* and *PRDM1* expression within the TAD was correlated, suggesting coregulation.

The 7q26 lead variant rs7781265 maps to *SMARCD3* intron 2. The SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3 gene (***SMARCD3)*** encodes a protein (SMARCD3 or **BAF60C**) involved in chromatin remodelling and cell cycle progression at the G1/S checkpoint (Tropee et al., 2021). BAF60C is part of a BRG1/BRM-associated factor (BAF) complex,

one of the SWI/SNF complexes, which participates in chromatin remodelling in the nucleosome by recruitment to specific enhancers (Tropee et al., 2021). The BAF complex is important in for example cardio-, neuro- and myogenesis (Kadoch & Crabtree, 2015; X. Sun et al., 2018; Toto, Puri, & Albini, 2016). A variant in the *SMARCD3* region (in an *ABCF2* intron), rs1122979, has also been associated with N-glycosylation of IgG, another plasma cell-related process. (Lauc et al., 2013). This variant is in LD with the MM lead variant (D'/r^2 = 0.90/0.95, using LDpair provided by NIH in the CEU population).

The colon cancer associated transcript 1 (***CCAT1***) does not encode a protein but an oncogenic long non-coding RNA (lncRNA) and has been associated with a long range of cancer tissues including MM (Z. Liu, Chen, & Hann, 2019). *CCAT1* is located at 8q21.24, 515 kb upstream from the *MYC* gene. *CCAT1* mediates the binding between the *MYC* promoter and an upstream enhancer, increasing *MYC* expression (Z. Liu et al., 2019). In MM cells and in hepatocytic cancer cells (HCCs), CCAT1 has been shown to work as a sponge to the microRNA miR-181a-5p thereby inhibiting its regulatory function (L. Chen, Hu, Wang, Zhao, & Gu, 2018; J. Guo, Ma, Peng, Jin, & Liu, 2019). MicroRNAs bind mRNA and can block translation and promote mRNA break-down (Wahid, Shehzad, Khan, & Kim, 2010). In MM cells, *CCAT1* is upregulated and promotes cell cycle progression (L. Chen et al., 2018). Consistent with this, *CCAT1* knockdown induces G0/G1 arrest (L. Chen et al., 2018). While CCAT1-mediated inhibition of miR-181a-5p promotes autophagy in HCCs, CCAT1 has been reported to inhibit autophagy in podocytes (J. Guo et al., 2019).

In our study, we found that the region *CCAT1* maps to, interacts with a region containing *MYC*. We also found looping chromatin interactions and TADs in the region. Interestingly, the MM lead variant rs1948915 is not in LD with risk alleles identified in other cancer types e.g., Hodgkin's Lymphoma and Colorectal Cancer. In addition, the risk allele did not associate with an upregulation of *CCAT1*, nor did it associate with overall survival as has been suggested in more recent studies (L. Chen et al., 2018).

The ***CDKN2A*** gene has been discussed earlier in this thesis. The lead variant rs2811710 maps to intron 1β of *CDKN2A* ($p14^{ARF}$). In this region, 9p21.3, we found looping interactions in GM12878 cells and a TAD in MM plasma cells. Interestingly, the MM risk variant is not in LD with *CDKN2A* risk variants for other cancer types such as lung cancer, breast cancer, melanoma, basal cell carcinoma, glioma, and acute lymphocytic leukaemia (ALL) (Sherborne et al., 2010).

The WW domain containing adaptor with coiled-coil gene (***WAC***) is involved in cell cycle control, genomic stability, and autophagy. The WAC protein is a part of the RNF20/RNF40/WAC-complex, one of the histone H2B ubiquitination E3

ligases that associate with RNA Polymerase II via its WW domain, regulating H2B ubiquitination and transcription (F. Zhang & Yu, 2011).

DNA damage has been reported to trigger transfer of WAC to the *CDKN1A*-locus. Loss of *WAC* has been shown to disrupt p53-dependent transcription of *CDKN1A* (p21). During DNA damage, p53 is activated through post-translational modifications such as ubiquitination, and subsequently targets genes involved in cell cycle arrest and DNA repair (F. Zhang & Yu, 2011). The RNF20/RNF40/WAC complex has recently been reported to directly bind p53 via its several coiled coil regions. The key binding site harbours one of the common sites for amino acid substitutions in p53 (Meng 2021). In mouse B cells, So et al suggest that the monoubiquitination of H2B (H2Bub) by RNF20/RNF40 is required for double-strand break (DSB) repair during class-switch recombination (So, Ramachandran, & Martin, 2019).

In the Golgi apparatus, WAC regulates autophagy in response to amino acid starvation. WAC binds GM130 and thereby promotes GABARAP delivery to the phagophore leading to activation of the ULK initiation complex (Joachim et al., 2015).

The ring finger and WD repeat domain 3 gene (***RFWD3***) encodes an E3 ubiquitin-protein ligase implicated in p53 stabilization and cell cycle control. The protein forms a complex with MDM2 and p53, thereby protecting p53 from MDM2-induced polyubiquitination and subsequent degradation (Fu et al., 2010). This stabilization regulates p53 at the G1/S DNA damage checkpoint. Additionally, RFWD3 is involved in RPA-mediated DNA repair at stalled replication forks (Duan & Pathania, 2020).

The risk variant (rs7193541) located in the *RFWD3* gene, is a missense mutation (I564V), that is predicted to be non-pathogenic by PolyPhen-2 and SIFT (Adzhubei et al., 2010; Sim et al., 2012). In whole blood, we detected a strong meQTL for *RFWD3* methylation attenuating *RFWD3* expression. In later work from the research group, a *RFWD3* plasma cell *cis*-eQTL, was detected for the risk variant (Ajore et al., *Nature Communications*, accepted).

The *RFWD3* gene has been associated with other diseases and malignancies. Our risk variant is in LD with a variant (rs4888262) associated with testicular germ cell tumours, as well as a variant that associated with leukocyte telomere length in a recent GWAS (Chung et al., 2013). A compound heterozygous mutation in *RFWD3* has been implicated in congenital Fanconi Anaemia-like disease in a 12-year-old child (Knies et al., 2017).

The phosphatidylinositol-3,4,5-triphosphate dependent Rac exchange factor 1 gene (***PREX1***) encodes a protein that exchanges bound GDP for free GTP for the RHO family of small GTP-binding proteins (RACs) (Srijakotre et al., 2017). RACs

regulate many cell responses such as cell migration, cell adhesion, reactive oxygen species production and actin cytoskeleton rearrangement (Srijakotre et al., 2017). The lead variant rs6066835 maps to *PREX1* intron 3, and we found that it associated with higher *PREX1* expression in MM plasma cells. *PREX1* is somatically mutated or overexpressed in several other cancer types, including prostate cancer, breast cancer and melanoma (Srijakotre et al., 2017). In various xenograft cancer cells, the encoded protein P-Rex1 associates with increased cell migration, possibly through its direct binding to mTOR, and proliferation through ERK1/2 signalling (Srijakotre et al., 2017).

PAPER III

In Paper III, we analysed 38 fMM cases by whole-exome sequencing and SNP microarrays. We identified cases primarily by cross-linking the Multi-generation Registry and the Swedish Cancer Registry, and subsequently cherry-picked samples from the Multiple Myeloma Biobank. Additional cases were obtained from collaborators.

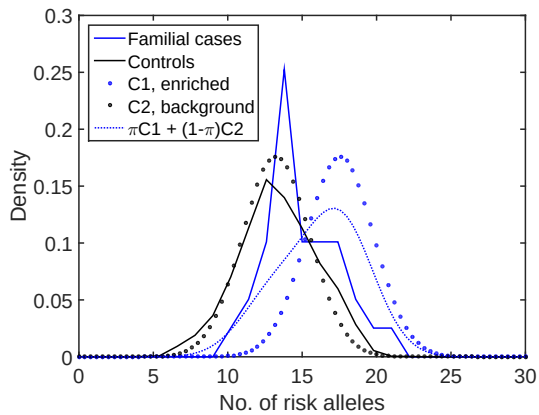


Figure 7. Estimation of the proportion of familial MM cases associated with enrichment of common MM risk alleles and other etiologies, respectively.

While whole-exome sequencing did not reveal any convincing new candidate genes, we observed a significant enrichment of GWAS MM risk allele in familial cases compared with 1,530 sporadic MM cases, as well as with 10,171 population-based controls. For this analysis, we used two different polygenic risk scores (PRSs): one calculated as the sum of risk alleles carried by each individual, and a second score that was calculated as the sum risk alleles weighted by their GWAS effect sizes, as estimated in Paper II.

Furthermore, we hypothesize that some familial cases are caused by an enrichment of common MM risk alleles, whereas others are caused by other factors (e.g., as-yet-unidentified monogenic variants or common environmental exposure). Under this hypothesis, the PRS distribution across a study population of familial cases should follow a composite distribution: cases that are caused by an enrichment of common MM risk alleles should follow PRS distribution shifted to the right; cases that are caused by other factors should follow the background distribution. To estimate the proportion of fMM cases associated with an enrichment of GWAS risk alleles, we used a two-component Gaussian mixture model (Figure 7) and estimated that about one-third of fMM cases follow the PRS distribution to the right, suggesting polygenic aetiology.

PAPER IV

In Paper IV, we identified six new loci associated with MM risk. We used data from Paper II and added additional samples totalling 9,974 MM cases for meta-analysis also in the previous multi-centre collaboration. The novel risk loci map to: *CEP120* (5q23.2), *POT1* (7q31.33), *CCDC71L* (7q22.3), *SP3* (2q31.1), *KLF2* (19p13.11) and *PRR14*/*RNF40* (16p11.2). In addition, a suggestive association was found near *IRF4* (6p25.3), a gene frequently mutated in MM cells (Lohr et al., 2014).

The centrosomal protein 120 gene (***CEP120***) encodes a protein that is important for centriole elongation during mitosis (Mahjoub, Xie, & Stearns, 2010). We found eQTL effects for *CEP120* in $CD138^{+}$ plasma cells, where the risk allele associated with a higher expression. A higher expression of *CEP120* results in aberrant centriole elongation (Comartin et al., 2013).

The protection of telomeres 1 gene (***POT1***) encodes a DNA-binding protein that is a part of several complexes that protect telomeres and regulate their length (C. Li et al., 2020). The shelterin complex is one of these, and is important for maintaining chromosomal stability (Pinzaru et al., 2016; Rice et al., 2017). While *POT1* has not been reported to be recurrently somatically muted in MM, acquisition of somatic mutations are common in B-cell chronic lymphatic leukaemia (CLL) (S. Chang, 2013; Speedy et al., 2014; Speedy et al., 2016).

Germline variants in *POT1* have been associated with several malignancies, including CLL, Hodgkin lymphoma, colorectal cancer, glioma, melanoma, and Li-Fraumeni-like syndrome (Calvete et al., 2017; McMaster et al., 2018; Robles-Espinoza et al., 2014; Speedy et al., 2014; Speedy et al., 2016).

At the 7q22.3 locus, the risk variant maps to an intron of an uncharacterized non-coding RNA *CTB-30L5.1*, that spans over a coding gene, the coiled-coil domain-containing protein 71L gene (***CCDC71L***). The variant is located about 3 kBp downstream of CCDC*71L*. This gene has been shown to be upregulated in triple-

negative breast cancer by the long intergenic non-coding RNA LINC00514, which sequesters inhibitory micro-RNAs (miR's) by sponging (X. Luo & Wang, 2021). The authors suggest an miR-6504-5p/miR-3139/*CCDC71L* axis, where *CCDC71L* expression leads to pro-tumour effects including cell migration and cell growth. LINC00514 has been associated with progression of several other tumour types, for example prostate, gastric, pancreatic, and oesophageal cancer (Han, Li, Xiong, & Song, 2020; Ramnarine et al., 2018; Wang et al., 2021; Yuan et al., 2021).

