

Department of Molecular Medicine and Surgery
Karolinska Institutet, Stockholm, Sweden

NOVEL GENETIC CAUSES OF CHILDHOOD CANCER PREDISPOSITION

Carolina Maya González



**Karolinska
Institutet**

Stockholm 2025

Cover art and illustration by my little sister, Andrea Maya González, & *Openart.ai*.

Figures 1–3, 5–14, 16–21 in this thesis were created with *Biorender.com*. The graph in eFigure 1 in Paper I; the graphical abstract and Figure 4A in Paper VI; and the graphs in Figure 3, Figure 4, and supplementary Figures S5, S6, S9 and S10 in Paper VI were created with *Biorender.com*.

The graph in Figure 4 in this thesis and the graph in eFigure 2 in Paper I were created with GraphPad Prism version 10.4.1.

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.

Printed by Universitetsservice US-AB, 2025.

© Carolina Maya González, 2025

ISBN: 978-91-8017-470-1

DOI: <https://doi.org/10.69622/28194191>

Novel genetic causes of childhood cancer predisposition

Thesis for Doctoral Degree (Ph.D.)

By

Carolina Maya González

The thesis will be defended in public at Eva & Georg Klein Lecture Hall, Karolinska Institutet, Stockholm, on March 21, 2025, at 9:00 AM.

Principal Supervisor:

Professor Ann Nordgren
Karolinska Institutet
Department of Molecular Medicine and Surgery

Co-supervisor(s):

Bianca Tesi, MD, Ph.D.
Karolinska Institutet
Department of Molecular Medicine and Surgery

Professor David Gisselsson Nord
Lund University
Department of Clinical Genetics

Docent Fulya Taylan
Karolinska Institutet
Department of Molecular Medicine and Surgery

Jesper Eisfeldt, Ph.D.
Karolinska Institutet
Department of Molecular Medicine and Surgery

Opponent:

Associate Professor Robert Martin Frans Wolthuis
Amsterdam University Medical Centers
Department of Human Genetics

Examination Board:

Professor Catharina Larsson
Karolinska Institutet
Department of Oncology-pathology

Professor Klas Blomgren
Karolinska Institutet
Department of Women's and Children's Health

Docent Maria K Haanpää
Helsinki University Hospital
Department of Clinical Genetics

*To my family, who encouraged
me to follow my dreams*

Popular science summary of the thesis

Cancer develops when cells become wired to divide uncontrollably. For this to happen, many changes take place in the genome of the cell; a process that can take a long time. Therefore, cancer risk increases with age. Why do children develop cancer then? The answer to this question is complex. However, one of the answers relates to their genome.

About 10% of pediatric patients with cancer have inherited genomic changes that create a shortcut for cancer development. As a result, the likelihood of developing certain tumors increases considerably, leading to conditions known as Cancer Predisposition Syndromes (CPS). Knowing that a patient has a CPS is important for oncologists, as it can result in changes in treatment, patient monitoring, and genetic counseling. However, CPS diagnosis can be tricky. On the one hand, there are more than 150 known CPS, each with their own cancer risk spectrum. On the other hand, there is still important information missing about many CPS.

My thesis focused on the analysis of CPS using cutting-edge genomic methods, molecular and cell biology techniques, and register-based studies.

In **Study I**, we compiled a broad childhood CPS research panel with 881 genes, based on multiple relevant sources of information.

In **Studies II – IV**, we reviewed the literature and presented the clinical and genetic characteristics of children with rare diseases and cancer, to understand the role of these congenital syndromes in cancer development.

In **Studies II and V**, we used Swedish nationwide registries to determine the occurrence of cancer in selected rare diseases, where an association to cancer development was suspected.

In **Study VI**, we searched the genome for the cause of a novel disease including immune deficiency, developmental delay, and leukemia. Then, we studied the cellular mechanisms leading to disease.

The ultimate goal of this thesis was to contribute to a better understanding of inherited childhood cancer predisposition.

Resumen divulgativo de la tesis

El cáncer se desarrolla a partir de un proceso celular de división descontrolada. Usualmente, para que esto suceda, son necesarios múltiples cambios en el genoma. Por esto, el desarrollo de cáncer suele llevar mucho tiempo, y el riesgo aumenta con la edad. ¿Entonces, por qué existe el cáncer en niños? Si bien esta pregunta es compleja, una de las respuestas está asociada con su genoma.

Se estima que alrededor del 10% de los pacientes con cáncer pediátrico heredan cambios genéticos que crean un atajo para la evolución del cáncer. El resultado es que la probabilidad de desarrollar ciertos tumores en estos pacientes aumenta considerablemente, dando lugar a Síndromes de Predisposición al Cáncer (CPS, por sus siglas en inglés). El diagnóstico de CPS puede tener connotaciones clínicas importantes, como cambios en el tratamiento, seguimiento del paciente y asesoramiento genético. Sin embargo, el diagnóstico de CPS es complicado. Por un lado, puesto que hay más de 150 síndromes conocidos, cada uno asociado a diferentes tumores. Por otra parte, ya que aún nos falta conocimiento importante sobre muchos CPS.

Esta tesis de doctorado se centró en el estudio de CPS, utilizando técnicas avanzadas de genómica, registros nacionales suecos, biología molecular y celular.

En el **Proyecto I**, compilamos un panel para investigación con 881 genes posiblemente asociados con CPS pediátrico, basado en múltiples fuentes de información.

En los **Proyectos II – IV**, hicimos una revisión literaria y caracterizamos a niños con enfermedades raras y cáncer, con el objetivo de entender la asociación de estos síndromes congénitos con el desarrollo de cáncer.

En los **Proyectos II y V**, utilizamos registros nacionales suecos para determinar la incidencia de cáncer en pacientes con ciertas enfermedades raras, en las que se sospechaba una asociación con el desarrollo de cáncer.

En el **Proyecto VI**, encontramos la causa genética de una nueva enfermedad congénita con inmunodeficiencia, retraso del desarrollo y leucemia. Luego, estudiamos los mecanismos celulares y moleculares que causan la enfermedad.

El objetivo final de esta tesis fue contribuir a un mejor entendimiento de la predisposición hereditaria al cáncer infantil.

Populärvetenskaplig sammanfattning

Cancer utvecklas när celler delar sig okontrollerat. För att detta ska hända, krävs många förändringar i genomet. Denna process tar relativt lång tid, och därför ökar cancerrisken i regel med åldern. Varför utvecklar barn cancer då? Svaret på denna fråga är komplex, men ett av svaren finns i deras genom.

Cirka 10% av alla barn med cancer har medfödda genetiska förändringar som skapar en genväg för cancerutveckling. Detta resulterar i en ökad sannolikhet för att utveckla vissa tumörer, vilket leder till ett tillstånd som kallas för syndrom med cancerpredisposition (CPS, för dess förkortning på engelska). CPS diagnosen är viktig, eftersom den kan leda till förändringar i behandling, kontrollprogram och genetisk vägledning. Men det är komplicerat att diagnosticera CPS. Å ena sidan finns det mer än 150 kända CPS, där varje diagnos har sitt cancerriskspektrum. Å andra sidan saknar vi kunskap om många CPS.

Min avhandling fokuserade på att studera CPS med hjälp av modern banbrytande sekvenseringsteknik, molekylär- och cellbiologiska metoder samt registerbaserade studier.

I **Studie I**, skapade vi en bred genpanel innehållande 881 gener associerade med barncancerpredisposition, baserade på flera informationskällor.

I **Studie II – IV**, granskade vi litteraturen och presenterade klinisk bild och genetiska egenskaper hos individer med sällsynta sjukdomar som också utvecklat cancer under barndomen, med syftet att förstå sambanden mellan dessa syndrom och cancerutveckling.

I **studie II** och **V** använde vi svenska nationella myndighetsregister för att fastställa förekomsten av cancer vid utvalda sällsynta sjukdomar, som vi misstänkte var associerade med cancerpredisposition.

I **Studie VI** sökte vi i genomet och hittade den genetiska orsaken till en ny sjukdom hos ett barn med immunbrist, intellektuell funktionsnedsättning och leukemi. Därefter utforskade vi de cellulära och molekylära mekanismer som leder till sjukdomen.

Denna avhandling hade som slutligt mål att bidra till en bättre förståelse av ärftlig barncancerpredisposition.

Abstract

Childhood cancer predisposition syndromes (CPS) refer to rare diseases increasing the risk of developing pediatric cancer. Genomic studies have estimated that about 8-18% of children with cancer carry pathogenic variants in CPS genes. This broad diagnostic yield is caused by differences in study designs such as patient inclusion criteria, sequencing methods, number of genes analyzed, and the definition of positive findings. Moreover, the diagnosis of CPS can have important clinical implications for patients including adjusted diagnostic procedures, treatment, surveillance and genetic counselling.

With the ultimate goal of increasing our knowledge on pediatric cancer predisposition, my thesis focused on the study of childhood CPS, using cross-disciplinary methods including register-based, genetic and molecular studies.

In **Study I**, we compiled a broad pediatric CPS research panel with 881 genes and developed a ranking system that prioritizes genes with established or suspected evidence for their association with childhood cancer predisposition. This panel can be used as a tool for the discovery of known and novel childhood CPS in massively parallel sequencing studies of large pediatric cancer cohorts.

In **Studies II – IV**, we report the occurrence of cancer in congenital syndromes with previously unknown cancer associations. Specifically, we describe multiple ovarian tumors in a 13-year-old girl with Prader-Willi syndrome (**Study II**), neuroblastoma in two female patients with Marfan syndrome (**Study III**), and a soft tissue sarcoma in a 17-year-old boy with Limb-girdle Muscular Dystrophy Recessive 1 (**Study IV**). Through these reports, we encourage further studies about the possible implication of these rare diseases in cancer development.

In **Studies II and V**, we used the Swedish national registries to determine the cancer risk spectrum in some of the congenital syndromes mentioned above. In patients with Prader-Willi syndrome (**Study II**), although we did not find an increased risk of cancer overall, we observed a high frequency of pediatric cancer. Moreover, the low number of pediatric patients with cancer precluded further statistical testing. In individuals with muscular dystrophy (**Study V**), we found an increased risk of pediatric astrocytomas and other gliomas, as well as an increased risk of adult pancreatic and nonthyroid endocrine tumors. In myotonic dystrophy (**Study V**), pediatric patients had an increased risk of brain tumors, while adults presented an overall increased cancer risk, explained by various malignancies.

In **Study VI**, we identified a homozygous variant in the *FLCN* gene as the genetic cause of a novel multisystemic syndrome in a boy with global developmental delay, short stature, severe immunodeficiency, and leukemia at 1-year of age. We showed that the *FLCN* p.G15S variant leads to nuclear retention of TFE3/TFEB, resulting in altered expression of genes involved in the lysosomal biogenesis and autophagy pathways. Further, we hypothesize that the *FLCN* p.G15S variant is hypomorphic, leading to this rare autosomal recessive syndrome.

All in all, this thesis aimed to contribute to a better understanding of childhood CPS.

List of constituent scientific papers in this thesis

- I. **Assembling a gene panel for the discovery of novel pediatric cancer predisposition syndromes.**
Maya-González C, Tesi B, Poluha A, Lagerstedt-Robinson K, Nordgren A[#], Taylan F[#]. *Submitted for publication*.
- II. **Register-based and genetic studies of Prader-Willi syndrome show a high frequency of gonadal tumors and a possible mechanism for tumorigenesis through imprinting relaxation.**
Maya-González C, Wessman S, Lagerstedt-Robinson K, Taylan F, Tesi B, Kuchinskaya E, McCluggage WG, Poluha A, Holm S, Nergårdh R, Díaz De Ståhl T, Höybye C, Tettamanti G, Delgado-Vega AM, Skarin Nordenvall A, Nordgren A. *Frontiers in Medicine*, 2023. DOI: 10.3389/fmed.2023.1172565.
- III. **Occurrence of cancer in Marfan syndrome: Report of two females with neuroblastoma and review of the literature.**
Maya-González C, Delgado-Vega AM, Taylan F, Lagerstedt Robinson K, Hansson L, Pal N, Fagman H, Puls F, Wessman S, Stenman J, Georgantzi K, Fransson S, Díaz De Ståhl T, Ek T, Palmer R, Tesi B, Kogner P[#], Martinsson T[#], Nordgren A[#]. *American Journal of Medical Genetics Part A*, 2024. DOI: 10.1002/ajmg.a.63812.
- IV. **Pediatric Soft Tissue Sarcoma in Limb-Girdle Muscular Dystrophy: Molecular Findings and Clinical Implications.**
Maya-González C, Díaz De Ståhl T, Wessman S, Taylan F, Tesi B, Lagerstedt-Robinson K, Tettamanti G, Dukic M, Poluha A, Ljungman G, Nordgren A. *American Journal of Case Reports*, 2024. DOI: 10.12659/AJCR.945715.
- V. **Cancer Risk in Patients With Muscular Dystrophy and Myotonic Dystrophy: A Register-Based Cohort Study.**
Maya-González C^{*}, Tettamanti G^{*}, Taylan F, Skarin Nordenvall A, Sejersen T, Nordgren A. *Neurology*, 2024. DOI: 10.1212/WNL.0000000000209883.
- VI. **Biallelic variants in the *FLCN* gene lead to a novel syndrome with global developmental delay, short stature and immunodeficiency.**
Maya-González C, Boeckemeier L, Ten Berk de Boer E, Eisfeldt J, Pozzani F, Mhashal A, Campbell T, Bobeck J, Ekholm K, Bryceson Y, Orellana L, Lindqvist A, Nordgren A[#], Taylan F[#]. *In manuscript*.