The Sp3 transcription factor gene (***SP3***) is a transcription factor that is involved in cell-cycle regulation, chromatin remodelling and B cell differentiation, among other processes (L. Li & Davie, 2010; S. R. Park et al., 2009). SP3 is a member of a family of similar transcription factors, the SP/KLF family, and often share DNA binding sites with SP1 (L. Li & Davie, 2010). For example, they both bind the promoter of *CDKN1A* ($p21^{WAF1/CIP1}$) and modulate its expression (Hammill, Jain, Armstrong, & Mueller, 2005). p21 inhibits CDKs in cell cycle progression (Abbas & Dutta, 2009).

In plasma cell development, SP3 has been reported to bind to AID. The AID enzyme (encoded by *AICDA*) is involved in somatic hypermutation and class-switch recombimation and is important for an adequate antibody response (S. R. Park et al., 2009). SP3 also binds the Germinal centre-associated signalling and motility protein gene (*GCSAM*) (Steinke et al., 2004). GCSAM participates in B cell signalling and maturation and is highly expressed in Germinal Centre B cells (Steinke et al., 2004).

The Kruppel like factor 2 gene (***KLF2***) is a transcription factor that is important in B cell differentiation and reactivity (Hart, Wang, Hogquist, & Jameson, 2011). KLF2 is also important T cells by preventing apoptosis in circulating cells and regulating $CD4^+$ and $CD8^+$ T cell quiescence (J. Wu & Lingrel, 2004).

The risk variant rs11086029, is located in the promoter of *KLF2*. Although we could not find an eQTL effect in MM plasma cells, the variant has an eQTL-effect in whole blood in a study about a potentially shared genetic aetiology with myeloproliferative neoplasms (Macauda et al., 2021).

In MM cell lines, *KLF2* knockdown results in apoptosis suggesting that *KLF2* is essential for MM cell survival (Ohguchi et al., 2016). $KLF2^{-/-}$ mouse embryos die in utero from haemorrhage due to defects in the blood vessel wall (Kuo et al., 1997).

KLF2 directly binds IRF4, and together they are regulated by, and regulate, the expression of the Lysine-specific demethylase 3A (*KDM3A*) in a positive feedback-loop. This KDM3A-KLF2-IRF4 axis has been suggested to regulate MM cell adhesion and homing in the bone marrow (Ohguchi et al., 2016). *KDM3A* encodes a histone demethylase, and although it is associated with several other malignancies, its role in MM remains unclear (Sui, Gu, & Janknecht, 2021).

Interestingly, *KLF2* belongs to the same family as *SP3*, and they both bind the promoter of the cyclin-dependent kinase inhibitor gene, *CDKN1A* ($p21^{WAF1/CIP1}$). KLF2 has been found to regulate *CDKN1A* promoter activity and thereby directly regulate its expression (J. Wu & Lingrel, 2004).

The proline rich 14 protein gene (***PRR14***) encodes a protein involved in regulation of heterochromatin attachment to the nuclear lamina throughout the cell cycle (Poleshko et al., 2013). In early mitosis, PRR14 is liberated, and the peripheral heterochromatin is separated from the nuclear lamina (Poleshko et al., 2013). In late mitosis, PRR14 then mediates its reattachment (Poleshko & Katz, 2014). It has been suggested that PRR14 participates in heterochromatin organization, and thereby regulates the mitotic exit in the cell cycle (Poleshko et al., 2013). Yang et al detected an overexpression of *PPR14* in lung cancer and propose that PPR14 binds GRB2, thereby activating the PI3K/AKT/mTOR pathway (Yang, Lewinska, Fan, Zhu, & Yuan, 2016). In addition, PRR14 is suggested to interact with the phosphatase complex PP2A (Herzog et al., 2012). Interestingly, a rare germline missense variant in one of the PP2A subunits has been detected in one MM family (Catalano et al., 2021).

RNF40 has already been discussed. In short, *RNF40* is located in the same haploblock as *PPR14*, encoding a protein required for double-stranded DNA break (DSB) repair by non-homologous end joining (NHEJ) (So et al., 2019).

We identified he interferon regulatory factor 4 gene (***IRF4***) as a suggestive MM association that did not, however, quite reach genome-wide significance ($p = 5.93 \times 10^{-8}$). *IRF4* encodes a transcription factor and is highly expressed in plasma and B cells. IRF4 is important during immunoglobulin class-switching and is essential for plasma cell differentiation and survival (Agnarelli, Chevassut, & Mancini, 2018).

General Discussion and Concluding Remarks

MM represents approximately 1% of malignant diseases globally. Familial MM (fMM) only constitutes 1 to 2% of MM cases, corresponding to six to twelve cases per year in Sweden. Consequently, compiling substantial sets of fMM samples for genetic studies is challenging.

As a result, the genes hitherto suggested as fMM candidate genes, have only been reported in a single or a handful of families. This could suggest that development of fMM is polygenic, that causal monogenic variants are highly variable (private), or that the studies conducted so far have been underpowered.

Whereas mutations detected in fMM are assumed to be rare with large effects, GWASs primarily identify common variants with moderate individual effects. Yet, inheriting several such variants in combination, can confer a substantial risk increase. Consistent with this, we observe enrichment of GWAS variants in familial cases in Paper III and IV.

Identifying the mechanisms that underlie GWAS findings is challenging. Each association is typically represented by multiple (often tens to hundreds) of variants in high LD, often spanning genomic regions containing multiple protein-coding genes. The vast majority of the linked variants are non-coding, and thus likely exert their effects by changing the transcriptional regulation of nearby target genes. Moving forward, an interesting challenge is to delineate causal variants within LD blocks, their target genes, their effects on transcriptional regulatory elements, and their downstream functional effects.

Several genes discussed in this thesis are involved in cell cycle-regulation. Timely progression in the cell cycle is crucial for DNA integrity. Chromosomal aberrations can arise as a result from abnormal centriole duplication, spindle formation, and polarity. Chromatin integrity and remodelling is important in both the cell cycle and in transcription. Epigenetic regulation is important by keeping the relevant chromatin open or wound up on histones to form nucleosomes, guided by histone marks and DNA methylation. DNA repair of double strand breaks is important during class-switch recombination and somatic hypermutation. In cancer cells, many of these processes are aberrant, and somatic mutations in oncogenes or tumour suppressor genes occur, leading to chromosomal instability.

Another subset of the discussed genes contains genes that are important in the cell's degrading systems, including autophagy. Autophagy is important in plasma cells, and its disruption leads to short-lived plasma cells with an overproduction of immunoglobulins at the expense of long-lived plasma cells homing in the bone marrow. Proteasomal degradation is targeted in MM treatment and proteasome inhibitors have improved survival and quality-of-life markedly in MM patients.

Less than 20% of heritability can be attributed to the associated risk variants identified so far by GWAS. However, statistical power to detect additional MM risk variants continues to improve by increasing sample sizes, mainly through international collaborations.

While association studies are important for identification of MM candidate genes, *in silico* fine-mapping and functional studies in particular, are crucial to determine the variants' mechanism of action. This will contribute to a better understanding of MM pathophysiology, and by extension clinical applications.

Populärvetenskaplig sammanfattning

Multipelt Myelom (MM) är en form av blodcancer som diagnosticeras hos drygt 600 personer i Sverige varje år. Trots att behandlingen förbättrats avsevärt de senaste 15 åren är det fortfarande en dödlig sjukdom. MM utgår från s k plasmaceller som producerar de antikroppar vi normalt behöver för att skydda oss från infektioner.

Sedan 1920-talet har det rapporterats att MM ibland uppträder hos flera medlemmar inom samma familj, s k familjärt MM. Man har bedömt att förstagradssläktingar till patienter med MM löper en 2 till 4 gånger ökad risk att själva drabbas. Man har därför misstänkt att en del av orsaken kan finnas i vår arvsmassa.

Bättre och billigare genetiska analyser, har gjort det möjligt att jämföra DNA från stora antal patienter med MM mot friska kontroller. I stora associationsstudier, s k GWAS, undersöker man om vanliga variationer i vårt genom kan förklara en del av den ökade risken att drabbas av MM. Även om vår genetiska kod inte skiljer sig särskilt mycket åt från person till person, bär vi alla på små ”stavfel”, eller åtminstone ”stavningsvarianter”, i vårt DNA. Dessa variationer har oftast uppträtt för många generationer sedan och delas därför oftare inom en släkt eller inom en folkgrupp. De kan påverka vanliga drag som t ex ögon-, och hårfärg, men de kan också påverka nivåer av proteiner som är viktiga för kroppsliga funktioner. Eftersom de är utspridda i genomet kan man använda dem för att märka upp olika delar av arvsmassan. På så sätt kan man försöka identifiera gener i dessa regioner som kan ha betydelse för t ex MM.

I avhandlingens första artikel, identifierade vi varianter i två regioner som var vanligare hos patienter med MM. I den andra och fjärde artikeln samarbetade vi med forskargrupper i Island, Norge, Danmark, Storbritannien, Tyskland, Nederländerna, och USA, och hittade tillsammans åtta respektive sex nya varianter som ökar risken att drabbas av MM. I den tredje artikeln undersökte vi familjärt MM och uppskattade att orsaken i ungefär en tredjedel av fallen skulle kunna vara orsakade av att de har fler riskvarianter än andra patienter med MM och friska kontroller. Detta tyder på att MM, åtminstone i en del av fallen, är en sjukdom som kan orsakas av att flera genvarianter samverkar, en s k polygen sjukdom, och inte en sjukdom som orsakas av en enda stark genvariant.

I och med att vi kan undersöka allt fler patienters arvsmassa upptäcker vi fler och fler gener som kan vara av betydelse för MM. Förhoppningsvis leder detta till en ökad förståelse av orsakerna till MM, vilket på längre sikt skulle kunna leda till bättre möjligheter att förebygga och behandla sjukdomen.

Acknowledgements

Most of you have probably already figured out that I am writing this approximately an hour before handing it in for press, and that I have not slept in 24 hours. Unable to get my mind together, I will go on the fly.

Thank you, Björn Nilsson, my supervisor for (all) these years, it hasn't always been easy, but we have had some good laughs too. And interesting conversations, of course, in many respects.

Thank you, all past and present members of the Hematogenomics group:

From the very start: my co-supervisor Markus Hansson, Ellinor, Mikael, Magnus, Linnea, Mina, Ram. And along the way: Jenny, Ludde, Bhairavi, Zain, Anna-Karin, Laura, Aitzkoa, Maurolitsa, Malte, Evelina, Angelica, Nerea, Cat, Abhishek, (in no planned order), and you who I have forgotten, I'm sorry.

Thank you, all other past and present members of the Hematology Division, including Prof. Urban Gullberg, Calle, Helena, Giorgia, Annarita, Åsa.

Thank you, Johan Flygare for helping me in the end.

Thank you, all nice people in the coffee room, especially those who didn't work simultaneously, and those who put on coffee when I didn't.

Thank you, collaborators, and frenemies. Thank you, Rosalind Franklin, Ada Lovelace, Fredrick Sanger etc. Ok, I'm losing it..

Thank you, my dear friends and (once) fellow AT-läkare Paula and Samuel. You've helped me more than you know the last couple of years.

Thank you, Tove! Finally, I'm here! At the end! It wouldn't have been possible without you and your sound mind and advice. But first and foremost, for being my friend and always being there for me, no questions asked.

To my brothers, for being you, and always supportive. And to my extended family of course.

Thank you, mom and dad, for always believing in me, loving me and supporting me. This wouldn't have been possible without you. And sorry for the blisters you must have had in your ears.

Last but never least, I'm done with my "book", so I'm all yours, now and forever, Jesper, Sigrid, August and Selma. Thank you Jesper, for all the coffee you have provided, especially in (every) morning.

Tonight, we will play the violin, and play the nice games where no one dies in Roblox, if you help me that is.