* Shared first authors.

[#] Shared senior authors.

Related scientific papers

The following paper is referred to as **Study VII** across this thesis.

VII. Diagnostic yield and clinical impact of germline sequencing in children with CNS and extracranial solid tumors—a nationwide, prospective Swedish study.

Tesi B*, Lagerstedt Robinson K*, Abel F*, Díaz de Ståhl T, Orrsjö S, Poluha A, Hellberg M, Wessman S, Samuelsson S, Frisk T, Vogt H, Henning K, Sabel M, Ek T, Pal N, Nyman P, Giraud G, Wille J, Pronk CJ, Norén-Nyström U, Borssén M, Fili M, Stålhammar G, Herold N, Tettamanti G, Maya-Gonzalez C, Arvidsson L, Rosén A, Ekholm K, Kuchinskaya E, Hallbeck AL, Nordling M, Palmebäck P, Kogner P, Kanter Smoler G, Lähteenmäki P, Fransson S, Martinsson T, Shamik A, Mertens F, Rosenquist R, Wirta V, Tham E, Grillner P, Sandgren J, Ljungman G#, Gisselsson D#, Taylan F#, Nordgren A#. *The Lancet Regional Health – Europe*, 2024. DOI: 10.1016/j.lanep.2024.100881.

Additional scientific papers

VIII. Genetic testing for childhood cancer predisposition syndromes: Controversies and recommendations from the SIOPE Host Genome Working Group meeting 2022.

Bakhuizen JJ, Bourdeaut F, Wadt KAW, Kratz CP, Jongmans MCJ, Waespe N, SIOPE Host Genome Working Group. *EJC Paediatric Oncology*, 2024. DOI: 10.1016/j.ejcped.2024.100176.

IX. Pushing the boundaries of rare disease diagnostics with the help of the first Undiagnosed Hackathon.

Delgado-Vega AM*, Cederroth H*, Taylan F*, (...) Maya-González C, (...) Buske OJ#, Mikk Cederroth M#, Nordgren A#. *Nature Genetics*, 2024. DOI: 10.1038/s41588-024-01941-1.

* Shared first authors.

#Shared senior authors.

Table of contents

1	Background	5
1.1	The history of human genetics	5
1.2	The human genome	7
1.3	Human genetic variation and disease	7
1.3.1	Effects of DNA changes in proteins	8
1.3.2	Rare diseases	9
1.4	Cancer – a genetic disease of somatic cells	10
1.4.1	Childhood cancer	12
1.5	Inherited cancer risk	13
1.6	Cancer predisposition syndromes as rare diseases	17
1.6.2	Genetic diagnosis of cancer predisposition	20
1.7	Towards novel findings in pediatric cancer predisposition	21
1.7.1	Multi-omics approaches	22
1.7.2	CPS disease modelling	22
1.7.3	Epidemiological studies of childhood CPS	23
1.8	Precision medicine	24
1.8.1	Patient tailored cancer diagnosis and treatment	25
1.8.2	Cancer surveillance in patients with pediatric CPS	26
2	Research Aims	27
3	Materials and Methods	29
3.1	Study participants	29
3.1.1	ChiCaP Cohort	29
3.1.2	Population-based studies	29
3.1.3	Ethical approval	30
3.2	Compilation of a pediatric CPS gene panel	31
3.3	Histological assessment of the tumors	31
3.4	Culture of human dermal fibroblasts	31
3.5	Genome sequencing and molecular methods	31
3.5.1	DNA sequencing	31
3.5.2	Gene expression assays	33
3.5.3	Methylation analyses	34
3.5.4	Protein detection	35
4	Results and Discussion	37
4.1	Compilation of a broad childhood CPS gene panel	39
4.2	The ChiCaP project	39

4.3	Cancer risk in congenital syndromes studied in this thesis	40
4.3.1	Prader-Willi Syndrome	40
4.3.2	Marfan Syndrome.....	41
4.3.3	Limb-Girdle Muscular Dystrophy Recessive 1 (LGMDR1).....	43
4.3.4	Muscular dystrophy and myotonic disorders.....	44
4.4	Discovery and characterization of a novel syndrome	45
4.4.1	Recessive <i>FLCN</i> -related disorder	45
5	Ethical considerations.....	49
6	Conclusions	51
7	Points of perspective	53
8	Acknowledgements.....	55
9	References	59

List of abbreviations

BHDS	Birt-Hogg-Dubé Syndrome
BMD	Becker Muscular Dystrophy
CI	Confidence Interval
CNS	Central Nervous System
CPS	Cancer Predisposition Syndrome
ChiCaP	(Swedish) Childhood Cancer Predisposition Project
ddPCR	Droplet Digital Polymerase Chain Reaction
DMD	Duchenne Muscular Dystrophy
DNA	Deoxyribonucleic acid
HR	Hazard Ratio
InDels	Insertions and Deletions
LOEUF	Loss-of-function observed/expected upper bound fraction
LoF	Loss-of-function (variant)
LGMD	Limb-Girdle Muscular Dystrophy
LOI	Loss-of-imprinting marks
MFS	Marfan Syndrome
MLPA	Multiplex ligation-dependent probe amplification
MS-MLPA	Methylation-specific MLPA
OG	Oncogene
PCR	Polymerase Chain Reaction
PWS	Prader-Willi Syndrome
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
SNV	Single Nucleotide Variant
SV	Structural Variants
TGRS	Total Gene Rank Score
TSG	Tumor Suppressor Gene

WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

1 Background

A long road has been transited for us to be able to understand genetic diseases to the depth we do today. In the first paragraphs of my thesis, I would like to take you with me on a trip across centuries, from peas-counting monks to synthetic biology. But let's not get ahead of ourselves, we start from the beginning.

1.1 The history of human genetics

Although the idea of “hereditary traits” existed much before [1], it was only in 1859 that the theory of evolution was proposed by Charles Darwin [2], after a five-year voyage around the world. Loosely, the theory formulates that all living organisms arose from a common ancestor, which evolved into different species by natural selection; according to which, only the most suitable characteristics for an environment are passed on to the offspring. Not long after, a monk called Gregor Mendel, realized during his experiments with peas, that certain crosses always resulted in the same phenotypic outcome. Meticulously, he designed experiments to understand the nature of this phenomenon, resulting in the formulation of the principles of inheritance [3].

Multiple researchers paved the way of genomics during the decades that followed, including the selected ground-breaking discoveries presented below. First, Friedrich Miescher isolated an acidic substance from white blood cells and called it “nuclein” [4], today known as DNA. Then, while studying inborn errors of metabolism, Archibald Garrod associated the principles of inheritance with disease, publishing the first known monogenic disorder: Alkaptonuria [5]. Shortly after, Thomas Hunt Morgan and his student, Alfred H. Sturtevant, through experiments on fruit flies, discovered that DNA is stored in units called chromosomes [6, 7] and drew the first chromosome map [8].

Further, Oswald Avery, Colin MacLeod and Maclyn McCarty identified that DNA causes bacterial transformation, suggesting it as the carrier of genetic information [9]. Next, the chemical structure of the DNA building blocks – deoxynucleotides – was resolved, and Erwin Chargaff proposed that their amount was not random. DNA contained the same number of purines and pyrimidines [10, 11].

The next step in this historical trip takes us to Cambridge University in the 1950s, where Rosalind Franklin generated an iconic DNA X-ray diffraction photograph (Photo 51 [12]). Meanwhile, James Watson and Francis Crick worked on deciphering the molecular model of DNA. After many tries, using crystallography work from

Rosalind Franklin and Maurice Wilkins, they finally cracked the structure of DNA [13]. The next game changer in the history of genomics took place when Frederick Sanger figured out the order of amino acids in insulin [14, 15]. He concluded then that the building blocks of proteins had a specific order, and therefore, DNA should too. In 1977, following these studies, Sanger published the first method for DNA sequencing [16], the basis of which is still used today.

Multiple important techniques and resources for the study of the human genome were developed in the following years, including Fluorescence *in-situ* hybridization [17, 18], the Polymerase chain reaction (PCR) [19], automated sequencing instruments [20], human genetic maps based on DNA markers [21], and positional cloning of disease-causing genes [22]. Meanwhile, international consortia to study human genetics were created. The biggest research program was the Human Genome Project, launched in 1990 with the aim to sequence the entire human genome –That is, roughly 3 billion DNA letters– in 15 years. This project became a race when a scientist named Craig Venter founded a private company aiming to sequence the genome faster, with his own scientific method. The technique was based on shedding the entire human genome into small fragments, sequencing them in parallel, and using computational methods to piece them together. Thanks to this race, the first draft of the human genome was astonishingly completed by both parties in only 10 years [23, 24].

A myriad of technological advances in the field of genetics have taken place after the launching of the genome project [25]. Some examples are the cloning of Dolly the sheep [26, 27], the sequencing of the genome of various eukaryotic organisms [28], the development of novel sequencing approaches [28], the introduction of the “omics sciences”: genomics, transcriptomics, proteomics, metabolomics, and epigenomics [29], and the generation of tools for targeted genome editing [30, 31]. Figure 1 presents a graphical summary of milestones in human genetics.

This trip of scientific progress places us in a very exciting time to work in clinical genetics. The use of massive parallel sequencing for diagnosis has become part of standard clinical care in many countries, and medicine is moving towards a personalized approach according to an individual’s genetic profile. Further, advanced therapies based on modified cells, nucleic acids or proteins are becoming available. The main caveat to the past statements being the inequality of the clinical applicability of genomic advances across the globe [32]; a challenge which must be addressed in the years to come.

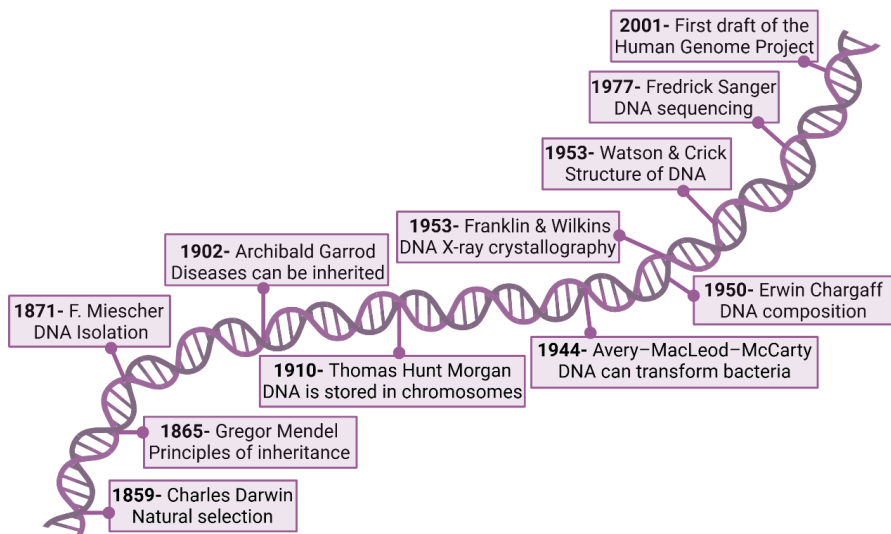


Figure 1. Timeline of landmarks in the history of human genetics. Most dates correspond to the year of publication of the respective landmark.

1.2 The human genome

The genome of an organism consists of the entire set of DNA instructions contained in the cell. Although mitochondria –an organelle which generates energy for the cell– has its own set of DNA instructions, the main manual on how to create a human is preserved in the nucleus. These instructions comprise about 3.2 billion pairs of building blocks known as deoxynucleotides, stored in 23 chromosome pairs; a set inherited from each parent.

The architecture of the genome is complex, some parts arranged in ways that still escape our understanding. This organization is very important, as it allows the orchestrated generation of thousands of proteins, which carry out the vast majority of functions in living organisms. Incredibly, only about 2% of the total genome encodes proteins. In recent years, we have finally started to understand the fascinating functions of the remaining 98% [33, 34].

1.3 Human genetic variation and disease

Although most of the DNA sequence is shared between individuals, about 0.1% of the genome differs from person to person. These sites are known as *genetic variants* and arise by spontaneous changes in unborn babies, but stay within the genetic pool, as they are passed on to new generations. Genetic variation is important, as it creates human genetic diversity, but sometimes it can also result in disease. Hereditary disorders occur when inherited DNA changes are harmful.