References

Abbas, T., & Dutta, A. (2009). p21 in cancer: intricate networks and multiple activities. *Nature Reviews: Cancer, 9*(6), 400-414. doi:10.1038/nrc2657

Abreu, I., Goncalves, F., da Silva, L., & Viseu, F. (1992). Multiple myeloma in two sisters. *European Journal of Cancer, 28A*(10), 1774-1775. doi:10.1016/0959-8049(92)90098-m

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods, 7*(4), 248-249. doi:10.1038/nmeth0410-248

Agnarelli, A., Chevassut, T., & Mancini, E. J. (2018). IRF4 in multiple myeloma-Biology, disease and therapeutic target. *Leukemia Research, 72*, 52-58. doi:10.1016/j.leukres.2018.07.025

Alexander, L. L., & Benninghoff, D. L. (1965). Familial multiple myeloma. *Journal of the National Medical Association, 57*(6), 471-475.

Alexander, L. L., & Benninghoff, D. L. (1967). Familial multiple myeloma. II. Final pathological findings in two brothers and a sister. *Journal of the National Medical Association, 59*(4), 278-281.

Ali, M., Ajore, R., Wihlborg, A. K., Niroula, A., Swaminathan, B., Johnsson, E., . . . Nilsson, B. (2018). The multiple myeloma risk allele at 5q15 lowers ELL2 expression and increases ribosomal gene expression. *Nat Commun, 9*(1), 1649. doi:10.1038/s41467-018-04082-2

Andreotti, G., Birmann, B., De Roos, A. J., Spinelli, J., Cozen, W., Camp, N. J., . . . Purdue, M. P. (2013). A pooled analysis of alcohol consumption and risk of multiple myeloma in the international multiple myeloma consortium. *Cancer Epidemiology, Biomarkers and Prevention, 22*(9), 1620-1627. doi:10.1158/1055-9965.EPI-13-0334

Andreotti, G., Birmann, B. M., Cozen, W., De Roos, A. J., Chiu, B. C., Costas, L., . . . Purdue, M. P. (2015). A pooled analysis of cigarette smoking and risk of multiple myeloma from the international multiple myeloma consortium. *Cancer Epidemiology, Biomarkers and Prevention, 24*(3), 631-634. doi:10.1158/1055-9965.EPI-14-1145

Aschard, H., Chen, J., Cornelis, Marilyn C., Chibnik, Lori B., Karlson, Elizabeth W., & Kraft, P. (2012). Inclusion of Gene-Gene and Gene-Environment Interactions Unlikely to Dramatically Improve Risk Prediction for Complex Diseases. *The American Journal of Human Genetics, 90*(6), 962-972. doi:https://doi.org/10.1016/j.ajhg.2012.04.017

Attar, N., & Kurdistani, S. K. (2017). Exploitation of EP300 and CREBBP Lysine Acetyltransferases by Cancer. *Cold Spring Harbor Perspectives in Medicine, 7*(3). doi:10.1101/cshperspect.a026534

Axelsson, U., Bachmann, R., & Hallen, J. (1966). Frequency of pathological proteins (M-components) om 6,995 sera from an adult population. *Acta Medica Scandinavica, 179*(2), 235-247.

Azhar, A., Begum, N. A., & Husain, A. (2021). Nucleotide Pool Imbalance and Antibody Gene Diversification. *Vaccines (Basel), 9*(10). doi:10.3390/vaccines9101050

Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A., & Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet, 12*(11), 745-755. doi:10.1038/nrg3031

Barbieri, D., & Grampa, A. (1972). [Familial multiple myeloma: beta myeloma in 2 sisters]. *Haematologica, 58*(9), 565-566.

Barnum, K. J., & O'Connell, M. J. (2014). Cell cycle regulation by checkpoints. *Methods in Molecular Biology, 1170*, 29-40. doi:10.1007/978-1-4939-0888-2_2

Baroy, T., Misceo, D., Stromme, P., Stray-Pedersen, A., Holmgren, A., Rodningen, O. K., . . . Frengen, E. (2013). Haploinsufficiency of two histone modifier genes on 6p22.3, ATXN1 and JARID2, is associated with intellectual disability. *Orphanet Journal of Rare Diseases, 8*, 3. doi:10.1186/1750-1172-8-3

Barwick, B. G., Gupta, V. A., Vertino, P. M., & Boise, L. H. (2019). Cell of Origin and Genetic Alterations in the Pathogenesis of Multiple Myeloma. *Front Immunol, 10*, 1121. doi:10.3389/fimmu.2019.01121

Basu, S., & Murphy, M. E. (2016). Genetic Modifiers of the p53 Pathway. *Cold Spring Harbor Perspectives in Medicine, 6*(4), a026302. doi:10.1101/cshperspect.a026302

Baudat, F., Imai, Y., & de Massy, B. (2013). Meiotic recombination in mammals: localization and regulation. *Nat Rev Genet, 14*(11), 794-806. doi:10.1038/nrg3573

Berlin, S. O., Odeberg, H., & Weingart, L. (1968). Familial occurrence of M-components. *Acta Medica Scandinavica, 183*(4), 347-350.

Boga, M., Jako, J., Doman, J., Magyar, E., & Konyar, E. (1973). Familial myeloma. *Folia Haematologica: Internationales Magazin für Klinische und Morphologische Blutforschung, 100*(3), 201-212.

Bolli, N., Barcella, M., Salvi, E., D'Avila, F., Vendramin, A., De Philippis, C., . . . Montefusco, V. (2017). Next-generation sequencing of a family with a high penetrance of monoclonal gammopathies for the identification of candidate risk alleles. *Cancer, 123*(19), 3701-3708. doi:10.1002/cncr.30777

Bourguet, C. C., Grufferman, S., Delzell, E., DeLong, E. R., & Cohen, H. J. (1985). Multiple myeloma and family history of cancer. A case-control study. *Cancer, 56*(8), 2133-2139. doi:10.1002/1097-0142(19851015)56:8<2133::aid-cncr2820560842>3.0.co;2-f

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 68*(6), 394-424. doi:10.3322/caac.21492

Broderick, P., Chubb, D., Johnson, D. C., Weinhold, N., Forsti, A., Lloyd, A., . . . Houlston, R. S. (2011). Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. *Nature Genetics, 44*(1), 58-61. doi:10.1038/ng.993

Calame, K. L., Lin, K. I., & Tunyaplin, C. (2003). Regulatory mechanisms that determine the development and function of plasma cells. *Annual Review of Immunology, 21*, 205-230. doi:10.1146/annurev.immunol.21.120601.141138

Calvete, O., Garcia-Pavia, P., Dominguez, F., Bougeard, G., Kunze, K., Braeuninger, A., . . . Benitez, J. (2017). The wide spectrum of POT1 gene variants correlates with multiple cancer types. *European Journal of Human Genetics, 25*(11), 1278-1281. doi:10.1038/ejhg.2017.134

Camp, N. J., Werner, T. L., & Cannon-Albright, L. A. (2008). Familial myeloma. *New England Journal of Medicine, 359*(16), 1734-1735; author reply 1735. doi:10.1056/NEJMc081677

Castleman, M. (1959). CASE RECORDS of the Massachusetts General Hospital; case 45261. *New England Journal of Medicine, 260*(26), 1336-1342. doi:10.1056/NEJM195906252602611

Catalano, C., Paramasivam, N., Blocka, J., Giangiobbe, S., Huhn, S., Schlesner, M., . . . Forsti, A. (2021). Characterization of rare germline variants in familial multiple myeloma. *Blood Cancer J, 11*(2), 33. doi:10.1038/s41408-021-00422-6

Celik, H., Koh, W. K., Kramer, A. C., Ostrander, E. L., Mallaney, C., Fisher, D. A. C., . . . Challen, G. A. (2018). JARID2 Functions as a Tumor Suppressor in Myeloid Neoplasms by Repressing Self-Renewal in Hematopoietic Progenitor Cells. *Cancer Cell, 34*(5), 741-756 e748. doi:10.1016/j.ccell.2018.10.008

Chan, S. H., Chiang, J., & Ngeow, J. (2021). CDKN2A germline alterations and the relevance of genotype-phenotype associations in cancer predisposition. *Hereditary Cancer in Clinical Practice, 19*(1), 21. doi:10.1186/s13053-021-00178-x

Chang, D., Ma, K., Gong, M., Cui, Y., Liu, Z. H., Zhou, X. G., . . . Wang, T. Y. (2010). SLP-2 overexpression is associated with tumour distant metastasis and poor prognosis in pulmonary squamous cell carcinoma. *Biomarkers, 15*(2), 104-110. doi:10.3109/13547500903311910

Chang, S. (2013). Cancer chromosomes going to POT1. *Nature Genetics, 45*(5), 473-475. doi:10.1038/ng.2617

Chatterjee, N., Shi, J., & Garcia-Closas, M. (2016). Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet, 17*(7), 392-406. doi:10.1038/nrg.2016.27

Chen, J. (2016). The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. *Cold Spring Harbor Perspectives in Medicine, 6*(3), a026104. doi:10.1101/cshperspect.a026104

Chen, L., Hu, N., Wang, C., Zhao, H., & Gu, Y. (2018). Long non-coding RNA CCAT1 promotes multiple myeloma progression by acting as a molecular sponge of miR-181a-5p to modulate HOXA1 expression. *Cell Cycle, 17*(3), 319-329. doi:10.1080/15384101.2017.1407893

Christie, D. A., Mitsopoulos, P., Blagih, J., Dunn, S. D., St-Pierre, J., Jones, R. G., . . . Madrenas, J. (2012). Stomatin-like protein 2 deficiency in T cells is associated with altered mitochondrial respiration and defective CD4+ T cell responses. *Journal of Immunology, 189*(9), 4349-4360. doi:10.4049/jimmunol.1103829

Chubb, D., Weinhold, N., Broderick, P., Chen, B., Johnson, D. C., Forsti, A., . . . Goldschmidt, H. (2013). Common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1 influences multiple myeloma risk. *Nature Genetics, 45*(10), 1221-1225. doi:10.1038/ng.2733

Chung, C. C., Kanetsky, P. A., Wang, Z., Hildebrandt, M. A., Koster, R., Skotheim, R. I., . . . Nathanson, K. L. (2013). Meta-analysis identifies four new loci associated with testicular germ cell tumor. *Nature Genetics, 45*(6), 680-685. doi:10.1038/ng.2634

Coleman, E. A., Lynch, H., Enderlin, C., Stewart, B., Thome, S. D., Kennedy, R., . . . Barlogie, B. (2009). Initial report of a family registry of multiple myeloma. *Cancer Nursing, 32*(6), 456-464. doi:10.1097/NCC.0b013e3181a90168

Comartin, D., Gupta, G. D., Fussner, E., Coyaud, E., Hasegan, M., Archinti, M., . . . Pelletier, L. (2013). CEP120 and SPICE1 cooperate with CPAP in centriole elongation. *Current Biology, 23*(14), 1360-1366. doi:10.1016/j.cub.2013.06.002

Comotti, B., Bassan, R., Buzzetti, M., Finazzi, G., & Barbui, T. (1987). Multiple myeloma in a pair of twins. *British Journal of Haematology, 65*(1), 123-124.