Genetic variants can happen at different scales. The smallest changes in DNA are Single Nucleotide Variants (SNV), in which only one base pair differs between individuals. Going up, Insertions and Deletions (InDels) are differences no longer than 50 nucleotides in size. Larger changes are called Structural Variants (SV), which are balanced when there is no loss or gain of DNA sets or unbalanced when genetic material is lost or added.

Unbalanced variation includes copy number variants –insertions and deletions–, while balanced changes comprise translocations and inversions. As the name implies, insertions occur when pieces of DNA are added to the genome, while deletions happen when DNA sequences are removed. On the other hand, translocations refer to exchanges of genomic material between chromosomes, while inversions encompass changes where DNA pieces within the same chromosome are flipped.

As mentioned in the historical background; to understand genomic variation, we need to decode the order of nucleotides in the genome in a process known as DNA sequencing. There are multiple sequencing methods used nowadays [35]. They can be focused on specific pieces of DNA or encompass the entirety of the genome. They can also generate small fragments of information –of about 150 nucleotides in size– or read much longer DNA molecules –currently around 100.000 nucleotides –[36, 37]. The selected approach usually depends both on the objective of the study and the budget.

1.3.1 Effects of DNA changes in proteins

Except for rare cases, the information contained in the coding region of DNA is converted into RNA in the nucleus, in a process called transcription. RNA molecules are then transported to the cell cytoplasm, while they are tweaked and cut in a process known as RNA maturation. Mature RNAs serve as messengers for the creation of proteins. This step, known as translation, involves the puzzling together of specific amino acids to form functional polypeptides, enabling the generation of proteins with thousands of functions.

As mentioned before, about 0.1% of the genome differs between individuals. This is possible because most changes at DNA level do not affect RNA nor proteins. A few variants do, however, modify proteins. Some changes impact the amount of protein produced, while others affect the function of the molecule. Amongst the last group, loss-of-function (LoF) variants lead to the disruption of the protein's

role, while gain-of-function mutations can increase its baseline activity (hypermorphic) or generate a completely new function (neomorphic).

Additional concepts commonly used to refer to the effect of a variant at protein level are dominant-negative, haploinsufficiency and hypomorphic variants. Dominant-negative mutations lead to a protein function which interferes with the correct activity of the typical (wild-type) protein. On the other hand, haploinsufficiency refers to changes in which one copy of the protein is altered, while the remaining copy is not enough to maintain normal functioning. Finally, variants leading to partial loss of protein function are known as hypomorphic. The impact of variants on the phenotype is a complex topic, reviewed in [38].

1.3.2 Rare diseases

"When you hear hoofbeats, don't expect to see a Zebra" – Theodore Woodward¹

According to the Operational Description of Rare diseases and the European Union, rare diseases are conditions with an incidence of less than 1:2000 individuals [39, 40]. Although individually rare (or even very rare), collectively these diseases are quite common, affecting about 3–8% of the world's population [40–42]. Rare diseases are very diverse, with over 7000 types described [43, 44]. For instance, all forms of pediatric cancer are included in this category [45]. About 70% of all affected individuals are children; approximately 1/3 dying before the age of five [46, 47]. In general, patients with rare diseases go through a "diagnosis odyssey", waiting on average 4–5 years before understanding the cause for their disease [47]. Moreover, even after extensive genetic investigations, including whole genome sequencing (WGS), the majority remain undiagnosed.

Of all rare diseases, it is estimated that 72% are of genetic origin [48], usually caused by harmful genomic mutations. If the variants causing disease are located in a single gene, the rare disorder is known as monogenic. On the other hand, if more locations are involved, it is polygenic. Further, rare diseases can be inherited –if the mutation is found in parents–, or *de novo* –if the variant is new to the patient.

¹ This citation is a metaphor used to teach medical students. When examining patients, they should think about the most common cause of their symptoms first –in nature, that would be thinking of horses when hearing hoofbeats–. At the rare diseases group, we specialize in discovering zebras when doctors find no horses.

In this thesis, we studied rare genetic variation leading to disease. Specifically, those diseases which result in an increased risk of cancer in children. Therefore, the next parts of this background section will focus on describing this complex disease group, from the tumor point of view –on somatic cells–, and the inherited perspective –germline variants–.

1.4 Cancer – a genetic disease of somatic cells

“Cancer, perhaps, is an ultimate perversion of genetics – a genome that becomes pathologically obsessed with replicating itself” – Siddhartha Mukherjee.

Cancer is a disease characterized by uncontrollable cell division, making it hard to contain. Usually, single mutations do not transform a cell into its malignant counterpart. Instead, cancer develops when a series of key genes is affected sequentially, resulting in rapidly dividing cells, which colonize their niche and expand to new tissues i.e., metastasize [49] (Figure 2). In most cases, this process takes time. Therefore, age is one of the major risk factors for adult cancer.

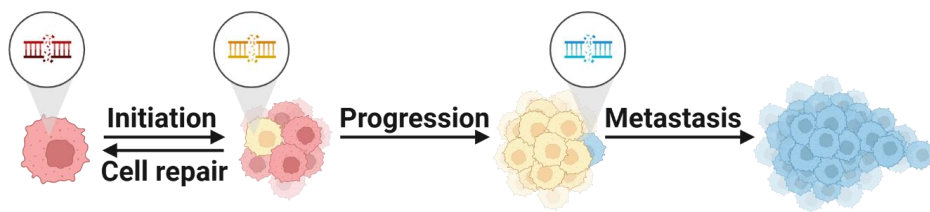


Figure 2. Carcinogenesis, simplified scheme. Benign cells acquire somatic variants leading to the initiation of carcinogenesis. This process can be reversed by cell repair, while additional mutations result in cancer progression and ultimately metastasis.

In contrast to inherited disorders, cancer development is mostly a result of genetic alterations acquired during our lifetime i.e., somatically [50]. Lifestyle-related factors increasing cancer risk include, for instance, smoking, drinking, obesity, and physical inactivity [51]. Hence, cancer can be seen as a genetic disease of somatic cells; and therefore, the most common genetic disorder, affecting more than 19 million people per year, and resulting in close to 10 million deaths worldwide [52].

Although the genome of a cancer cell usually carries a high number of genetic changes (sometimes thousands of new mutations!), the number of variants needed to turn a normal cell into its malignant counterpart is estimated at less than a dozen [53]. The set of genes necessary for this transformation is known as cancer drivers. Depending on their effects on cells, drivers are divided into Tumor

Suppressor Genes (TSG) and oncogenes (OG). TSG work as caretakers of the genome, ensuring its stability; or gatekeepers, controlling cell growth and division. On the other hand, OG operate as positive regulators of cell division or growth (Figure 3). Cancer cells combine the inactivation of TSG and the chronic activation of OG in a synergic process to ultimately achieve their malignant state [50].



Figure 3. Genetic changes in cancer evolution. The set of genetic changes enabling carcinogenesis is known as cancer drivers, classified as oncogenes and tumor suppressor genes. All other changes in cancer cells are known as passenger variants.

Driver genes allow malignant cells to sustain the physiological changes necessary for the development of cancer. Initially, six hallmarks of tumors were described: (1) cell growth independently of external signals, (2) ability to grow despite antiproliferative signals, (3) bypassing programmed cell death, (4) limitless ability to divide, (5) continuous formation of new vessels, and (6) the ability to metastasize [49]. Two additional hallmarks were added a decade later, namely (7) modification of metabolism to support limitless cell growth, and (8) evading destruction by the immune system [54]. These hallmarks are enabled by the presence of two important characteristics of malignant cells: genomic instability, which results in rapidly evolving tumor genetics; and sustained inflammatory responses [54].

As more knowledge is gathered about the mechanisms leading to cancer, new traits intrinsic to carcinogenesis are described. For instance, phenotypic plasticity, disrupted cell differentiation [55] and aberrant alternative splicing [56]. Moreover, these hallmarks are a simplified way to understand a very complex disease. A complete view of cancer would need to be considered systemically, with additional dimensions than the molecular biology of the malignant cell [57]. For this thesis however, we will adhere to the concept of cancer as explained by its ground hallmarks.

1.4.1 Childhood cancer

As mentioned before, one of the main risk factors for the development of cancer is aging. It is not surprising thereof that childhood cancer is rare, with about 280.000 new cases per year worldwide [52]. In Sweden, 300–350 children are diagnosed with cancer every year, corresponding to an incidence of about 17 in 100.000 individuals. Data from the Swedish childhood registry shows a similar proportion of leukemias (30%), brain tumors (27%) and other solid malignancies (42,5%) in children with cancer [58], and a male-to-female ratio of 1.15 (years 2000–2020) [59]. However, the age at onset varies between diagnostic groups (Figure 4). While solid tumors have a constant frequency across ages, leukemias have a peak incidence between 1–3 years of age, and central nervous system (CNS) tumors at the age of 1–4 years [58]. Thus, it is more common for children under the age of four to develop cancer.

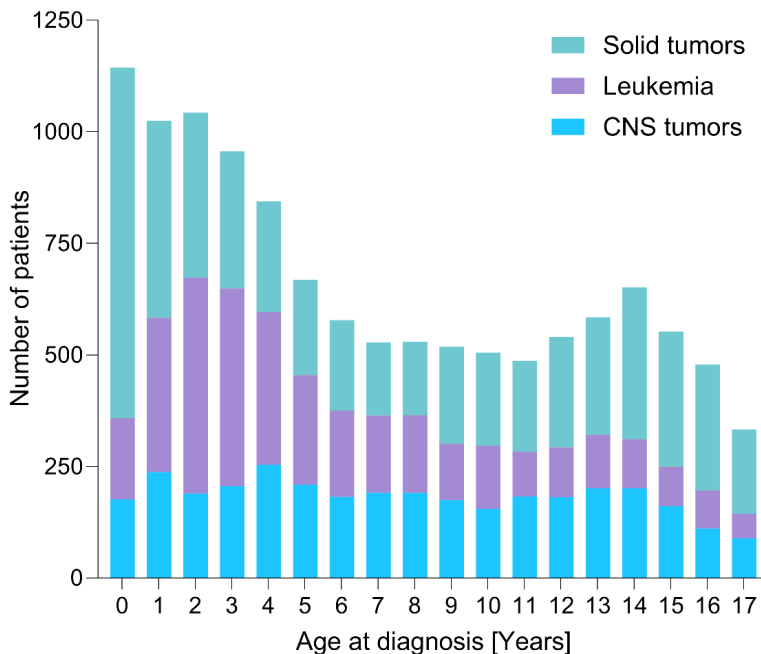


Figure 4. Childhood cancer diagnoses in Sweden. Total number of children diagnosed with central nervous system (CNS) tumors, leukemia, and solid tumors in Sweden, 1982–2020, according to the 2022 yearly report of the Swedish childhood registry [58] by age at cancer diagnosis.

Although their hallmarks remain unchanged [49, 54], the biology and genetics of childhood and adult cancer differ markedly. In contrast to their adult counterparts, childhood tumors develop over a short period of time, and harbor on average 14–

times fewer variants. Further, most driver mutations in pediatric tumors are exclusive to a specific childhood cancer type, as compared to frequent co-mutations driving adult cancer. Finally, driver genes in childhood and adult cancers do not completely overlap [60]. Therefore, not all the knowledge collected in adult cancer research translates to children, creating the need to study pediatric cancer independently.

1.5 Inherited cancer risk

"If you have lung cancer, the most important thing you can know is your genetic code" – Craig Venter.

As previously mentioned, cancer formation is a complex process that takes time. Why do children and young adults develop cancer then? It is currently thought that pediatric cancer is caused in part by exposure to environmental mutagens leading to DNA damage, viral infections promoting the development of certain malignancies, somatic mutations on progenitor cells, and hereditary predisposition [61].

The last mentioned refers to cancer predisposition syndromes (CPS). These are caused by genetic variants leading to an increased risk of developing malignancies. Most CPS are inherited in an autosomal-dominant manner [62] –that is, genetic changes in one copy of the gene are enough to increase the risk of cancer. Although CPS can be a risk factor for cancer across all ages [63, 64], the following paragraphs will center on childhood CPS.

The first observations suggesting that cancer could be inherited were made in children with retinoblastoma, a type of cancer appearing in the retina. Alfred G. Knudson statistically analyzed the incidence of retinoblastoma [65], finding two forms. An inherited class, where one mutation occurs in the germ cells and a second somatically; and a sporadic form, in which both mutations must occur somatically, in the same cell, for the disease to develop (Figure 5).

As in retinoblastoma, two "hits" to the DNA are necessary for carcinogenesis involving many TSG. In sporadic forms, separate mutations to both alleles are required to inactivate the TSG in a single cell, initiating tumor formation. As the chance that two mutations occur in an individual cell is relatively small, sporadic cancer involving TSG inactivation is rare. In hereditary forms, however, a first mutation is present in a germ cell, while the "second hit" is somatically acquired

(Figure 5). As somatic changes are rather frequent, individuals with germline mutations on TSG often develop multiple tumors.