Consortium, T. G. P. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature, 491*(7422), 56-65. doi:10.1038/nature11632

Consortium, U. K., Walter, K., Min, J. L., Huang, J., Crooks, L., Memari, Y., . . . Soranzo, N. (2015). The UK10K project identifies rare variants in health and disease. *Nature, 526*(7571), 82-90. doi:10.1038/nature14962

Conte, C., Baird, M. A., Davidson, M. W., & Griffis, E. R. (2018). Spindly is required for rapid migration of human cells. *Biol Open, 7*(5). doi:10.1242/bio.033233

Conte, C., Griffis, E. R., Hickson, I., & Perez-Oliva, A. B. (2018). USP45 and Spindly are part of the same complex implicated in cell migration. *Scientific Reports, 8*(1), 14375. doi:10.1038/s41598-018-32685-8

Conway, K. L., Kuballa, P., Khor, B., Zhang, M., Shi, H. N., Virgin, H. W., & Xavier, R. J. (2013). ATG5 regulates plasma cell differentiation. *Autophagy, 9*(4), 528-537. doi:10.4161/auto.23484

Coop, G., & Przeworski, M. (2007). An evolutionary view of human recombination. *Nat Rev Genet, 8*(1), 23-34. doi:10.1038/nrg1947

Corley, M., & Kroll, K. L. (2015). The roles and regulation of Polycomb complexes in neural development. *Cell and Tissue Research, 359*(1), 65-85. doi:10.1007/s00441-014-2011-9

Crick, F. H., Barnett, L., Brenner, S., & Watts-Tobin, R. J. (1961). General nature of the genetic code for proteins. *Nature, 192*, 1227-1232. doi:10.1038/1921227a0

Crozes-Bony, P., Palazzo, E., Meyer, O., De Bandt, M., & Kahn, M. F. (1995). Familial multiple myeloma. Report of a case in a father and daughter. Review of the literature. *Revue du Rhumatisme. English Edition, 62*(6), 439-445.

Dalerba, P., Cho, R. W., & Clarke, M. F. (2007). Cancer stem cells: models and concepts. *Annual Review of Medicine, 58*, 267-284. doi:10.1146/annurev.med.58.062105.204854

Dewan, A., Liu, M., Hartman, S., Zhang, S. S., Liu, D. T., Zhao, C., . . . Hoh, J. (2006). HTRA1 promoter polymorphism in wet age-related macular degeneration. *Science, 314*(5801), 989-992. doi:10.1126/science.1133807

Dilworth, D., Liu, L., Stewart, A. K., Berenson, J. R., Lassam, N., & Hogg, D. (2000). Germline CDKN2A mutation implicated in predisposition to multiple myeloma. *Blood, 95*(5), 1869-1871.

Duan, H., & Pathania, S. (2020). RPA, RFWD3 and BRCA2 at stalled forks: a balancing act. *Mol Cell Oncol, 7*(6), 1801089. doi:10.1080/23723556.2020.1801089

Duran-Lozano, L., Thorleifsson, G., Lopez de Lapuente Portilla, A., Niroula, A., Went, M., Thodberg, M., . . . Nilsson, B. (2021). Germline variants at SOHLH2 influence multiple myeloma risk. *Blood Cancer J, 11*(4), 76. doi:10.1038/s41408-021-00468-6

Dutta, R., Tiu, B., & Sakamoto, K. M. (2016). CBP/p300 acetyltransferase activity in hematologic malignancies. *Molecular Genetics and Metabolism, 119*(1-2), 37-43. doi:10.1016/j.ymgme.2016.06.013

Dutto, I., Scalera, C., & Prosperi, E. (2018). CREBBP and p300 lysine acetyl transferases in the DNA damage response. *Cellular and Molecular Life Sciences, 75*(8), 1325-1338. doi:10.1007/s00018-017-2717-4

Ekbom, A. (2011). The Swedish Multi-generation Register. *Methods in Molecular Biology, 675*, 215-220. doi:10.1007/978-1-59745-423-0_10

Esteller, M. (2011). Non-coding RNAs in human disease. *Nat Rev Genet, 12*(12), 861-874. doi:10.1038/nrg3074

Evangelou, E., & Ioannidis, J. P. (2013). Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet, 14*(6), 379-389. doi:10.1038/nrg3472

Feinberg, A. P., & Tycko, B. (2004). The history of cancer epigenetics. *Nature Reviews: Cancer, 4*(2), 143-153. doi:10.1038/nrc1279

Frank, C., Fallah, M., Chen, T., Mai, E. K., Sundquist, J., Forsti, A., & Hemminki, K. (2016). Search for familial clustering of multiple myeloma with any cancer. *Leukemia, 30*(3), 627-632. doi:10.1038/leu.2015.279

Fu, X., Yucer, N., Liu, S., Li, M., Yi, P., Mu, J. J., . . . Wang, Y. (2010). RFWD3-Mdm2 ubiquitin ligase complex positively regulates p53 stability in response to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America, 107*(10), 4579-4584. doi:10.1073/pnas.0912094107

Gao, X., Tate, P., Hu, P., Tjian, R., Skarnes, W. C., & Wang, Z. (2008). ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proceedings of the National Academy of Sciences of the United States of America, 105*(18), 6656-6661. doi:10.1073/pnas.0801802105

Gebregziabher, M., Bernstein, L., Wang, Y., & Cozen, W. (2006). Risk patterns of multiple myeloma in Los Angeles County, 1972-1999 (United States). *Cancer Causes and Control, 17*(7), 931-938. doi:10.1007/s10552-006-0030-x

Genin, E. (2020). Missing heritability of complex diseases: case solved? *Human Genetics, 139*(1), 103-113. doi:10.1007/s00439-019-02034-4

Genomes Project, C., Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., . . . McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. *Nature, 467*(7319), 1061-1073. doi:10.1038/nature09534

Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

Gerkes, E. H., de Jong, M. M., Sijmons, R. H., & Vellenga, E. (2007). Familial multiple myeloma: report on two families and discussion of screening options. *Hereditary Cancer in Clinical Practice, 5*(2), 72-78. doi:10.1186/1897-4287-5-2-72

Geschickter, C. F., & Copeland, M. M. (1928). Multiple Myeloma. *Archives of Surgery, 16*, 807-863.

Gilbert, W. (1978). Why genes in pieces? *Nature, 271*(5645), 501. doi:10.1038/271501a0

Giuffre, A., Ong, J., Guadalupe, M., Ravi, H., Visitacion, M., Pabón-Peña, C., . . . Happe, S. (2012). Development of SureSelect Target Capture Methods for Sequencing on the Ion Torrent PGM System. Retrieved from https://www.agilent.com/cs/library/posters/Public/5990-9660EN_1-4-12Low.pdf

Gobbi, P. G., Montecucco, C., Trespi, E., Cavalli, C., Gendarini, A., Bellotti, V., . . . Ascari, E. (1985). Familial myeloma. A new observation in two distinct pairs of siblings. *Haematologica, 70*(4), 324-328.

Goldstone, A. H., Wood, J. K., & Cook, M. K. (1973). Myeloma in mother and daughter. *Acta Haematologica, 49*(3), 176-181. doi:10.1159/000208399

Grass, S., Preuss, K. D., Ahlgrimm, M., Fadle, N., Regitz, E., Pfoehler, C., . . . Pfreundschuh, M. (2009). Association of a dominantly inherited hyperphosphorylated paraprotein target with sporadic and familial multiple myeloma and monoclonal gammopathy of undetermined significance: a case-control study. *Lancet Oncology, 10*(10), 950-956. doi:10.1016/S1470-2045(09)70234-7

Grass, S., Preuss, K. D., Thome, S., Weisenburger, D. D., Witt, V., Lynch, J., . . . Pfreundschuh, M. (2011). Paraproteins of familial MGUS/multiple myeloma target family-typical antigens: hyperphosphorylation of autoantigens is a consistent finding in familial and sporadic MGUS/MM. *Blood, 118*(3), 635-637. doi:10.1182/blood-2011-01-331454

Grosbois, B., Gueguen, M., Fauchet, R., Lebouc, H., Guenot, A., Lancelin, F., . . . Genetet, B. (1986). Multiple myeloma in two brothers. An immunochemical and immunogenetic familial study. *Cancer, 58*(11), 2417-2421. doi:10.1002/1097-0142(19861201)58:11<2417::aid-cncr2820581111>3.0.co;2-s

Grosbois, B., Jego, P., Attal, M., Payen, C., Rapp, M. J., Fuzibet, J. G., . . . Bataille, R. (1999). Familial multiple myeloma: report of fifteen families. *British Journal of Haematology, 105*(3), 768-770. doi:10.1046/j.1365-2141.1999.01415.x

Grossman, L. A., Ownby, F. D., Grossman, M., Kaplan, H. J., & Wolfe, L. K. (1963). Multiple Myeloma in Brothers. *Journal of the Tennessee Medical Association, 56*, 398-400.

Guan, B., Wang, T. L., & Shih Ie, M. (2011). ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. *Cancer Research, 71*(21), 6718-6727. doi:10.1158/0008-5472.CAN-11-1562

Guo, J., Ma, Y., Peng, X., Jin, H., & Liu, J. (2019). LncRNA CCAT1 promotes autophagy via regulating ATG7 by sponging miR-181 in hepatocellular carcinoma. *Journal of Cellular Biochemistry, 120*(10), 17975-17983. doi:10.1002/jcb.29064

Guo, W., Wang, H., Chen, P., Shen, X., Zhang, B., Liu, J., . . . Xiao, X. (2021). Identification and Characterization of Multiple Myeloma Stem Cell-Like Cells. *Cancers, 13*(14). doi:10.3390/cancers13143523

Hall, J. B., & Bush, W. S. (2016). Analysis of Heritability Using Genome-Wide Data. *Curr Protoc Hum Genet, 91*, 1 30 31-31 30 10. doi:10.1002/cphg.25

Hammill, D., Jain, N., Armstrong, S., & Mueller, C. R. (2005). The D-domain of Sp3 modulates its protein levels and activation of the p21(CIP1/WAF1) promoter. *Biochemical and Biophysical Research Communications, 335*(2), 377-384. doi:10.1016/j.bbrc.2005.07.084

Han, Q., Li, J., Xiong, J., & Song, Z. (2020). Long noncoding RNA LINC00514 accelerates pancreatic cancer progression by acting as a ceRNA of miR-28-5p to upregulate Rap1b expression. *Journal of Experimental and Clinical Cancer Research, 39*(1), 151. doi:10.1186/s13046-020-01660-5

Hansen, M., Rubinsztein, D. C., & Walker, D. W. (2018). Autophagy as a promoter of longevity: insights from model organisms. *Nature Reviews: Molecular Cell Biology, 19*(9), 579-593. doi:10.1038/s41580-018-0033-y

Hart, G. T., Wang, X., Hogquist, K. A., & Jameson, S. C. (2011). Kruppel-like factor 2 (KLF2) regulates B-cell reactivity, subset differentiation, and trafficking molecule expression. *Proceedings of the National Academy of Sciences of the United States of America, 108*(2), 716-721. doi:10.1073/pnas.1013168108

Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics, 107*(1), 1-8. doi:10.1016/j.ygeno.2015.11.003

Hemminki, K., Försti, A., & Hansson, M. (2021). Incidence, mortality and survival in multiple myeloma compared to other hematopoietic neoplasms in Sweden up to year 2016. *Scientific Reports, 11*(1), 17272. doi:10.1038/s41598-021-96804-8

Hemminki, K., Forsti, A., Houlston, R., & Sud, A. (2021). Epidemiology, genetics and treatment of multiple myeloma and precursor diseases. *International Journal of Cancer, 149*(12), 1980-1996. doi:10.1002/ijc.33762

Herrell, W. E., Ruff, J. D., & Bayrd, E. D. (1958). Multiple myeloma in siblings. *Journal of the American Medical Association, 167*(12), 1485-1487.

Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., . . . Aebersold, R. (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science, 337*(6100), 1348-1352. doi:10.1126/science.1221483

Hirsch, W., & Schwarz, G. (1959). [Multiple myeloma in siblings (contribution to the diagnosis, clinical aspects and etiology of plasmocytoma)]. *Medizinische Klinik, 54*, 1624-1626.

Hoerter, J. E., & Ellis, S. R. (2021). Biochemistry, Protein Synthesis. In *StatPearls*. Treasure Island (FL).

Horwitz, L. J., Levy, R. N., & Rosner, F. (1985). Multiple myeloma in three siblings. *Archives of Internal Medicine, 145*(8), 1449-1450.

Hubert, D., Thomas, M., Krulik, M., de Gramont, A., Brissaud, P., Sirinelli, A., & Debray, J. (1985). [Familial myeloma. Apropos of a case]. *Revue de Médecine Interne, 6*(4), 447-453. doi:10.1016/s0248-8663(85)80105-3

Huschtscha, L. I., & Reddel, R. R. (1999). p16(INK4a) and the control of cellular proliferative life span. *Carcinogenesis, 20*(6), 921-926. doi:10.1093/carcin/20.6.921

Illumina. (2021). Infinium Global Screening Array-24 Kit. Retrieved from https://www.illumina.com/products/by-type/microarray-kits/infinium-global-screening.html

International HapMap, C., Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., . . . Stewart, J. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature, 449*(7164), 851-861. doi:10.1038/nature06258

Isobe, T., Shaura, K., Itoh, T., & Fujita, T. (1981). Myeloma in siblings. *Nihon Ketsueki Gakkai Zasshi: Journal of Japan Haematological Society, 44*(1), 98-101.

Jain, M., Ascensao, J., & Schechter, G. P. (2009). Familial myeloma and monoclonal gammopathy: a report of eight African American families. *American Journal of Hematology, 84*(1), 34-38. doi:10.1002/ajh.21325

Joachim, J., Jefferies, H. B., Razi, M., Frith, D., Snijders, A. P., Chakravarty, P., . . . Tooze, S. A. (2015). Activation of ULK Kinase and Autophagy by GABARAP Trafficking from the Centrosome Is Regulated by WAC and GM130. *Molecular Cell, 60*(6), 899-913. doi:10.1016/j.molcel.2015.11.018

Johnsen, H. E., Bogsted, M., Schmitz, A., Bodker, J. S., El-Galaly, T. C., Johansen, P., . . . Dybkaer, K. (2016). The myeloma stem cell concept, revisited: from phenomenology to operational terms. *Haematologica, 101*(12), 1451-1459. doi:10.3324/haematol.2015.138826

Johnson, R. C., Nelson, G. W., Troyer, J. L., Lautenberger, J. A., Kessing, B. D., Winkler, C. A., & O'Brien, S. J. (2010). Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics, 11*, 724. doi:10.1186/1471-2164-11-724

Judson, I. R., Wiltshaw, E., & Newland, A. C. (1985). Multiple myeloma in a pair of monozygotic twins: the first reported case. *British Journal of Haematology, 60*(3), 551-554.

Kadoch, C., & Crabtree, G. R. (2015). Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv, 1*(5), e1500447. doi:10.1126/sciadv.1500447

Kanoh, T., Funahashi, H., Okada, T., & Horii, S. (1988). [Familial multiple myeloma]. *Nihon Naika Gakkai Zasshi. Journal of the Japanese Society of Internal Medicine, 77*(5), 740-741. doi:10.2169/naika.77.740

Keskitalo, S., Haapaniemi, E. M., Glumoff, V., Liu, X., Lehtinen, V., Fogarty, C., . . . Varjosalo, M. (2019). Dominant TOM1 mutation associated with combined immunodeficiency and autoimmune disease. *NPJ Genom Med, 4*, 14. doi:10.1038/s41525-019-0088-5

Kim, D., Kim, K. I., & Baek, S. H. (2021). Roles of lysine-specific demethylase 1 (LSD1) in homeostasis and diseases. *Journal of Biomedical Science, 28*(1), 41. doi:10.1186/s12929-021-00737-3

Kinkel, S. A., Galeev, R., Flensburg, C., Keniry, A., Breslin, K., Gilan, O., . . . Blewitt, M. E. (2015). Jarid2 regulates hematopoietic stem cell function by acting with polycomb repressive complex 2. *Blood, 125*(12), 1890-1900. doi:http://dx.doi.org/10.1182/blood-2014-10-603969.

Klingler, W., & Daweke, H. (1985). [Familial plasmocytoma in mother and son]. *Onkologie, 8*(2), 85-87. doi:10.1159/000215624

Knies, K., Inano, S., Ramirez, M. J., Ishiai, M., Surralles, J., Takata, M., & Schindler, D. (2017). Biallelic mutations in the ubiquitin ligase RFWD3 cause Fanconi anemia. *Journal of Clinical Investigation, 127*(8), 3013-3027. doi:10.1172/JCI92069

Kouznetsova, V. L., Tchekanov, A., Li, X., Yan, X., & Tsigelny, I. F. (2019). Polycomb repressive 2 complex-Molecular mechanisms of function. *Protein Science, 28*(8), 1387-1399. doi:10.1002/pro.3647

Kristinsson, S. Y., Bjorkholm, M., Goldin, L. R., Blimark, C., Mellqvist, U. H., Wahlin, A., . . . Landgren, O. (2009). Patterns of hematologic malignancies and solid tumors among 37,838 first-degree relatives of 13,896 patients with multiple myeloma in Sweden. *International Journal of Cancer, 125*(9), 2147-2150. doi:10.1002/ijc.24514

Kuo, C. T., Veselits, M. L., Barton, K. P., Lu, M. M., Clendenin, C., & Leiden, J. M. (1997). The LKLF transcription factor is required for normal tunica media formation and blood vessel stabilization during murine embryogenesis. *Genes and Development, 11*(22), 2996-3006. doi:10.1101/gad.11.22.2996

Kyle, R. A., & Rajkumar, S. V. (2007). Epidemiology of the plasma-cell disorders. *Best Practice & Research: Clinical Haematology, 20*(4), 637-664. doi:10.1016/j.beha.2007.08.001

Kyle, R. A., Therneau, T. M., Rajkumar, S. V., Larson, D. R., Plevak, M. F., Offord, J. R., . . . III., L. J. M. (2006). Prevalence of Monoclonal Gammopathy of Undetermined Significance. *New England Journal of Medicine, 354*, 1362-1369.

Kyle, R. A., Therneau, T. M., Rajkumar, S. V., Offord, J. R., Larson, D. R., Plevak, M. F., & Melton, L. J., 3rd. (2002). A long-term study of prognosis in monoclonal gammopathy of undetermined significance. *New England Journal of Medicine, 346*(8), 564-569. doi:10.1056/NEJMoa01133202

Laffleur, B., & Basu, U. (2019). Biology of RNA Surveillance in Development and Disease. *Trends in Cell Biology, 29*(5), 428-445. doi:10.1016/j.tcb.2019.01.004

Laffleur, B., Lim, J., Zhang, W., Chen, Y., Pefanis, E., Bizarro, J., . . . Basu, U. (2021). Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. *Nature Genetics, 53*(2), 230-242. doi:10.1038/s41588-020-00772-0

LaFramboise, T. (2009). Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Research, 37*(13), 4181-4193. doi:10.1093/nar/gkp552

Laird, N. M., & Lange, C. (2011). *The Fundamentals of Modern Statistical Genetics*.

Lauby-Secretan, B., Scoccianti, C., Loomis, D., Grosse, Y., Bianchini, F., Straif, K., & International Agency for Research on Cancer Handbook Working, G. (2016). Body Fatness and Cancer--Viewpoint of the IARC Working Group. *The New England journal of medicine, 375*(8), 794-798. doi:10.1056/NEJMsr1606602

Lauc, G., Huffman, J. E., Pucic, M., Zgaga, L., Adamczyk, B., Muzinic, A., . . . Rudan, I. (2013). Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *Plos Genetics, 9*(1), e1003225. doi:10.1371/journal.pgen.1003225

Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., . . . Investigators, G. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic Epidemiology, 34*(6), 591-602. doi:10.1002/gepi.20516

Law, M. I. (1976). Familial occurrence of multiple myeloma. *Southern Medical Journal, 69*(1), 46-48. doi:10.1097/00007611-197601000-00016

Leoncini, D. L. (1964). Korngold L: Multiple Myeloma in 2 Sisters. An Immunochemical Study. *Cancer, 17*, 733-737. doi:10.1002/1097-0142(196406)17:6<733::aid-cncr2820170608>3.0.co;2-3

Li, C., Stoma, S., Lotta, L. A., Warner, S., Albrecht, E., Allione, A., . . . Codd, V. (2020). Genome-wide Association Analysis in Humans Links Nucleotide Metabolism to Leukocyte Telomere Length. *American Journal of Human Genetics, 106*(3), 389-404. doi:10.1016/j.ajhg.2020.02.006

Li, L., & Davie, J. R. (2010). The role of Sp1 and Sp3 in normal and cancer cell biology. *Ann Anat, 192*(5), 275-283. doi:10.1016/j.aanat.2010.07.010

Li, Y., Willer, C., Sanna, S., & Abecasis, G. (2009). Genotype imputation. *Annu Rev Genomics Hum Genet, 10*, 387-406. doi:10.1146/annurev.genom.9.081307.164242

Li, Y. Y., Fan, H. Q., Wang, C., Bai, J., & Liu, C. S. (2017). [Brothers with multiple myeloma: two cases report]. *Zhonghua Xue Ye Xue Za Zhi, 38*(6), 557. doi:10.3760/cma.j.issn.0253-2727.2017.06.020

Li, Z., Ji, X., Wang, D., Liu, J., & Zhang, X. (2016). Autophagic flux is highly active in early mitosis and differentially regulated throughout the cell cycle. *Oncotarget, 7*(26), 39705-39718. doi:10.18632/oncotarget.9451

Li, Z., Tian, X., Ji, X., Wang, J., Chen, H., Wang, D., & Zhang, X. (2020). ULK1-ATG13 and their mitotic phospho-regulation by CDK1 connect autophagy to cell cycle. *PLoS Biology, 18*(6), e3000288. doi:10.1371/journal.pbio.3000288

Lin, D. Y., & Zeng, D. (2010). Meta-analysis of genome-wide association studies: no efficiency gain in using individual participant data. *Genetic Epidemiology, 34*(1), 60-66. doi:10.1002/gepi.20435

Liu, D., Zhang, L., Shen, Z., Tan, F., Hu, Y., Yu, J., & Li, G. (2013). Increased levels of SLP-2 correlate with poor prognosis in gastric cancer. *Gastric Cancer, 16*(4), 498-504. doi:10.1007/s10120-013-0232-3

Liu, Z., Chen, Q., & Hann, S. S. (2019). The functions and oncogenic roles of CCAT1 in human cancer. *Biomedicine and Pharmacotherapy, 115*, 108943. doi:10.1016/j.biopha.2019.108943

Lohr, J. G., Stojanov, P., Carter, S. L., Cruz-Gordillo, P., Lawrence, M. S., Auclair, D., . . . Golub, T. R. (2014). Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell, 25*(1), 91-101. doi:10.1016/j.ccr.2013.12.015

Loth, T. S., Perrotta, A. L., Lima, J., Whiteaker, R. S., & Robinson, A. (1991). Genetic aspects of familial multiple myeloma. *Military Medicine, 156*(8), 430-433.

Luo, X., & Wang, H. (2021). LINC00514 upregulates CCDC71L to promote cell proliferation, migration and invasion in triple-negative breast cancer by sponging miR-6504-5p and miR-3139. *Cancer Cell International, 21*(1), 180. doi:10.1186/s12935-021-01875-2

Luo, Z., Lin, C., & Shilatifard, A. (2012). The super elongation complex (SEC) family in transcriptional control. *Nature Reviews: Molecular Cell Biology, 13*(9), 543-547. doi:10.1038/nrm3417

Lynch, H. T., Brand, R. E., Hogg, D., Deters, C. A., Fusaro, R. M., Lynch, J. F., . . . Kern, S. (2002). Phenotypic variation in eight extended CDKN2A germline mutation familial atypical multiple mole melanoma–pancreatic carcinoma–prone families. *Cancer, 94*(1), 84-96. doi:https://doi.org/10.1002/cncr.10159

Lynch, H. T., Ferrara, K., Barlogie, B., Coleman, E. A., Lynch, J. F., Weisenburger, D., . . . Thomé, S. (2008). Familial Myeloma. *New England Journal of Medicine, 359*(2), 152-157.