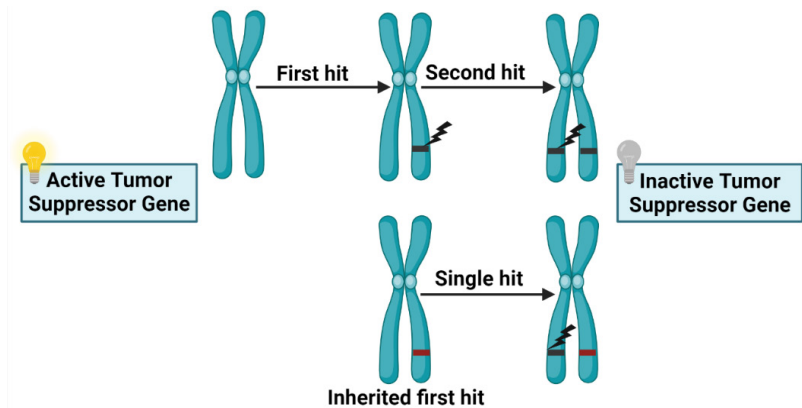


Figure 5. Scheme of Knudson's two-hit hypothesis. Inactivation of tumor suppressor genes can lead to cancer. Usually, two pathogenic inactivating changes are required. However, if an individual inherits the first variant, only an extra hit is needed.

After Knudson studied retinoblastoma, many CPS have been described. Today, there are more than 150 known pediatric CPS, with some of the most common described in Table 1 (See also [66] for a review on childhood CPS). Importantly, each CPS has a unique cancer risk fold and spectrum. This variability can hinder their clinical diagnosis at the pediatric oncology unit.

In general, CPS are more common in children with multiple malignancies, family history of cancer, high treatment-related toxicity, specific tumor types, and concomitant genetic syndromes [67]. Based on these risk factors, selection tools have been developed, such as The Childhood Cancer Screening checklist [68], the Jongman's criteria [67, 69], the McGill Interactive Pediatric OncoGenetic Guidelines [70] and the Childhood Cancer Predisposition (ChiCaP) criteria (**Study VII**). Although these tools have improved pediatric CPS diagnosis [71], cancer predisposition can still be missed for instance due to lack of family history of cancer, absence of specific symptoms, and second malignancies mistaken as relapses from the initial tumor.

Table 1. Common inherited childhood cancer predisposition syndromes

Syndrome	Genetic cause	Inheritance	Lifetime cancer risk	Cancer spectrum	Other clinical features	References
Li-Fraumeni Syndrome	<i>TP53</i>	AD	~70% (Men) ~90% (Women)	Broad spectrum. Especially ACC, BC, CNS, OS, and STS.	None.	[72]
Constitutional Mismatch Repair Deficiency	Mismatch repair genes (<i>MLH1</i> , <i>MSH2</i> , <i>MSH6</i> , <i>PMS1</i> , <i>EPCAM</i> , <i>PMS2</i>)	AR	Varies depending on the genetic cause	Lymphoma, ALL, GB, CRC, GIT, among others.	Hypopigmentation or café-au-lait spots.	[73]
Fanconi Anemia	DNA crosslink repair genes (21 genes known. Most common, <i>FANCA</i>)	Mostly AR	~13% AML by age 50	Mainly AML, but also HNT, SC, and GUC.	Congenital malformations, such as cardiac defects, limb and skeletal malformations, ophthalmic and urinary tract abnormalities, short stature, and bone marrow failure.	[74]
<i>DICER1</i> tumor predisposition	<i>DICER1</i>	AD	~26%/10% by age 50 (Women/men)	PPB, TC, SCST, CBME, and other malignancies.	Macrocephaly, cystic nephroma, multinodular goiter, retinal and kidney abnormalities.	[75]
Heritable retinoblastoma	<i>RB1</i>	AD	Almost complete penetrance for Rb	Mainly Rb, but also PB, OS, STS, and SC.	None.	[76]
Xeroderma pigmentosum	DNA excision repair genes (<i>DDB2</i> , <i>ERCC1-5</i> , <i>POLH</i> , <i>XPA</i> , and <i>XPC</i>)	AR	~65% for SKC	Mainly SKC, SKCM, OC, and OSCC. Other malignancies reported.	Sun hypersensitivity, ocular abnormalities, and possible neurological abnormalities.	[77]
Neurofibromatosis type 1	<i>NFI</i>	AD	~38.8% by age 50	Mainly CNS and PNS/ Other malignancies reported.	Neurodevelopmental disorders, café-au-lait spots, and cutaneous neurofibromas.	[78, 79]
Noonan syndrome	Genes in the RAS/MAPK pathway	AD	~4% by age 20	Mainly JMML, AML, ALL, CNS, and Nb. Other malignancies reported.	Eye abnormalities, hypotonia, hyperflexible joints, pectoral deformity, cryptorchidism, congenital heart disease, and neurodevelopmental disorders.	[80–82]
WT1 disorder	<i>WT1</i>	AD	Varies depending on the genetic variant	Mainly Wilms tumor and GoB.	Steroid-resistant nephrotic syndrome, proteinuria, abnormal testicular development, and kidney and urinary tract abnormalities.	[83]

Abbreviations: ACC, Adrenocortical carcinoma; AD, Autosomal Dominant; ALL, Acute Lymphoblastic Leukemia; AML, Acute myeloid leukemia; AR, Autosomal recessive; BC, Breast cancer; CBME, Ciliary body medulloepithelioma; CNS, central nervous system; CRC, Colorectal carcinoma; EC, Endometrial cancer; GB, Glioblastoma; GoB, Gonadoblastoma; GIT, Gastrointestinal tumors; GUC, Genito-urinary tract cancer; HNT, Head and neck tumors; JMML, Juvenile Myelomonocytic Leukemia; Nb, Neuroblastoma; OC, Ocular cancer; OS, Osteosarcoma; OSCC, Oral squamous cell carcinoma; PB, Pineoblastoma; PC, Pancreatic cancer; PNS/1, peripheral nerve sheath tumor; PPB, Pleuropulmonary blastoma; PTC, Prostate cancer; Rb, Retinoblastoma; SC, Skin cancer; SCST, Sex cord-stromal tumors; SKC, Skin cancer; SKCM, Skin cutaneous melanoma; STS, Soft-tissue sarcoma; TC, Thyroid Cancer; TML, Transient myeloproliferative disorder.

Table 2. Selected studies of germline genetic diagnosis of childhood cancer predisposition syndromes

Study	Cohort	Diagnostic method	Percentage CPS variants	Most common germline findings	Reference
St. Jude–Washington University Pediatric Cancer Genome Project	1120 unselected pediatric cancer patients (<20 years).	WGS/WES/Both. Panel of 60 genes.	8.5%	TP53 (1/2 of all positive diagnoses), APC, BRCA2, NF1, PMS2, RBI, and RUNX1.	[84]
Baylor College of Medicine Advancing Sequencing in Childhood Cancer Care (BASiC3)	150 patients with newly diagnosed solid tumors (<18 years).	WES. Panel of 112 genes.	10%	VHL, TP53, DICER1, MSH2, WT1, KRAS, BRCA1, BRCA2, SMARCA4, and CHEK2	[85]
Laboratory of Personalized Genomic Medicine at Columbia University Medical Center (CUMC)	90 high-risk patients: prognosis <50% 5-year survival, outlier phenotype, rare cancer, possible CPS, relapse (<18 years).	WES and RNA-seq. CPS filtering panel not specified.	14%	APC, ATM, CIQA, PMS2, XIAP, RUNX1, MLL2.	[86]
Pediatric Pan Cancer (pedPanCan)	914 pediatric cancer patients (<25 years). Multiple cohorts.	WGS or WES. Panel of 162 genes.	7.6%	TP53, NF1, BRCA2, RBI, LZTR1, MSH6 and SMARCB1.	[60]
Zero Childhood Cancer	247 pediatric patients with high-risk or rare cancer types (<22 years at diagnosis).	WGS and RNA-seq. Panel of 161 genes.	16.2%	CHEK2, SMARCB1, BRCA2, PMS2, NF1, ATM, FANCA, MSH2, MSH6, PALB2, SDHB, SMARCA4, HAVCR2.	[87]
Sequencing Tumor and Germline DNA—Implications and National Guidelines (STAGING)	198 unselected pediatric cancer patients (<18 years).	WGS. Panel of 314 genes.	14.6% (10.6% pediatric and 4.5% adult CPS).	CDC73, DDX41, LZTR1, NF1, RBI, SDHC, SMARCA4, TP53, TSC2, UPD 11p, trisomy 21.	[88]
Germline mutations in children with cancer	160 unselected pediatric cancer patients (<19 years).	Trio WES. Panel of 295 genes.	13.8% (include 6.9% variants of unknown significance).	PTPN71, MSH6, TP53, NF1, DICER1 and, ATM.	[89]
Memorial Sloan Kettering–integrated mutation profiling of actionable cancer targets (MSK-IMPACT)	751 patients with solid tumors (<19 years).	MSK-IMPACT capture-based assay. Panel of 88 genes.	18% (includes low-penetrance and heterozygous variants in recessive genes)	RBI, NF1, TP53, CHEK2, PMS2, SDHA, PHOX2B, MUTHY, RECQL4, MTF, APC, 1307K, and FANCA.	[90]
Childhood Cancer Predisposition Project (ChiCaP)	309 children with solid tumors (<18 years).	WGS. Panel of 189 genes.	11%	NF1, RBI, WT1, DICER1, MSH2, PMS2, TP53	[91]
Genomes for Kids (G4K)	300 unselected pediatric cancer patients (<22 years, except one patient).	WGS and WES. Panel of 156 genes.	18% (10% with confirmed relevance for tumor formation).	RBI, MUTHY, NF1, CHEK2, PMS2, RECQL4, APC, ATM, ERCC2, FANCM	[92]
Abbreviations: CPS, Cancer Predisposition Syndromes; RNA-seq, RNA sequencing; WES, Whole Exome Sequencing; WGS, Whole Genome Sequencing.					

1.6 Cancer predisposition syndromes as rare diseases

Pediatric CPS are rare diseases. What is more, certain rare congenital syndromes increase cancer risk and can thus be considered CPS. For instance, patients with Down syndrome present a 10–400-fold increased risk for leukemia [93], while 2–10% of Wilms tumors are associated with an inherited disorder [94, 95]. With the improved life expectancy in children with congenital abnormalities, novel associations with cancer predisposition arise. The following subsections give an overview of the congenital syndromes which were the focus of my research.

1.6.1.1 Imprinting disorders: Prader Willi Syndrome

Study II focused on the association between cancer and Prader–Willi Syndrome (PWS), a disorder affecting multiple body systems. Neonatally, patients present with poor sucking and hypotonia –that is, poor muscle tone–. In childhood, they develop hyperphagia –insatiable hunger–, short stature, hypogonadism –low hormonal production in testicles and ovaries–, characteristic facial traits and intellectual disability [96] (Figure 6). Although PWS is not considered a CPS, epidemiological studies have found an increased incidence of leukemia [97] and other malignancies [98, 99] in this patient group.

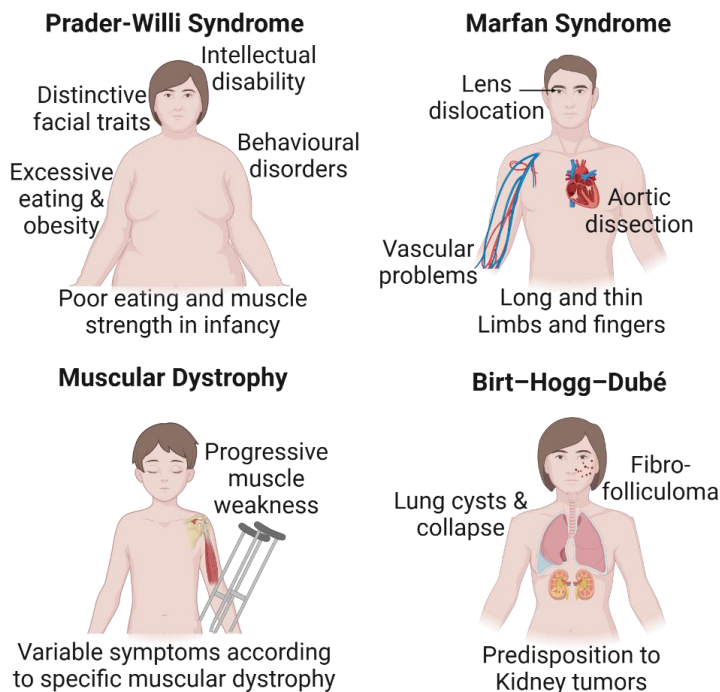


Figure 6. Characteristic symptoms of the congenital disorders studied in this thesis.