Lynch, H. T., Watson, P., Tarantolo, S., Wiernik, P. H., Quinn-Laquer, B., Isgur Bergsagel, K., . . . Weisenburger, D. (2005). Phenotypic heterogeneity in multiple myeloma families. *Journal of Clinical Oncology, 23*(4), 685-693. doi:10.1200/JCO.2005.10.126

Macauda, A., Giaccherini, M., Sainz, J., Gemignani, F., Sgherza, N., Sanchez-Maldonado, J. M., . . . Campa, D. (2021). Do myeloproliferative neoplasms and multiple myeloma share the same genetic susceptibility loci? *International Journal of Cancer, 148*(7), 1616-1624. doi:10.1002/ijc.33337

Mahjoub, M. R., Xie, Z., & Stearns, T. (2010). Cep120 is asymmetrically localized to the daughter centriole and is essential for centriole assembly. *Journal of Cell Biology, 191*(2), 331-346. doi:10.1083/jcb.201003009

Maldonado, J. E., & Kyle, R. A. (1974). Familial myeloma. Report of eight families and a study of serum proteins in their relatives. *American Journal of Medicine, 57*(6), 875-884.

Male, D. (2014). *Immunology: an illustrated outline* (Fifth ed.): Garland Science, Taylor & Francis Group, LCC.

Mandema, E., & Wildervanck, L. S. (1954). [Kahler's disease (multiple myeloma) in two sisters]. *Journal de Génétique Humaine, 3*(3), 170-175.

Manigand, G., Sarrazin, A., Demuth, L., Schoen, E., Deparis, M., Ropartz, C., & Rivat, L. (1970). [Familial Kahler's disease. Immunological and karyotype study of a case]. *Presse Medicale, 78*(40), 1741-1743.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature, 461*(7265), 747-753. doi:10.1038/nature08494

Manson, D. I. (1961). Multiple myeloma in sisters. *Scottish Medical Journal, 6*, 188. doi:10.1177/003693306100600502

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., . . . Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature, 437*(7057), 376-380. doi:10.1038/nature03959

Markham, R., & Smith, J. (1951). Structure of ribonucleic acid. *Nature, 168*(4271), 406.

Martin, C. J., & Moorehead, R. A. (2020). Polycomb repressor complex 2 function in breast cancer (Review). *International Journal of Oncology, 57*(5), 1085-1094. doi:10.3892/ijo.2020.5122

Martincic, K., Alkan, S. A., Cheatle, A., Borghesi, L., & Milcarek, C. (2009). Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. *Nature Immunology, 10*(10), 1102-1109. doi:10.1038/ni.1786

Martins, G., & Calame, K. (2008). Regulation and functions of Blimp-1 in T and B lymphocytes. *Annual Review of Immunology, 26*, 133-169. doi:10.1146/annurev.immunol.26.021607.090241

McCombie, W. R., McPherson, J. D., & Mardis, E. R. (2019). Next-Generation Sequencing Technologies. *Cold Spring Harbor Perspectives in Medicine, 9*(11). doi:10.1101/cshperspect.a036798

McCrea, A. P., & Morris, T. C. (1986). Concurrent familial myeloma in Northern Ireland. *Cancer, 58*(2), 394-396. doi:10.1002/1097-0142(19860715)58:2<394::aid-cncr2820580234>3.0.co;2-8

McMaster, M. L., Sun, C., Landi, M. T., Savage, S. A., Rotunno, M., Yang, X. R., . . . Liu, Y. (2018). Germline mutations in Protection of Telomeres 1 in two families with Hodgkin lymphoma. *British Journal of Haematology, 181*(3), 372-377. doi:10.1111/bjh.15203

Meffre, E., Casellas, R., & Nussenzweig, M. C. (2000). Antibody regulation of B cell development. *Nature Immunology, 1*(5), 379-385. doi:10.1038/80816

Meijers, K. A., De Leeu, M. B., & Voormolen-Kalova, M. (1972). The multiple occurrence of myeloma and asymptomatic paraproteinaemia within one family. *Clinical and Experimental Immunology, 12*(2), 185-193.

Meyerding, H. W. (1925). Multiple Myeloma. *Radiology, 5*, 132-146.

Miller, B. C., Zhao, Z., Stephenson, L. M., Cadwell, K., & Pua, H. H. (2008). The autophagy gene ATG5 plays an essential role in B lymphocyte development. *Autophagy, 4*(3), 309-314.

Mitsopoulos, P., Chang, Y. H., Wai, T., Konig, T., Dunn, S. D., Langer, T., & Madrenas, J. (2015). Stomatin-like protein 2 is required for in vivo mitochondrial respiratory chain supercomplex formation and optimal cell function. *Molecular and Cellular Biology, 35*(10), 1838-1847. doi:10.1128/MCB.00047-15

Mohanta, T. K., Mishra, A. K., & Al-Harrasi, A. (2021). The 3D Genome: From Structure to Function. *International Journal of Molecular Sciences, 22*(21). doi:10.3390/ijms222111585

Morgan, G. J., Johnson, D. C., Weinhold, N., Goldschmidt, H., Landgren, O., Lynch, H. T., . . . Houlston, R. S. (2014). Inherited genetic susceptibility to multiple myeloma. *Leukemia, 28*(3), 518-524. doi:10.1038/leu.2013.344

Nadeau, L. A., Magalini, S. I., & Stefanini, M. (1956). Familial multiple myeloma. *A.M.A. Archives of Pathology, 61*(2), 101-106.

Newbold, A., Martin, B. P., Cullinane, C., & Bots, M. (2014). Measuring apoptosis in mammals in vivo. *Cold Spring Harb Protoc, 2014*(11), 1125-1127. doi:10.1101/pdb.top070417

Ni, G., Zeng, J., Revez, J. A., Wang, Y., Zheng, Z., Ge, T., . . . Wray, N. R. (2021). A Comparison of Ten Polygenic Score Methods for Psychiatric Disorders Applied Across Multiple Cohorts. *Biological Psychiatry*. doi:10.1016/j.biopsych.2021.04.018

Ninkovic, S., & Quach, H. (2020). Shaping the Treatment Paradigm Based on the Current Understanding of the Pathobiology of Multiple Myeloma: An Overview. *Cancers, 12*(11). doi:10.3390/cancers12113488

Niroula, A., Urolagin, S., & Vihinen, M. (2015). PON-P2: prediction method for fast and reliable identification of harmful variants. *PloS One, 10*(2), e0117380. doi:10.1371/journal.pone.0117380

Ohguchi, H., Hideshima, T., Bhasin, M. K., Gorgun, G. T., Santo, L., Cea, M., . . . Anderson, K. C. (2016). The KDM3A-KLF2-IRF4 axis maintains myeloma cell survival. *Nat Commun, 7*, 10258. doi:10.1038/ncomms10258

Ohkura, H. (2015). Meiosis: an overview of key differences from mitosis. *Cold Spring Harbor Perspectives in Biology, 7*(5). doi:10.1101/cshperspect.a015859

Oliva, L., & Cenci, S. (2014). Autophagy in plasma cell pathophysiology. *Front Immunol, 5*, 103. doi:10.3389/fimmu.2014.00103

Olujohungbe, A. B., Gledhill, T., & Satchithananathan, G. (2006). Temporal development of myeloma in a syngeneic twin pair. *British Journal of Haematology, 133*(2), 211-212. doi:10.1111/j.1365-2141.2006.06034.x

Ozet, A., Guran, S., & Beksac, M. (2008). Familial multiple myeloma associated with disorders of chronic inflammation: first report from Turkey. *Clinical Lymphoma & Myeloma, 8*(4), 246-248. doi:10.3816/CLM.2008.n.033

Pagliaroli, L., & Trizzino, M. (2021). The Evolutionary Conserved SWI/SNF Subunits ARID1A and ARID1B Are Key Modulators of Pluripotency and Cell-Fate Determination. *Front Cell Dev Biol, 9*, 643361. doi:10.3389/fcell.2021.643361

Parikh, R., Tariq, S. M., Marinac, C. R., & Shah, U. A. (2021). A comprehensive review of the impact of obesity on plasma cell disorders. *Leukemia*. doi:10.1038/s41375-021-01443-7

Park, K. S., Bayles, I., Szlachta-McGinn, A., Paul, J., Boiko, J., Santos, P., . . . Milcarek, C. (2014). Transcription Elongation Factor ELL2 Drives Ig Secretory-Specific mRNA Production and the Unfolded Protein Response. *Journal of Immunology*. doi:10.4049/jimmunol.1401608

Park, S. R., Zan, H., Pal, Z., Zhang, J., Al-Qahtani, A., Pone, E. J., . . . Casali, P. (2009). HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nature Immunology, 10*(5), 540-550. doi:10.1038/ni.1725

Pengo, N., Scolari, M., Oliva, L., Milan, E., Mainoldi, F., Raimondi, A., . . . Cenci, S. (2013). Plasma cells require autophagy for sustainable immunoglobulin production. *Nature Immunology, 14*(3), 298-305. doi:10.1038/ni.2524

Perez-Oliva, A. B., Lachaud, C., Szyniarowski, P., Munoz, I., Macartney, T., Hickson, I., . . . Alessi, D. R. (2015). USP45 deubiquitylase controls ERCC1-XPF endonuclease-mediated DNA damage responses. *EMBO J, 34*(3), 326-343. doi:10.15252/embj.201489184

Pertea, M., Shumate, A., Pertea, G., Varabyou, A., Breitwieser, F. P., Chang, Y. C., . . . Salzberg, S. L. (2018). CHESS: a new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Genome Biology, 19*(1), 208. doi:10.1186/s13059-018-1590-2

Pertesi, M., Vallee, M., Wei, X., Revuelta, M. V., Galia, P., Demangel, D., . . . Dumontet, C. (2019). Exome sequencing identifies germline variants in DIS3 in familial multiple myeloma. *Leukemia*. doi:10.1038/s41375-019-0452-6

Pfeiffer, R. M., Goldin, L. R., Chatterjee, N., Daugherty, S., Hemminki, K., Pee, D., . . . Gail, M. H. (2004). Methods for testing familial aggregation of diseases in population-based samples: application to Hodgkin lymphoma in Swedish registry data. *Annals of Human Genetics, 68*(Pt 5), 498-508. doi:10.1046/j.1529-8817.2003.00111.x

Pinzaru, A. M., Hom, R. A., Beal, A., Phillips, A. F., Ni, E., Cardozo, T., . . . Denchi, E. L. (2016). Telomere Replication Stress Induced by POT1 Inactivation Accelerates Tumorigenesis. *Cell Reports, 15*(10), 2170-2184. doi:10.1016/j.celrep.2016.05.008

Poleshko, A., & Katz, R. A. (2014). Specifying peripheral heterochromatin during nuclear lamina reassembly. *Nucleus, 5*(1), 32-39. doi:10.4161/nucl.28167

Poleshko, A., Mansfield, K. M., Burlingame, C. C., Andrake, M. D., Shah, N. R., & Katz, R. A. (2013). The human protein PRR14 tethers heterochromatin to the nuclear lamina during interphase and mitotic exit. *Cell Reports, 5*(2), 292-301. doi:10.1016/j.celrep.2013.09.024

Preuss, K. D., Pfreundschuh, M., Fadle, N., Regitz, E., Raudies, S., Murwaski, N., . . . Grass, S. (2011). Hyperphosphorylation of autoantigenic targets of paraproteins is due to inactivation of PP2A. *Blood, 118*(12), 3340-3346. doi:10.1182/blood-2011-04-351668

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics, 38*(8), 904-909. doi:10.1038/ng1847

Prideaux, S. M., Conway O'Brien, E., & Chevassut, T. J. (2014). The genetic architecture of multiple myeloma. *Advances in Hematology, 2014*, 864058. doi:10.1155/2014/864058

Puente, C., Hendrickson, R. C., & Jiang, X. (2016). Nutrient-regulated Phosphorylation of ATG13 Inhibits Starvation-induced Autophagy. *Journal of Biological Chemistry, 291*(11), 6026-6035. doi:10.1074/jbc.M115.689646

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*(3), 559-575. doi:10.1086/519795

Ramnarine, V. R., Alshalalfa, M., Mo, F., Nabavi, N., Erho, N., Takhar, M., . . . Collins, C. C. (2018). The long noncoding RNA landscape of neuroendocrine prostate cancer and its clinical implications. *Gigascience, 7*(6). doi:10.1093/gigascience/giy050

Rebel, V. I., Kung, A. L., Tanner, E. A., Yang, H., Bronson, R. T., & Livingston, D. M. (2002). Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. *Proceedings of the National Academy of Sciences of the United States of America, 99*(23), 14789-14794. doi:10.1073/pnas.232568499

Rebuzzini, P., Zuccotti, M., & Garagna, S. (2021). Building Pluripotency Identity in the Early Embryo and Derived Stem Cells. *Cells, 10*(8). doi:10.3390/cells10082049

Regionala Cancercentrum. (2021). Myelom - Nationellt vårdprogram. 3. Retrieved from https://kunskapsbanken.cancercentrum.se/globalassets/cancerdiagnoser/blod-lymfom-myelom/myelom/vardprogram/nationellt-vardprogram-myelom.pdf

Riccardi, C., Di Girolamo, R., & Parziale, A. (1966). Mieloma multiplo familiare. *Progresso Medico, 22*, 822.