PWS is an imprinting disorder. To understand what this implies, we need to go through some definitions. First, epigenetics, which is the study of characteristics or cell changes, independently of DNA variation. Genomic imprinting is an epigenetic mechanism leading to the expression of certain chromosomal regions depending on the sex of the parent from which they are inherited [100] (Figure 7). Imprinting is established during the formation of the gametes, and it is maintained during our lifetime in almost all tissues. Genetic changes affecting these patterns of gene expression can result in human disease, including for instance Beckwith-Wiedemann syndrome (when the imprinted region at 11p15.5 in our genome is affected); and Angelman syndrome/PWS (when the imprinted region at 15q11-q13 is affected on the maternal and paternal allele, respectively) [101].

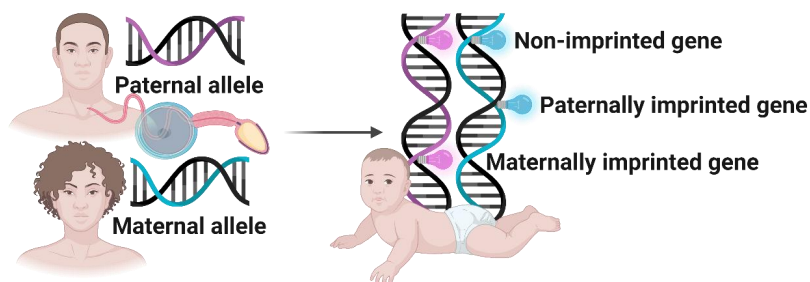


Figure 7. Schematic illustration of genetic imprinting. In the figure, gene expression is depicted by a blue (maternal) or pink (paternal) light bulb. Both alleles are expressed ("turned-on") in non-imprinted genes. Moreover, in maternally imprinted genes, only the paternal allele is expressed, and vice versa.

Changes in imprinting patterns have also been reported in cancer. In the described cases, loss-of-imprinting marks (LOI) lead to the re-expression of normally silent alleles. If the re-expression is advantageous to the cell, LOI can work as a cancer driver [102]. A well-studied example is Wilms tumor, in which LOI at the 11p15.5 imprinted region results in reduced expression of the TSG H19, conferring growing advantages to the cell [103, 104]. Recently, the first case of LOI in a testicular tumor of an individual with PWS was reported [105]. The implications of LOI at this site in cancer development in patients with (or without) PWS remain to be studied.

1.6.1.2 Cancer predisposition in Marfan syndrome

In **Study III**, we investigated the cancer risk in individuals with Marfan syndrome (MFS), a rare disease with a prevalence of 10–20 in 100.000 individuals in Sweden [106]. MFS is an autosomal dominant disorder affecting the connective tissue, with a broad disease spectrum, from mild symptoms to a fatal presentation –known as

neonatal MFS-. The main systems affected are the eyes, skeleton, and cardiovascular system. Clinical diagnostic criteria include the presence of aortic aneurysm and dissection -bulge or breakage of the aortic artery, respectively-, overgrowth, and ectopia lentis -dislocation of the crystalline lens of the eye- [107, 108] (Figure 6).

FBNI, the disease-causing gene for MFS, encodes a protein called fibrillin, which makes up cell microfibrils, which in turn form the extracellular matrix [109]. Microfibrils are the fabric of tissues, providing elasticity, stability, and anchoring different important proteins. Dysregulation of *FBNI* results, for instance, in changes in the activation of TGF- β , a known cancer driver [110]. However, only one epidemiological study on the cancer risk in individuals with MFS has been performed, reporting an overrepresentation of head and neck, and urinary track malignancies in patients with MFS [111]. Additional information is therefore required to clarify the link between MFS and cancer development.

1.6.1.3 *Cancer risk in muscular dystrophies and myotonic disorders*

In **Studies IV** and **V**, we investigated the development of cancer in individuals with muscular dystrophies and myotonic disorders. Muscular dystrophies are genetic diseases characterized by progressive muscular weakness, due to the degeneration of muscle cells (Figure 6). There are many types of muscular dystrophies, with variable age at onset, symptoms and disease severity. These include for instance, Duchenne Muscular Dystrophy (DMD), Becker Muscular Dystrophy (BMD), and Limb-Girdle Muscular Dystrophy (LGMD). Myotonic dystrophy is also clinically classified as a muscular dystrophy; moreover, these terms refer to different types of muscle disorders. Specifically, myotonic dystrophy is characterized by myotonia -inability to relax muscles after contraction. From these disorders, only myotonic dystrophy is a known CPS, increasing the risk for multiple cancer types in adults [112–117].

Interestingly, mouse models of multiple muscular dystrophies develop soft tissue sarcomas, rare cancers affecting connective and supportive tissue as muscles, fat, blood vessels, nerves, tendons and joints. Initially, an increased risk for mixed sarcomas was observed in the muscles of aged DMD mice [118]. Follow up studies on different mouse strains confirmed these results, not only for DMD, but also for two types of LGMD, and a rare severe type of muscular dystrophy [119–121]. The sarcoma incidence varied between diagnoses and increased in mice lacking two

of the studied genes [119–121]. This predisposition to develop sarcomas has not been reported in patients.

1.6.1.4 Birt-Hogg-Dubé syndrome and the *FLCN* gene

In **Study VI**, we characterized a novel rare disease, caused by compound variants in *FLCN*, a known disease-causing gene. Single pathogenic variants in *FLCN* lead to Birt-Hogg-Dubé syndrome (BHDS), an autosomal dominant disease associated with kidney cancer predisposition, with a lifetime risk of 19 to 34% [122, 123]. *FLCN* acts as a TSG and most kidney tumors present somatic second hits [124, 125]. Additional symptoms of BHDS include fibrofolliculomas, lung cysts and spontaneous pneumothorax –lung collapse– [126] (Figure 6). For a review on cancer predisposition in patients with BHDS see [127].

Folliculin, the protein encoded by the *FLCN* gene, is a master regulator of cell metabolism. The functions of the protein are complex, so I will just take you through the main regulated pathways (See [124] for an in-depth gene review). Under nutrient rich conditions, folliculin activates the mTORC1 pathway via Rag GTPases, resulting in cell growth and proliferation [128]. On the other hand, during starvation, folliculin forms a complex with FNIP1/FNIP2, leading to AMPK mediated catabolism, production of new mitochondria, and oxidative phosphorylation through PGC1 α / β –metabolic pathways leading to energy production in the cell–. Additionally, under nutrient deficient conditions, the *FLCN* complex does not hinder the translocation of TFEB and TFE3 from the cytoplasm to the nucleus, leading to the transcription of genes mediating, for instance, lysosomal biogenesis [128]. In **Study VI**, we explore the function of these pathways in patient’s cells.

1.6.2 Genetic diagnosis of cancer predisposition

In Sweden, germline genetic testing of all pediatric cancer patients became clinical routine after the completion of **Study VII**. Upon cancer diagnosis, germline WGS is offered to patients. If a CPS is confirmed, genetic testing is also offered to other family members at risk.

As the genome has about 3.2 billion letters, strategies for genetic testing vary in the percentage covered. When a genetic disease affecting a specific gene or small group of genes is suspected, targeted panels covering those regions exclusively can be used. These are usually smaller, manageable, and can be cost-effective (For an example in CPS, see [129]).

More general approaches include the use of Whole Exome Sequencing (WES) or WGS, both allowing an unbiased analysis of DNA variants. In WES, the genomic region encoding proteins –known as exome– is sequenced. This piece constitutes less than 2% of the entire genome, and it is responsible for most known pathogenic genetic changes [130]. Moreover, WGS is a comprehensive method, where the entire genome is sequenced. Therefore, it enables the identification of most variant types, including copy number changes, structural rearrangements, and variants in the non-coding genome.

A convenient strategy for genetic diagnosis is the use of WES/WGS followed by a filtering step with *in-silico* gene panels. In this case, sequencing results are masked for variants in genes associated with the disease suspected, e.g., CPS. As only relevant regions are studied, the likelihood of incidental findings –that is, detection of variants that were not the aim of the test–, and the number of variants to analyze decreases. Importantly, it is also possible to re-analyze the results as new gene–disease associations arise. Clinical and research efforts to assemble gene panels for the study of CPS have been published recently [66, 89, 91, 92, 131, 132]. Selected research studies using different forms of this methodology for childhood CPS diagnosis are presented in Table 2.

In children with cancer, sequencing methods have applications in two fronts: diagnosis of cancer predisposition and tumor characterization. As previously described, germline genetic testing can help us understand the inherited cancer risk. Moreover, tumor genomic sequencing gives oncologists a comprehensive account of the tumor of a specific individual, further guiding treatment and surveillance [133]. This interplay of tumor and germline genetic diagnosis in cancer patients is central for the implementation of personalized medicine. However, the application of paired germline–tumor genetic testing is challenging. Limitations in terms of centers' capacity, cost of testing and, when applicable, reimbursement, still reduces the availability of genetic diagnosis to a few countries.

1.7 Towards novel findings in pediatric cancer predisposition

The diagnostic yield of pediatric CPS has increased in the last decades. This has been facilitated by the high throughput and price-effectiveness of novel sequencing methods, as well as the increased knowledge about the genetics of childhood CPS. However, challenges remain, such as the lack of knowledge on the prevalence of CPS in the general population, the clinical impact of surveillance programs, and the cancer risk amongst children with certain CPS.

These unsolved questions may result in childhood CPS underdiagnosis, which negatively impacts patient's clinical care. Specifically, it hinders personalized cancer treatment when available, access to surveillance and genetic counselling. Therefore, additional research on novel CPS and molecular mechanisms leading to childhood cancer is needed. For instance, germline genetic studies of large cohorts of children with cancer and functional studies on candidate genes. The following sections summarize selected methods for the discovery of novel CPS (See also Figure 8).

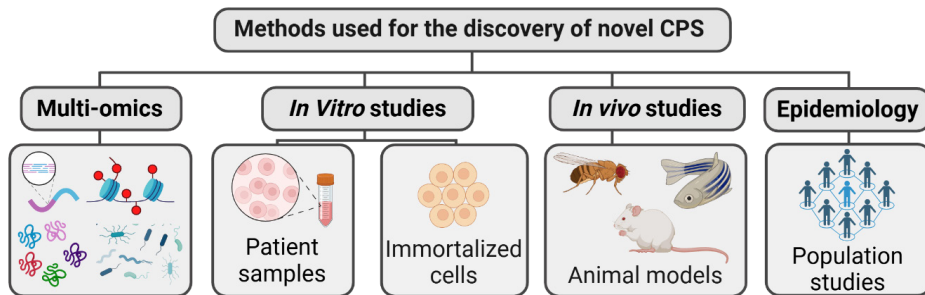


Figure 8. Selected methods used to study CPS. Graphical representation of selected methods for the study of gene–disease associations.

1.7.1 Multi-omics approaches

Apart from the ability to perform sequencing at a high throughput and a low cost, the last decade has also seen a revolution in the use of multi-omics approaches for diagnostic purposes. Additional layers of information can therefore be added to the study of genetic diseases, as cancer predisposition [60, 92]. These include, among others, the epigenetic landscape (epigenomics), gene expression and splicing isoforms (transcriptomics), peptide concentrations and interactions (proteomics), metabolite abundances (metabolomics), and the effects of our microbiome in disease (microbiomics) [134]. Compiling these aspects increases our ability to validate candidate disease-causing variants and confirm the involvement of novel genes in CPS.

1.7.2 CPS disease modelling

To understand the effects of genetic changes, researchers use disease models; biological systems mimicking some of the characteristics of the disease studied. These models can be divided into *in vitro* –mostly cell culture based–, and *in vivo* –using for instance, animals–. The main advantages of cellular models are their ease, low-price and the possibility to work with patient-derived material [135].

Animal models, on the other hand, enable more complex functional studies, including the effects on development, behavior, and body systems. These characteristics have been taken advantage of to study several diseases [136], including cancer [137].

An important model for the validation of candidate disease-causing variants in CPS are patient-derived cells, as they carry the genetic background from the patient (non-malignant cells) or patient's tumor itself (cells derived from tumor biopsies). These cells have a higher biological relevance than their immortalized counterparts –artificially manipulated to grow indefinitely–, although they are limited by their proliferation capacity [138]. The assays used to evaluate the effect of a variant depend greatly on the disease studied and the predicted consequences of the genetic change. Broadly, they include cell viability, proliferation, death, and senescence, along with specific analyses of the pathways involved.

Although cell lines are a pivotal tool during the initial stages of CPS discovery, animal models are uniquely suited to study the systemic and phenotypic effects of genetic changes. Long established *in vivo* models of human disease include the mouse (*Mus musculus*), the fruit fly (*Drosophila melanogaster*) and the Zebrafish (*Danio rerio*). These models have been widely used to study most genetic disorders, including cancer [139–141] and cancer predisposition [142, 143].

The use of models has revolutionized our ability to study disease. However, there are challenges in their adoption. First, intrinsic limitations to the use of research models. The main concern being failure to accurately mimic human disease phenotypes or predict therapeutic outcomes [144, 145]. Second, constraints linked to the study of cancer, as interspecies differences in carcinogenesis, distinct driver mutations and cancer genetics, and dissimilar metabolism and drug responses [141]. Finally, most CPS confer an increased risk of cancer, with few additional phenotypes. Simple biochemical tests to evaluate the pathogenicity of variants are usually lacking. Therefore, the combination of disease modelling, genetic tools, and epidemiology is of special importance for the study of CPS.