Rice, C., Shastrula, P. K., Kossenkov, A. V., Hills, R., Baird, D. M., Showe, L. C., . . . Skordalakes, E. (2017). Structural and functional analysis of the human POT1-TPP1 telomeric complex. *Nat Commun, 8*, 14928. doi:10.1038/ncomms14928

Rigby, P. G., Pratt, P. T., Rosenlof, R. C., & Lemon, H. M. (1968). Genetic relationships in familial leukemia and lymphoma. *Archives of Internal Medicine, 121*(1), 67-70.

Roach, T. G., Lang, H. K. M., Xiong, W., Ryhanen, S. J., & Capelluto, D. G. S. (2021). Protein Trafficking or Cell Signaling: A Dilemma for the Adaptor Protein TOM1. *Front Cell Dev Biol, 9*, 643769. doi:10.3389/fcell.2021.643769

Robbins, R. (1967). Familial multiple myeloma: the tenth reported occurrence. *American Journal of the Medical Sciences, 254*(6), 848-850. doi:10.1097/00000441-196712000-00012

Robles-Espinoza, C. D., Harland, M., Ramsay, A. J., Aoude, L. G., Quesada, V., Ding, Z., . . . Adams, D. J. (2014). POT1 loss-of-function variants predispose to familial melanoma. *Nature Genetics, 46*(5), 478-481. doi:10.1038/ng.2947

Roddie, P. H., Dang, R., & Parker, A. C. (1998). Multiple myeloma in three siblings. *Clinical and Laboratory Haematology, 20*(3), 191-193. doi:10.1046/j.1365-2257.1998.00109.x

Rostoker, G., Uzzan, B., Baumelou, E., & Chapman, A. (1986). [Familial IgG kappa myeloma in a mother and her daughter. Review of the literature]. *Nouvelle Revue Française d'Hématologie, 28*(1), 27-32.

Ryan, K. R., Giles, F., & Morgan, G. J. (2021). Targeting both BET and CBP/EP300 proteins with the novel dual inhibitors NEO2734 and NEO1132 leads to anti-tumor activity in multiple myeloma. *European Journal of Haematology, 106*(1), 90-99. doi:10.1111/ejh.13525

Samy, E. F., Ross, J., Bolton, E., Morris, E. J., & Oliver, S. E. (2015). Variation in incidence and survival by ethnicity for patients with myeloma in England (2002-2008). *Leukemia and Lymphoma, 56*(9), 2660-2667. doi:10.3109/10428194.2014.1003060

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America, 74*(12), 5463-5467. doi:10.1073/pnas.74.12.5463

Santo, L., Liao, L. M., Andreotti, G., Purdue, M. P., & Hofmann, J. N. (2019). Alcohol consumption and risk of multiple myeloma in the NIH-AARP Diet and Health Study. *International Journal of Cancer, 144*(1), 43-48. doi:10.1002/ijc.31648

Santos, P., Arumemi, F., Park, K. S., Borghesi, L., & Milcarek, C. (2011). Transcriptional and epigenetic regulation of B cell development. *Immunologic Research, 50*(2-3), 105-112. doi:10.1007/s12026-011-8225-y

SCB. (2017). Personnumret fyller 70 år. Retrieved from https://www.scb.se/hitta-statistik/artiklar/2017/Personnumret-fyller-70-ar/

SCB. (2021). Flergenerationsregistret. Retrieved from https://www.scb.se/vara-tjanster/bestall-data-och-statistik/bestalla-mikrodata/vilka-mikrodata-finns/individregister/flergenerationsregistret/

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biology, 14*(8), e1002533. doi:10.1371/journal.pbio.1002533

Shadrina, A. S., Zlobin, A. S., Zaytseva, O. O., Klaric, L., Sharapov, S. Z., E, D. P., . . . Tsepilov, Y. A. (2021). Multivariate genome-wide analysis of immunoglobulin G N-glycosylation identifies new loci pleiotropic with immune function. *Human Molecular Genetics, 30*(13), 1259-1270. doi:10.1093/hmg/ddab072

Shah, V., Boyd, K. D., Houlston, R. S., & Kaiser, M. F. (2017). Constitutional mutation in CDKN2A is associated with long term survivorship in multiple myeloma: a case report. *BMC Cancer, 17*(1), 718. doi:10.1186/s12885-017-3715-5

Shastry, B. S. (2002). SNP alleles in human disease and evolution. *Journal of Human Genetics, 47*(11), 561-566. doi:10.1007/s100380200086

Shell, S. A., Martincic, K., Tran, J., & Milcarek, C. (2007). Increased Phosphorylation of the Carboxyl-Terminal Domain of RNA Polymerase II and Loading of Polyadenylation and Cotranscriptional Factors Contribute to Regulation of the Ig Heavy Chain mRNA in Plasma Cells. *The Journal of Immunology, 179*(11), 7663-7673. doi:10.4049/jimmunol.179.11.7663

Sherborne, A. L., Hosking, F. J., Prasad, R. B., Kumar, R., Koehler, R., Vijayakrishnan, J., . . . Houlston, R. S. (2010). Variation in CDKN2A at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nature Genetics, 42*(6), 492-494. doi:10.1038/ng.585

Shoenfeld, Y., Berliner, S., Shaklai, M., Gallant, L. A., & Pinkhas, J. (1982). Familial multiple myeloma. A review of thirty-seven families. *Postgraduate Medical Journal, 58*(675), 12-16. doi:10.1136/pgmj.58.675.12

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research, 40*(Web Server issue), W452-W457. doi:10.1093/nar/gks539

Skatteverket. (2021). Folkbokföringens historia. Retrieved from https://www.skatteverket.se/privat/folkbokforing/attvarafolkbokford/folkbokforingenshistoria.4.18e1b10334ebe8bc80003006.html

Slunecka, J. L., van der Zee, M. D., Beck, J. J., Johnson, B. N., Finnicum, C. T., Pool, R., . . . Ehli, E. A. (2021). Implementation and implications for polygenic risk scores in healthcare. *Hum Genomics, 15*(1), 46. doi:10.1186/s40246-021-00339-y

Snowden, J. A., & Greaves, M. (1995). IgA lambda myeloma presenting concurrently in identical twins with subsequent transformation to 'aggressive phase' in one. *Clinical and Laboratory Haematology, 17*(1), 95-96.

So, C. C., Ramachandran, S., & Martin, A. (2019). E3 Ubiquitin Ligases RNF20 and RNF40 Are Required for Double-Stranded Break (DSB) Repair: Evidence for Monoubiquitination of Histone H2B Lysine 120 as a Novel Axis of DSB Signaling and Repair. *Molecular and Cellular Biology, 39*(8). doi:10.1128/MCB.00488-18

Sobol, H., Vey, N., Sauvan, R., Philip, N., Noguchi, T., & Eisinger, F. (2002). Re: familial multiple myeloma: a family study and review of the literature. *Journal of the National Cancer Institute, 94*(6), 461-462; author reply 463. doi:10.1093/jnci/94.6.461

Socialstyrelsen. (2019). The Swedish Cancer Register. Retrieved from https://www.socialstyrelsen.se/en/statistics-and-data/registers/register-information/swedish-cancer-register/

Speedy, H. E., Di Bernardo, M. C., Sava, G. P., Dyer, M. J., Holroyd, A., Wang, Y., . . . Houlston, R. S. (2014). A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. *Nature Genetics, 46*(1), 56-60. doi:10.1038/ng.2843

Speedy, H. E., Kinnersley, B., Chubb, D., Broderick, P., Law, P. J., Litchfield, K., . . . Houlston, R. S. (2016). Germ line mutations in shelterin complex genes are associated with familial chronic lymphocytic leukemia. *Blood, 128*(19), 2319-2326. doi:10.1182/blood-2016-01-695692

Spengler, G. A., Butler, R., Fischer, C., Ryssel, H. J., Schmid, E., & Siebner, H. (1966). On the question of familial occurrence of paraproteinemia. *Helvetica Medica Acta, 33*(3), 208-219.

Spitz, F., & Furlong, E. E. (2012). Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet, 13*(9), 613-626. doi:10.1038/nrg3207

Srijakotre, N., Man, J., Ooms, L. M., Lucato, C. M., Ellisdon, A. M., & Mitchell, C. A. (2017). P-Rex1 and P-Rex2 RacGEFs and cancer. *Biochemical Society Transactions, 45*(4), 963-977. doi:10.1042/BST20160269

Steinke, J. W., Hodsdon, W., Parenti, S., Ostraat, R., Lutz, R., Borish, L., & Hagman, J. (2004). Identification of an Sp factor-dependent promoter in GCET, a gene expressed at high levels in germinal center B cells. *Molecular Immunology, 41*(12), 1145-1153. doi:10.1016/j.molimm.2004.06.031

Su, C. L., Deng, T. R., Shang, Z., & Xiao, Y. (2015). JARID2 inhibits leukemia cell proliferation by regulating CCND1 expression. *International Journal of Hematology, 102*(1), 76-85. doi:10.1007/s12185-015-1797-x

Sui, Y., Gu, R., & Janknecht, R. (2021). Crucial Functions of the JMJD1/KDM3 Epigenetic Regulators in Cancer. *Molecular Cancer Research, 19*(1), 3-13. doi:10.1158/1541-7786.MCR-20-0404

Sun, F., Ding, W., He, J. H., Wang, X. J., Ma, Z. B., & Li, Y. F. (2015). Stomatin-like protein 2 is overexpressed in epithelial ovarian cancer and predicts poor patient survival. *BMC Cancer, 15*, 746. doi:10.1186/s12885-015-1723-x

Sun, X., Hota, S. K., Zhou, Y. Q., Novak, S., Miguel-Perez, D., Christodoulou, D., . . . Bruneau, B. G. (2018). Cardiac-enriched BAF chromatin-remodeling complex subunit Baf60c regulates gene expression programs essential for heart development and function. *Biol Open, 7*(1). doi:10.1242/bio.029512

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209-249. doi:10.3322/caac.21660

Swallow, D. M. (2003). Genetics of lactase persistence and lactose intolerance. *Annual Review of Genetics, 37*, 197-219. doi:10.1146/annurev.genet.37.110801.143820

Takekuma, N., Kagiomoto, T., Tajiri, M., Horie, T., Iseri, Y., Kawakita, Y., & Migita, T. (1968). IgD and IgG myeloma in siblings. *Acta Haematol Jap, 31*, 468.

Tellier, J., & Nutt, S. L. (2019). Plasma cells: The programming of an antibody-secreting machine. *European Journal of Immunology, 49*(1), 30-37. doi:10.1002/eji.201847517

Thomas, T. F. (1964). Multiple Myeloma in Siblings. *New York State Journal of Medicine, 64*, 2096-2099.

Thompson, J. R., Attia, J., & Minelli, C. (2011). The meta-analysis of genome-wide association studies. *Briefings in Bioinformatics, 12*(3), 259-269. doi:10.1093/bib/bbr020

Thorsteinsdottir, S., Dickman, P. W., Landgren, O., Blimark, C., Hultcrantz, M., Turesson, I., . . . Kristinsson, S. Y. (2018). Dramatically improved survival in multiple myeloma patients in the recent decade: results from a Swedish population-based study. *Haematologica, 103*(9), e412-e415. doi:10.3324/haematol.2017.183475

Tomecki, R., Kristiansen, M. S., Lykke-Andersen, S., Chlebowski, A., Larsen, K. M., Szczesny, R. J., . . . Jensen, T. H. (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J, 29*(14), 2342-2357. doi:10.1038/emboj.2010.121

Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The personal and clinical utility of polygenic risk scores. *Nat Rev Genet, 19*(9), 581-590. doi:10.1038/s41576-018-0018-x

Toto, P. C., Puri, P. L., & Albini, S. (2016). SWI/SNF-directed stem cell lineage specification: dynamic composition regulates specific stages of skeletal myogenesis. *Cellular and Molecular Life Sciences, 73*(20), 3887-3896. doi:10.1007/s00018-016-2273-3

Trevino, V. (2019). Integrative genomic analysis identifies associations of molecular alterations to APOBEC and BRCA1/2 mutational signatures in breast cancer. *Mol Genet Genomic Med, 7*(8), e810. doi:10.1002/mgg3.810

Tropee, R., de la Pena Avalos, B., Gough, M., Snell, C., Duijf, P. H. G., & Dray, E. (2021). The SWI/SNF subunit SMARCD3 regulates cell cycle progression and predicts survival outcome in ER+ breast cancer. *Breast Cancer Research and Treatment, 185*(3), 601-614. doi:10.1007/s10549-020-05997-5

Turesson, I., Linet, M. S., Bjorkholm, M., Kristinsson, S. Y., Goldin, L. R., Caporaso, N. E., & Landgren, O. (2007). Ascertainment and diagnostic accuracy for hematopoietic lymphoproliferative malignancies in Sweden 1964-2003. *International Journal of Cancer, 121*(10), 2260-2266. doi:10.1002/ijc.22912

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., . . . Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers, 1*(1), 59. doi:10.1038/s43586-021-00056-9

Unakami, H., Kikuta, Y., Ohnami, K., Umetsu, Y., Itoh, K., Kumagai, H., . . . Komatsu, M. (1987). [IgG (kappa) benign monoclonal gammopathy, IgG (kappa) multiple myeloma and Bence Jones (lambda) multiple myeloma in 3 brothers]. *Nihon Naika Gakkai Zasshi. Journal of the Japanese Society of Internal Medicine, 76*(7), 1087-1091. doi:10.2169/naika.76.1087

Várkonyi, J., Farkas, P., Tamáska, J., Masszi, T., Gopcsa, L., Padányi, Á., & Rajczy, K. (2009). Familial multiple myeloma. Two more families. *Central European Journal of Medicine, 4*(4), 501-505. doi:10.2478/s11536-009-0088-4

Varley, J. M. (2003). Germline TP53 mutations and Li-Fraumeni syndrome. *Human Mutation, 21*(3), 313-320. doi:10.1002/humu.10185

Verberne, E. A., Goh, S., England, J., van Ginkel, M., Rafael-Croes, L., Maas, S., . . . Campeau, P. M. (2021). JARID2 haploinsufficiency is associated with a clinically distinct neurodevelopmental syndrome. *Genetics in Medicine, 23*(2), 374-383. doi:10.1038/s41436-020-00992-z

Wahid, F., Shehzad, A., Khan, T., & Kim, Y. Y. (2010). MicroRNAs: Synthesis, mechanism, function, and recent clinical trials. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1803*(11), 1231-1243. doi:https://doi.org/10.1016/j.bbamcr.2010.06.013

Waller, R. G., Darlington, T. M., Wei, X., Madsen, M. J., Thomas, A., Curtin, K., . . . Camp, N. J. (2018). Novel pedigree analysis implicates DNA repair and chromatin remodeling in multiple myeloma risk. *Plos Genetics, 14*(2), e1007111. doi:10.1371/journal.pgen.1007111

Wang, X., Liu, H., Zhang, Q., Zhang, X., Qin, Y., Zhu, G., . . . Fan, R. (2021). LINC00514 promotes lipogenesis and tumor progression in esophageal squamous cell carcinoma by sponging miR378a5p to enhance SPHK1 expression. *International Journal of Oncology, 59*(5). doi:10.3892/ijo.2021.5266

Watanabe, T., Suzuki, H., & Unakami, H. (1988). [IgG-K benign monoclonal gammopathy, IgG-K multiple myeloma and Bence Jones-L multiple myeloma in three brothers]. *Rinsho Byori. Japanese Journal of Clinical Pathology, 36*(2), 233-236.

Watanabe, T., Suzuki, Y., Murakami, S., & Komatsu, M. (1999). [Multiple myeloma in siblings]. *Rinsho Ketsueki. Japanese Journal of Clinical Hematology, 40*(2), 135-139.

Watson, J. D., & Crick, F. H. (1953). Genetical implications of the structure of deoxyribonucleic acid. *Nature, 171*(4361), 964-967. doi:10.1038/171964b0

Weale, M. E. (2010). Quality control for genome-wide association studies. *Methods in Molecular Biology, 628*, 341-372. doi:10.1007/978-1-60327-367-1_19

Wei, X., Calvo-Vidal, M. N., Chen, S., Wu, G., Revuelta, M. V., Sun, J., . . . Lipkin, S. M. (2018). Germline Lysine-Specific Demethylase 1 (LSD1/KDM1A) Mutations Confer Susceptibility to Multiple Myeloma. *Cancer Research, 78*(10), 2747-2759. doi:10.1158/0008-5472.CAN-17-1900

Weinhold, N., Johnson, D. C., Chubb, D., Chen, B., Forsti, A., Hosking, F. J., . . . Hemminki, K. (2013). The CCND1 c.870G>A polymorphism is a risk factor for t(11;14)(q13;q32) multiple myeloma. *Nature Genetics, 45*(5), 522-525. doi:10.1038/ng.2583

Wellcome Trust Case Control, C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447*(7145), 661-678. doi:10.1038/nature05911

Went, M., Sud, A., Law, P. J., Johnson, D. C., Weinhold, N., Forsti, A., . . . Houlston, R. S. (2017). Assessing the effect of obesity-related traits on multiple myeloma using a Mendelian randomisation approach. *Blood Cancer J, 7*(6), e573. doi:10.1038/bcj.2017.48

West, S. L., Gartlehner, G., & Mansfield, A. J. (2010). *Comparative Effectiveness Review Methods: Clinical Heterogeneity [Internet]*: Rockville (MD): Agency for Healthcare Research and Quality (US).

Widom, J. (1998). Structure, dynamics, and function of chromatin in vitro. *Annual Review of Biophysics and Biomolecular Structure, 27*, 285-327. doi:10.1146/annurev.biophys.27.1.285

Wiedermann, D., Urban, P., Wiedermann, B., & Cidl, K. (1976). Multiple myeloma in two brothers. Immunoglobulin levels among their relatives. *Neoplasma, 23*(2), 197-207.

Wiernik, P. H., Wickramasinghe, D., & Dutcher, J. P. (2015). Families with both Hodgkin lymphoma and multiple myeloma in their pedigrees. *Clinical Advances in Hematology & Oncology, 13*(4), 257-260.

Willems, P. M., Kuypers, A. W., Meijerink, J. P., Holdrinet, R. S., & Mensink, E. J. (1993). Sporadic mutations of the p53 gene in multiple myeloma and no evidence for germline mutations in three familial multiple myeloma pedigrees. *Leukemia, 7*(7), 986-991.

Wittke-Thompson, J. K., Pluzhnikov, A., & Cox, N. J. (2005). Rational inferences about departures from Hardy-Weinberg equilibrium. *American Journal of Human Genetics, 76*(6), 967-986. doi:10.1086/430507

Wu, J., & Lingrel, J. B. (2004). KLF2 inhibits Jurkat T leukemia cell growth via upregulation of cyclin-dependent kinase inhibitor p21WAF1/CIP1. *Oncogene, 23*(49), 8088-8096. doi:10.1038/sj.onc.1207996

Wu, J. N., & Roberts, C. W. (2013). ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discovery, 3*(1), 35-43. doi:10.1158/2159-8290.CD-12-0361

Xu, S., & Tang, C. (2021). The Role of ARID1A in Tumors: Tumor Initiation or Tumor Suppression? *Frontiers in Oncology, 11*, 745187. doi:10.3389/fonc.2021.745187

Yamada, Y., Yasukochi, Y., Kato, K., Oguri, M., Horibe, H., Fujimaki, T., . . . Sakuma, J. (2018). Identification of 26 novel loci that confer susceptibility to early-onset coronary artery disease in a Japanese population. *Biomed Rep, 9*(5), 383-404. doi:10.3892/br.2018.1152

Yang, M., Lewinska, M., Fan, X., Zhu, J., & Yuan, Z. M. (2016). PRR14 is a novel activator of the PI3K pathway promoting lung carcinogenesis. *Oncogene, 35*(42), 5527-5538. doi:10.1038/onc.2016.93

Ye, X., Zhou, X. J., & Zhang, H. (2018). Exploring the Role of Autophagy-Related Gene 5 (ATG5) Yields Important Insights Into Autophagy in Autoimmune/Autoinflammatory Diseases. *Front Immunol, 9*, 2334. doi:10.3389/fimmu.2018.02334

Yi, Z., Ouyang, J., Sun, W., Xiao, X., Li, S., Jia, X., . . . Zhang, Q. (2019). Biallelic mutations in USP45, encoding a deubiquitinating enzyme, are associated with Leber congenital amaurosis. *Journal of Medical Genetics, 56*(5), 325-331. doi:10.1136/jmedgenet-2018-105709

Yuan, L., Li, J., Yang, Y., Chen, Y., Bu, Y., Ye, M., . . . Nan, Y. (2021). LINC00514 promotes gastric cancer cell growth and EMT progression via miR-204-3p/KRAS. *Aging, 13*(8), 12007-12015. doi:10.18632/aging.202905

Zawadzki, Z. A., Aizawa, Y., Kraj, M. A., Haradin, A. R., & Fisher, B. (1977). Familial immunopathies: report of nine families and survey of literature. *Cancer, 40*(5), 2094-2101. doi:10.1002/1097-0142(197711)40:5<2094::aid-cncr2820400518>3.0.co;2-5

Zepeda, V. H., & Dominguez, V. J. (2007). Familial multiple myeloma: first case reported in a Mexican family. *Annals of Hematology, 86*(7), 527-528. doi:10.1007/s00277-006-0221-2

Zhang, F., & Yu, X. (2011). WAC, a functional partner of RNF20/40, regulates histone H2B ubiquitination and gene transcription. *Molecular Cell, 41*(4), 384-397. doi:10.1016/j.molcel.2011.01.024

Zhang, L., Ding, F., Cao, W., Liu, Z., Liu, W., Yu, Z., . . . Liu, Z. (2006). Stomatin-like protein 2 is overexpressed in cancer and involved in regulating cell growth and cell adhesion in human esophageal squamous cell carcinoma. *Clinical Cancer Research, 12*(5), 1639-1646. doi:10.1158/1078-0432.CCR-05-1858