1.7.3 Epidemiological studies of childhood CPS

There are multiple ways in which epidemiological studies can be used in CPS research. For instance, to better understand the cancer risk in known syndromes, to discover novel CPS genes, and to find the prevalence of cancer predisposition in the general population. Registry-based studies can also be used to map

differences in the incidence of CPS and their cancer risk spectrum across populations. For instance, in myotonic dystrophy (**Study V**), we observed a partly overlapping, partly inconsistent, cancer risk spectrum in epidemiological studies of different populations.

The cancer incidence varies greatly among pediatric CPS. From a lifetime risk of 90% in Li-Fraumeni Syndrome [72] to a 3.5-fold increase in patients with Noonan syndrome due to *PTPN11* mutations [146]. Therefore, epidemiological studies to delineate cancer risks are urgent. In **Studies II** and **V**, we used information from the Swedish national registries to investigate the cancer incidence in individuals with PWS (**Study II**), muscular dystrophy and myotonic dystrophy (**Study V**).

Finally, the estimated CPS incidence in children with cancer varies considerably, from about 8% to 18%, depending on the gene panel used and the selected patient group (Table 2). To better calculate the incidence of known pediatric CPS, more studies of unselected cohorts –that is, all children diagnosed with cancer– are warranted. Evaluating broader gene panels in larger cohorts will also improve the estimations of the CPS incidence amongst pediatric patients with cancer and enable the discovery of novel CPS genes. With this aim, in **Study I**, we assembled a broad childhood CPS panel with research purposes, based on multiple published or publicly available sources of information.

1.8 Precision medicine

Precision medicine is a novel clinical approach where diagnosis and treatment are patient tailored. As cancer is a heterogeneous disease regarding presentation, aggressiveness, and treatment response, individualized treatments are especially useful, and precision medicine is slowly changing the paradigm in both diagnosis and treatment. For a long time, oncologists have understood that the approach “one-size-fits-all”, used until now, is at the best, suboptimal. New knowledge about molecular indicators for diagnosis, prognosis and treatment toxicity will enable the use of patient’s genetic information to improve their clinical outcomes [147], shifting cancer care towards individualized programs.

However, the development of precision medicine depends greatly on technological advances such as cost-effective genomic sequencing and an increased understanding of the genetic profile obtained. A big shift in this direction was the development of massive parallel sequencing. As described above and in contrast to initial sequencing methods, massive parallel sequencing

allows for high-throughput, scalable, and rapid sequencing at a relatively low price [148]. In fact, the fast improvement of these technologies had led to a decrease of six orders of magnitude in the sequencing costs within twenty years, reaching the 1000 USD/genome mark [149]. It is even expected that the price will soon reach 100 USD/genome [150, 151]. This economic leap has enabled the use of genetic information as a routine diagnostic tool, as well as the continuous implementation of novel diagnostic methods in the clinic [152].

1.8.1 Patient tailored cancer diagnosis and treatment

During the last 30 years, there has been a dramatic improvement in the treatment of children with cancer in most European countries, with an incredible rise from 30% to above 80% survival rate [153, 154]. Unfortunately, like with genetic diagnosis, this improvement is not matched around the globe, as the mortality rates in countries with low development index are doubled [155]. A better resource allocation and implementation of policies for equal access to cancer care will be imperative to enable the availability of personalized medicine in vulnerable populations [156].

New challenges in pediatric cancer treatment arise as mortality rates decrease, such as how to avoid and manage severe side-effects [153]. Chemotherapy and radiation, two of the most established therapies for cancer treatment, do not specifically target malignant cells, and thus damage normal tissues. Treatment sensitivity differs amongst individuals, resulting in highly variable acute and chronic side-effects. Around 2% of children with cancer die in the acute phase of treatment [157, 158]. Moreover, up to 40% of pediatric cancer survivors report severe, disabling, or fatal late effects, and more than 70% live with chronic health conditions [159].

An approach to decrease therapy-related complications is the development of diagnostic tools that can predict treatment toxicity or suggest targeted therapy options (Figure 9). The number of US cancer patients eligible for treatment with targeted therapies increased from 5.1% in 2006 to 13.6% in 2020 [160]. As of December 2024, according to data from the National Institutes of Health, more than 200 targeted cancer drugs were approved (www.cancer.gov/about-cancer/treatment/types/targeted-therapies/approved-drug-list). This number will continue to rise as novel therapeutic agents are developed and tested [161].

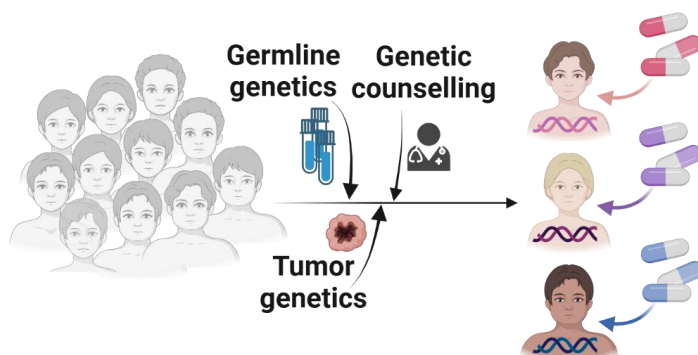


Figure 9. Targeted therapies for cancer treatment. Genetic analyses of paired tumor tissue and blood (i.e., the germline genome) are carried out. Results are then analyzed, and the patient is informed and genetic counselled. When available, personalized cancer treatment is offered.

Currently approved targeted cancer treatments include immunotherapies and immune checkpoint inhibitors. In the first group, cancer vaccines work by boosting the immune system to react to a specific tumor, while CAR T-cells are engineered to recognize and attack the tumor. On the other hand, immune checkpoint inhibitors block the dampening of the immune system around the tumor (See [162–164] for in-depth reviews). Finally, other types of personalized cancer treatments, such as monoclonal antibodies and small-molecule inhibitors, are designed to interfere with specific cancer drivers.

1.8.2 Cancer surveillance in patients with pediatric CPS

Besides treatment, childhood CPS diagnosis has important implications in cancer surveillance. As individuals with CPS have an increased risk of developing cancer, screening protocols have been introduced to detect malignancies at early stages. These protocols are specific to each diagnosis, depending on the cancer risk spectrum and age at onset. In highly penetrant pediatric CPS, surveillance has been linked to improved survival and reduced systemic toxicity [165].

Monitoring protocols are country specific. In the European Union, cancer surveillance is suggested for individuals with a cumulative cancer risk above 5% [165]. Although surveillance varies between member nations, consensus screening protocols have been published for the most common CPS [166, 167]. Similarly, the Pediatric Cancer Working Group of the American Association for Cancer Research has published expert suggestions for monitoring of pediatric CPS in the US [168].

2 Research Aims

The aim of this thesis was to identify novel genes, pathways and molecular mechanisms underlying cancer predisposition in children. This, with the long-term goal of identifying genetic risk factors for which treatment should be modified, surveillance started, and genetic counseling offered to the family.

The specific aims included:

- 1) **Generate a broad childhood CPS gene panel.** To develop a pediatric CPS panel with research purposes (**Study I**).
- 2) **Recognize new associations between congenital diseases and cancer.** To present clinical cases of children with known inherited syndromes, not previously associated with cancer, who developed cancer in childhood. Specifically, PWS (**Study II**), MFS (**Study III**), and LGMDR1 (**Study IV**).
- 3) **Investigate the cancer risk in patients with inherited diseases using information from the Swedish National Registries.** To determine the cancer risk and risk spectrum in patients with PWS (**Study II**), muscular dystrophy and myotonic dystrophy (**Study V**) through a literature review and population-based epidemiological studies.
- 4) **Characterize a novel genetic syndrome.** To decipher the molecular mechanisms leading to a novel multisystemic disease including leukemia, immunodeficiency, cognitive and metabolic symptoms, caused by compound pathogenic variants in the *FLCN* gene (**Study VI**).

Ultimately, we hope that the results presented in this thesis will improve the diagnosis and contribute to the understanding of pediatric CPS.

3 Materials and Methods

"In God we trust. All others must bring data" – W. Edwards Deming.

The workflow used across multiple studies in this thesis initiated with the diagnosis of pediatric cancer, followed by tumor histological assessment and generation of paired genetic data from the germline and the tumor. Stand-alone projects emerged when patients with congenital syndromes, not previously associated with cancer development, were diagnosed with cancer at the clinic.

Individual studies included the design of a CPS research gene panel (**Study I**), reporting of clinical cases (**Studies II – IV** and **Study VI**), molecular investigations (**Studies II, III** and **VI**), and register-based studies (**Studies II** and **V**). A general description of the methods is included below. For in-depth information, please refer to the methods section of the respective paper.

3.1 Study participants

3.1.1 ChiCaP Cohort

Between May 2021 and December 2022, all pediatric patients with solid tumors in Sweden were prospectively included in the ChiCaP project (**Study VII**). As part of the ChiCaP workflow, paired blood- and tumor-derived material was collected from the patients. Clinical information about CPS risk factors was also collected, including family history of cancer, previous primary malignancies, and congenital syndromes or other symptoms. Selected patients in the ChiCaP study previously diagnosed with congenital syndromes not linked to cancer were included in **Studies II – IV**. An additional patient outside the ChiCaP study, who had a novel congenital syndrome and developed leukemia, was also included in this thesis (**Study VI**).

3.1.2 Population-based studies

In **Studies II** and **V**, epidemiological information from the Swedish National registries was used to understand the incidence of different cancer types in PWS (**Study II**), muscular dystrophy and myotonic dystrophy (**Study V**). For this, congenital diagnoses were collected from the Swedish National Patient Register [169], the Medical Birth Registry [170] and from the Karolinska University laboratory information system. Further, information on cancer diagnoses was retrieved from the National Cancer Registry [171]. Table 3 summarizes the registers used in our population-based studies, and the information provided by each record.

Table 3. Swedish registers used in the epidemiological studies

Registers	Information used in Studies II and V	Reference
Swedish National patient Register	Congenital diagnoses. Details on inpatient and outpatient stays since 1987 and 2001.	[169]
Medical Birth Register	Congenital diagnoses in newborns. Coverage since 1973.	[170]
National Cancer Register	Cancer diagnoses, including age at onset, tumor site and histology. Coverage since 1958.	[171]
Total Population Register	Information on birth, death, and emigration dates, available since 1968.	[172]
Longitudinal Integration Database for Health Insurance and Labor Market Studies	Educational level (Patient's parents). Compiled information from 1990.	[173]
Multi Generation Register	Information across generations, linking patients and their parents. Individuals registered from 1961 and born after 1931 are covered.	[174]

Individuals with specific diagnoses were identified according to ICD-10, the 10th Revision of the International Statistical Classification of Diseases and Related Health Problems, which classifies diseases using unique codes. Patients diagnosed with the codes for PWS (**Study II**), muscular dystrophy and myotonic dystrophy (**Study V**) were included in the study and matched to 50 unaffected individuals, using the total population register. The association between the evaluated congenital disease and cancer was estimated using Cox Proportional Hazards; a regression model which calculates the probability of a hazard (e.g., cancer), considering covariates. Covariates differed between studies, including for instance birthyear, sex, parental education level and parental age.

3.1.3 Ethical approval

All studies included in this thesis were approved by the Regional Ethical Review Board in Stockholm, Sweden (Dnr numbers 2015-292-31-4, 2015-608-31-4, 2018-1849-32, 2019-04746, 2020-03827, 2021-05916-02, 2022-04349-02, 2023-0754-02). Additionally, written informed consent was obtained from each patient or their legal guardians prior to inclusion.

3.2 Compilation of a pediatric CPS gene panel

To compile a childhood CPS panel, we evaluated genes from multiple sources including nine publicly available gene panels [60, 66, 89, 91, 92, 132, 175, 176] and two cancer databases [177, 178]. Additionally, a set of genes was collected from case observations and conference visits.

The resulting list was ranked according to each gene's appearance in the sources of evidence presented above, and any associated phenotype description in Ensembl/BioMart [179, 180]. As some intellectual disability and primary immunodeficiency syndromes have been associated with cancer predisposition [181, 182], Genomics England panels [175] and inferred gene associations at the Human Phenotype Ontology database (<https://hpo.jax.org>) for these two groups of syndromes were also used for the ranking, resulting in a Total Gene Rank Score (TGRS). Finally, evolutionary constraint was evaluated using the loss-of-function observed vs. expected upper bound fraction (LOEUF) metric, which expresses a gene's intolerance to LoF variation. The cut-off for gene constraint was set at $LOEUF \leq 0.35$ according to [183].

3.3 Histological assessment of the tumors

Diagnostic biopsies were obtained from patient's tumors in **Studies II – IV**. Pathology examination was carried out by collaborators at the Karolinska Institutet; including Hematoxylin and eosin staining and immunohistochemistry with antibodies tailored to the diagnosis of the specific tumor type.

3.4 Culture of human dermal fibroblasts

Dermal fibroblasts isolated from skin biopsies were used in **Studies III and VI**. Cells were cultured in Dulbecco's modified eagle's medium with high glucose, supplemented with non-essential amino acids, 10% fetal bovine serum and 0.2% Primocin at 37°C and 5% CO₂. These dermal fibroblasts were used for cellular staining, and downstream methods requiring DNA, RNA or proteins.

3.5 Genome sequencing and molecular methods

3.5.1 DNA sequencing

3.5.1.1 *Germline and Somatic Whole Genome Sequencing*

As mentioned in the background, WGS is a genetic tool that decodes the entire 3.2 billion nucleotides in the human genome. In this thesis, WGS from paired blood

and tumor samples (for solid malignancies) was carried out in all patients (**Studies II – IV** and **Study VI**). Results from blood conferred information about inherited DNA changes, while tumor material was used to generate a genetic profile of the malignancy. This, to determine tumor aggressiveness, potential targetable DNA changes, and further specify the cancer type. Sometimes, somatic changes in the tumor can also be used to confirm a CPS diagnosis, for instance through tumor second hits or mutational profiles associated with specific CPS (**Study VII**).

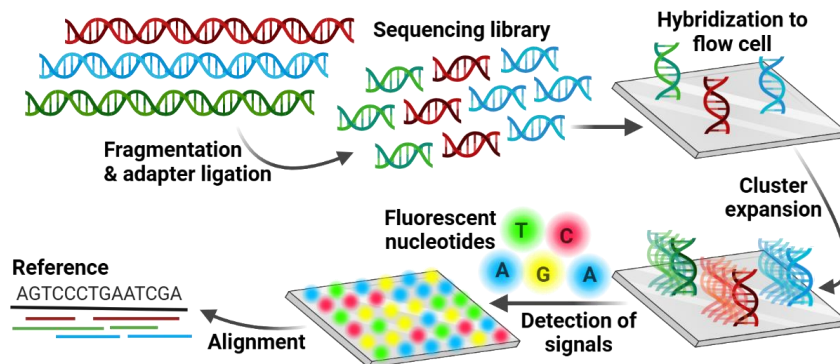


Figure 10. Short read whole genome sequencing workflow used. A sequencing library is built by fragmenting input DNA and tagging it with short DNA pieces (adapters). These adapters bind their complementary sequences onto a flow cell –a glass slide–, preparing DNA for *in-situ* cyclic amplification. Upon amplification, a cluster with the original sequence is formed, and sequencing can start. During each step, fluorescently marked nucleotides are added to the flow cell. As the correct nucleotide joins the growing DNA strand, its fluorescent signal is detected. A new set of nucleotides is then added, and the process continues until the entire DNA strand has been read. This process takes place in parallel for all generated clusters in the flow cell. Further, sequences are aligned to the reference genome, and deviations from the reference are identified as variants.

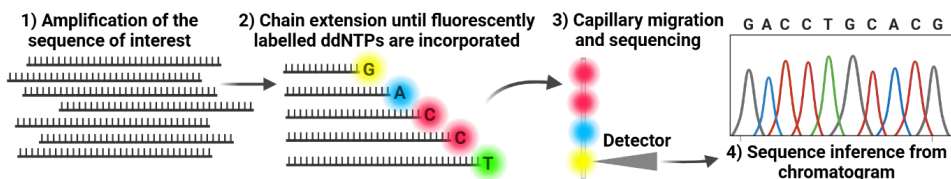


Figure 11. Sanger Sequencing Overview. The DNA sequence of interest is amplified. Then, the template is sequenced with a mixture of unaltered and modified fluorescently labelled DNA building blocks. Amplification continues until blocked by the modified bases, generating labelled molecules of different lengths. Following, DNA is loaded into capillaries where it migrates according to size. The fluorescence of the modified bases is then read by detectors, starting with the shortest molecule. That is, the first nucleotide of the strand. The DNA sequence can be inferred from the resulting colorful peaks (chromatogram).

Specifically, pair-end Illumina short-read WGS, with a coverage of 30x in blood and 90x in tumors was used (Figure 10). Once generated, genetic data was ranked, visualized, and filtered in the Scout platform from Clinical Genomics. Germline data was processed using the Mutation Identification Pipeline framework [184], while tumor data was analyzed by collaborators at the Department of Oncology-Pathology, at the Karolinska Institutet. Germline candidate variants were manually explored in Scout or with the Integrative Genomics Viewer [185] program.

In **Studies II, III** and **VI**, variants were confirmed by Sanger Sequencing (Figure 11).

3.5.2 Gene expression assays

3.5.2.1 RNA sequencing

In **Study IV**, tumor RNA was used to evaluate the presence of a cancer driver translocation. In **Study VI**, RNA sequencing (RNA-seq) was performed to analyze the gene expression profile of a patient with a novel syndrome. Broadly, RNA-seq works like WGS (Figure 9), except that the starting genetic material is RNA, which carries the information to make proteins, among others. In this way, sequencing results are a snapshot of the status of the cells when RNA was extracted. RNA expression is tissue type dependent, and it reflects the metabolic state of the cell.

3.5.2.2 Droplet Digital PCR

The Polymerase Chain Reaction (PCR) is a tool used to amplify a selected piece of DNA in a sample. It takes advantage of an enzyme called polymerase, which can elongate sequences given a starting DNA segment. Droplet digital PCR (ddPCR) is a novel version of this technique, which enables absolute quantification of the DNA molecules. In other words, it answers the question: how many molecules of my DNA of interest are there in the sample? For this, DNA is diluted and separated into tiny oil droplets, such as there is approximately one particle per drop. PCR amplification with fluorescently labelled primers takes place in each droplet and the count of positive signals is converted into DNA particle numbers (Figure 12).

In **Study III**, ddPCR was used to quantify the expression of novel splice isoforms in patient skin fibroblasts, as the pathogenic variant found was suspected to affect splicing (See the results section for more information on splicing). In **Study VI**, ddPCR enabled exact quantification of selected differentially expressed genes in the patient identified by RNA-seq.

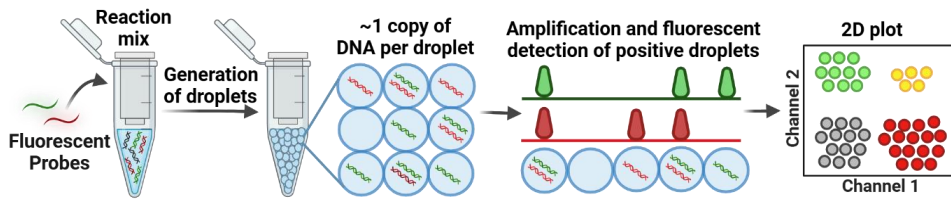


Figure 12. Overview of droplet digital PCR. A reaction mix including DNA and fluorescent probes targeting the genomic region of interest is prepared. Next, a droplet generator is used to partition the sample into ~20.000 droplets, each with about one copy of DNA. The sequences of interest are amplified inside each droplet, which is then read as positive or negative according to the quantified fluorescence. Finally, the results are plotted, and the absolute number of DNA particles is statistically calculated.

3.5.3 Methylation analyses

3.5.3.1 Methylation-specific MLPA

The acronym MS-MLPA stands for Methylation-specific Multiplex ligation-dependent probe amplification, and it refers to a molecular assay used to study the copy number and methylation profile of a selected genomic region, including multiple sites. MS-MLPA relies on the use of methylation-sensitive restriction enzymes, which digest unmethylated DNA at specific sites. Undigested DNA pieces are amplified and separated according to their length in capillaries. This separation is converted into signals, and the copy number and methylation percentage inferred from the ratio between a standard and the target (Figure 13).

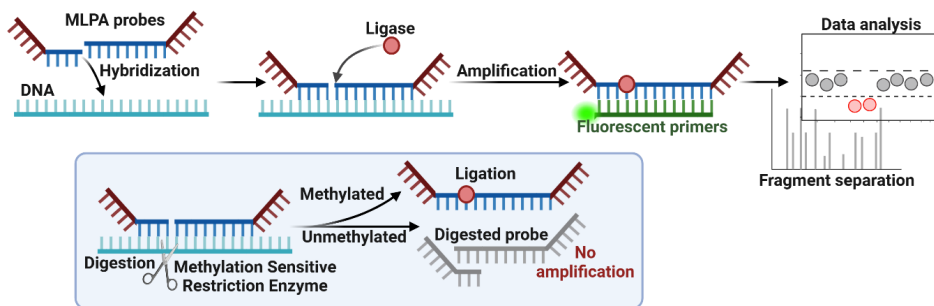


Figure 13. The principle of MS-MLPA. Each MLPA probe has two pieces which hybridize at a specific DNA location. If both sides attach correctly, they can be sealed by a ligase enzyme. Only ligated probes are amplified with fluorescent primers. For the probes that measure methylation (Boxed panel), digestion by a restriction enzyme will help distinguish unmethylated alleles (digested, and thus not amplified), from methylated forms (not digested). All probes have different lengths and can thus be sorted in capillaries according to their migration speed. In the capillaries, fluorescence is measured and plotted by detectors. Finally, the size of the peaks is quantified and converted into methylation ratios.

As described in the background, PWS is an imprinting disorder with aberrant methylation patterns at a specific genomic region, leading to abnormal expression of the encompassed genes [186]. In **Study II**, we used MS-MLPA to compare the methylation patterns at the PWS region and control imprinted sites in patient's blood and tumor tissue.

3.5.4 Protein detection

3.5.4.1 Western blotting

Western blots are based on the recognition of a target protein by an antibody. Antibodies are molecules produced by the immune system to mark foreign objects. In this case, the antibodies are generated to recognize target proteins (Figure 14, Upper panel). In **Study VI**, western blot was used to quantify the expression of multiple proteins in pathways possibly affected in the patient.

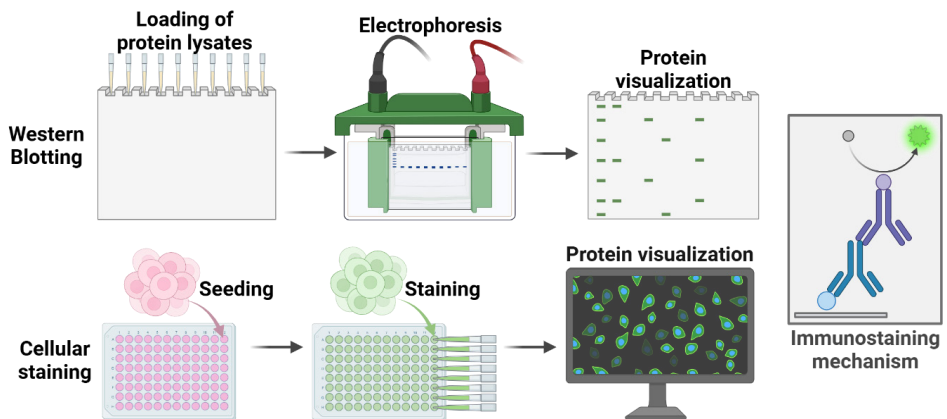


Figure 14. Methods for protein detection using antibodies. In Western Blotting (Upper panel), cell lysates are loaded into nitrocellulose membranes, and proteins separated according to size. For visualization, membranes are incubated with antibodies marked with chemiluminescent reagents. In immunofluorescent staining (Lower panel), cells are seeded, stained with fluorescently labelled antibodies, and imaged using a fluorescence microscope. In both methods, we used an indirect mechanism for detection, which is shown on the right. In short, a specific antibody is used to recognize the protein of interest. A second antibody, labelled with a luminescent tag, recognizes the first one. Finally, the tag is activated resulting in a colorimetric or fluorescent product.

3.5.4.2 *Immunofluorescent staining*

Immunofluorescent staining also uses antibodies to mark target proteins. However, in this case, cells are usually seeded on glass slides, marked with fluorescently labelled antibodies, and visualized with a fluorescent microscope (Figure 14, Lower panel). In **Study VI**, immunofluorescent staining was used to quantify the proteins of interest in patient's skin fibroblasts.

4 Results and Discussion

“That’s why we do science, because every now and then there’s this incredible joy of figuring something out” – Jennifer Doudna.

The main results of this thesis are:

1) **Generation of a broad childhood CPS gene panel.**

In **Study I**, we compiled a research panel of 881 genes associated with childhood cancer predisposition, and developed a ranking system that can be used to prioritize genes based on the evidence for their association with CPS. Using LOEUF as an indicator of mutational constraint, we found that 43.4% of the genes in our panel are constraint, as compared to 15.6% of all genes with LOEUF metrics.

2) **Recognition of new associations between congenital diseases and cancer.**

In **Study II**, we describe the case of a 13-year-old girl with PWS who developed an ovarian dysgerminoma and bilateral ovarian sex cord tumors with annular tubules. Additionally, MS-MLPA revealed locus-specific LOI at the PWS locus in the dysgerminoma.

In **Study III**, we report two females with MFS who developed the pediatric tumor neuroblastoma. Patient 1 presented with neonatal MFS, the most severe form of the disease, and was diagnosed with neuroblastoma at the age of 10 months. Patient 2 had classical MFS, and developed neuroblastoma at the age of 18 years. We suggest a possible association between neuroblastoma development and MFS.

In **Study IV**, we describe a 17-year-old boy with LGMDR1 caused by compound heterozygous variants in the *CAPN3* gene, who developed a desmoplastic small round cell tumor. Interestingly, mouse models of different muscular dystrophies also develop soft tissue sarcomas, indicating a possible association between muscular dystrophies and this tumor type, which was further investigated in **Study V**.

3) **Investigation of the cancer risk in patients with inherited diseases using information from the Swedish National Registries.**

In **Study II**, we explored the cancer risk in patients with PWS. No overall increased cancer risk was found, although our study suggests a possible pediatric cancer risk increase. We also observed a large proportion of germ cell tumors in young individuals with PWS.

In **Study V**, we presented the first report of an increased risk of pediatric astrocytomas and other gliomas, and adult nonthyroid endocrine and pancreatic cancer in patients with muscular dystrophy. In myotonic dystrophy, we confirmed the previously reported increased risk of brain tumors in children. We also confirmed an increased risk of cancer overall in adults, explained by an overrepresentation of diverse malignancies.

4) **Characterization of a novel genetic syndrome.**

In **Study VI**, we found a homozygous variant in the *FLCN* gene (*FLCN* p.G15S) in a patient with a novel multisystemic syndrome including immunodeficiency, global developmental delay, and dysmorphic features, who developed B-cell leukemia at 1-year of age. We then carried out *in silico* cellular and molecular analyses to show that the variant likely results in hypomorphic folliculin function, leading to this novel syndrome.

A summary of the most important findings from the studies is presented in the following sections. In depth results can be found in the respective publications, at the end of this thesis.

4.1 Compilation of a broad childhood CPS gene panel

With the long-term goal of discovering known and novel pediatric CPS, in **Study I** we compiled a research gene panel for childhood cancer predisposition. The panel was developed with information from publicly available gene panels [60, 66, 89, 91, 92, 132, 175, 176], cancer databases [177, 178], the Human Phenotype Ontology database (<https://hpo.jax.org>), and case observations.

The final panel comprises 881 genes, ranked according to a system that prioritizes genes with established evidence for their association with childhood cancer predisposition. The resulting TGRS ranges from 0.0 to 30.0, with an average of 8.66 (95% Confidence Interval [CI] 8.20–9.12). Some genes with a TGRS < 12 were included from literature reviews or conferences and thus present the lowest confidence for their involvement in CPS. Moreover, all genes with TGRS \geq 12 (n=199) have a known association with cancer predisposition, confirming the usefulness of the ranking system for gene prioritization during CPS genetic analyses.

Finally, we evaluated the mutational constraint using LOEUF metrics as an index. We found that 43.4% of genes in our panel were constraint (mean LOEUF 0.56), as compared to 15.6% of all genes with constraint metrics (mean LOEUF 0.95) [183]. This is in line with the hypothesis that genes leading to early onset severe illnesses, such as pediatric CPS, are often constraint [187].

In summary, in **Study I** we compiled a pediatric CPS panel with 881 genes, which can be used as a resource for the discovery of novel childhood CPS.

4.2 The ChiCaP project

As mentioned in the background, CPS diagnosis is of great importance due to clinical implications such as access to surveillance, genetic counselling, and tailored treatments. Therefore, the ChiCaP project, a national initiative to improve the diagnosis of pediatric CPS, started in Sweden in 2021. ChiCaP is a prospective study that combines the analysis of clinical and genetic data to identify pediatric cancer predisposition. The study included 309 children with solid malignancies diagnosed across the country and found a CPS prevalence of 11% in the 189 genes evaluated (**Study VII**).

Additional branches of the ChiCaP project are underway, including the analysis of CPS in children with leukemia, and studies of the psychosocial effects of genetic testing in pediatric CPS. In addition, WGS results from children with congenital syndromes previously unrelated with cancer were analyzed separately as part of

Studies II – IV, as presented in the following sections. Finally, a rare variants association study from germline WGS of pediatric cancer patients is in progress.

4.3 Cancer risk in congenital syndromes studied in this thesis

4.3.1 Prader–Willi Syndrome

In **Study II**, we investigated the cancer risk in individuals with PWS using information from the Swedish National Registries, and presented the clinical, molecular and genetic characteristics of a patient with PWS who developed multiple ovarian tumors (Figure 16).

In total, 360 patients diagnosed with PWS in the National Registries between 1997 and 2017 were included in the epidemiological part of the study, each matched to 50 comparisons by year of birth, sex, and birth county. The overall cancer incidence was similar between patients and comparisons (3.33% vs. 3.07%, respectively). Although we observed a high number of pediatric cancer cases in patients with PWS (3 [0.83%] vs. 48 [0.27%] in matched comparisons), the numbers were too low to perform further statistical testing. Similarly, the proportion of germ cell tumors among all individuals with cancer was increased in PWS (2/12 [17%] vs. 17/551 [3%] in matched comparisons), but further analyses were precluded by the low case number. In a literature review of PWS and cancer, we found 50 described patients. Cancer development at a young age (Average 24.5 years) and a high number of germ cell tumors (22.5%) were again observed.

Multiple lines of evidence have suggested an increased risk of testicular cancer in PWS [98, 99, 188]. In line with these reports, we observed a high number of testicular cancers and ovarian tumors in this patient group. Some hypotheses for this association include gonadal dysgenesis, and a high incidence of undescended testis in males with PWS [189]. Epidemiological studies on larger cohorts are needed to confirm all presented associations.

In the second part of the study, we describe the case of a girl with PWS due to a paternal deletion in chromosome 15 (15q11.2–q13). At 13 years of age, she developed a dysgerminoma –cancer of female germ cells– and multiple bilateral Sex cord tumors with annular tubules –rare, usually benign, tumors in the ovarian sex cords–. WGS of the dysgerminoma showed a complex genetic profile, with changes in chromosomal number and translocations at multiple genetic locations. A cancer driver mutation in the *KIT* gene (p.V559G) was detected. No pathogenic variants in CPS genes were found, when filtering with the panel used in **Study VII**.

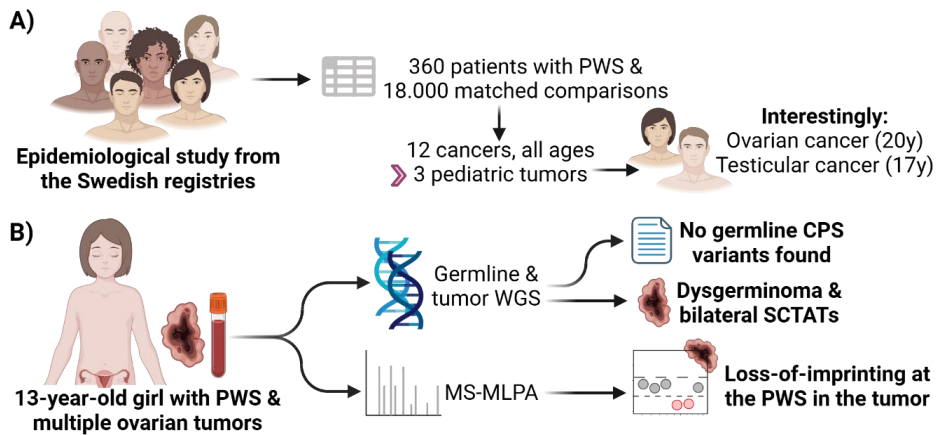


Figure 16. Graphical representation of Study II. A) Overview of the population-based study of cancer in patients with PWS. B) Pipeline for the genetic investigation of a girl with PWS who developed a dysgerminoma and bilateral Sex cord tumors with annular tubules (SCTATs) at the age of 13 years (y).

As explained in the background, PWS is an imprinting disorder. As imprinting changes have been previously described in cancer [104, 190–192], including in the germ cell tumor of a patient with PWS [105], we decided to investigate the imprinting status in the dysgerminoma of the patient. MS-MLPA results confirmed a partial LOI in the tumor at the PWS region.

In conclusion, in **Study II** we carried out a register-based study and literature review of cancer in PWS. Although we did not find an overall cancer risk increase in this patient group, we observed a high frequency of pediatric cancer and gonadal tumors. We also presented the case of a 13-year-old girl with PWS who developed a dysgerminoma and multiple bilateral Sex cord tumors with annular tubules. Interestingly, we reported the second case of LOI at the PWS locus in a germ cell tumor of a patient with PWS. Further studies are needed to understand the impact of LOI at the PWS region in cancer development in these patients.

4.3.2 Marfan Syndrome

In **Study III**, we characterized the first reported patients with MFS who developed neuroblastoma, a pediatric tumor of early nerve cells (Figure 17). We also reviewed the literature on the co-occurrence of MFS and cancer and looked for *FBN1* variants in pediatric tumors using cancer databases.

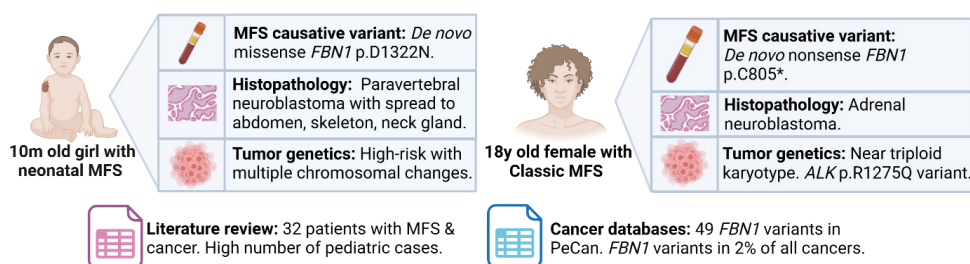


Figure 17. Main findings of Study III. (Upper panels) Patients germline and tumor genetic and histopathological results. (Lower panels) Main findings from literature review on Marfan Syndrome and cancer (left) and reported *FBN1* variants in cancer databases (right). In the upper panels, m refers to months and y to years.

The first patient described was a girl with neonatal MFS, caused by a *de novo* –not inherited– missense variant in *FBN1* exon 32 (p.D1322N). This variant is predicted to affect splicing, a term referring to the process in which RNA pieces are cut out of the final molecule used as a template to create proteins. Sanger Sequencing and ddPCR were used to detect and quantify the new splicing forms of the protein. We observed both an isoform with exon 32 skipping and one with complete intron 32 retention (accounting for 13% and 7.6% of total *FBN1* RNA, respectively). Further, the patient’s adrenal neuroblastoma was diagnosed at 10-months-of-age and WGS showed unfavorable tumor genetics with multiple chromosomal changes. Unfortunately, the patient died at 19-months due to cardiac failure, associated with her MFS.

The second patient was a woman with MFS caused by a *de novo* *FBN1* nonsense variant (p.C805*) –a mutation resulting in an early stop in RNA translation. She was diagnosed at 18-years-of-age with a paravertebral neuroblastoma with metastatic spread. Tumor WGS detected an activating mutation in the cancer driver *ALK* (p.R1275Q), a near triploid copy number, and multiple chromosomal changes. No known pathogenic CPS variants were found in the patients, using the gene panel described in **Study VII**.

In our literature review of MFS and cancer, we found 32 cases, 30% with pediatric presentations. Epidemiological studies are warranted to further study the association between childhood cancer and MFS. Next, we queried the PeCan [193] and cBioPortal [194] cancer databases for reported somatic *FBN1* changes. 49 variants were present in PeCan, 15 of which were germline pathogenic for MFS [195]. However, their effect on cancer development is uncertain. In all non-redundant datasets in the cBioPortal, *FBN1* was altered in 2% of patients.

All in all, in **Study III** we present the first two reported patients with MFS who developed the pediatric tumor neuroblastoma and highlight an early age at cancer diagnosis in reported patients with MFS. Epidemiological studies are needed to clarify the growing evidence linking MFS and pediatric cancer development.

4.3.3 Limb-Girdle Muscular Dystrophy Recessive 1 (LGMDR1)

In **Study IV**, we reported the case of a boy with muscular dystrophy who developed a soft tissue sarcoma –a tumor originating from connective tissue– (Figure 18). During infancy, the patient was diagnosed with LGMDR1, a type of muscular dystrophy characterized by progressive weakness on the proximal and shoulder girdle muscles. The disease-causing variants were two heterozygous pathogenic changes in the *CAPN3* gene; a frameshift variant (p.T184Rfs*36) and a missense variant (p.R448C).

At the age of 17 years, the patient was diagnosed with a type of soft tissue sarcoma known as desmoplastic small round cell tumor. One year after the initial diagnosis, the boy developed a first relapse in his left hip, and a second relapse in the scalp. RNA-seq revealed a diagnostic driver *EWSR1::WT1* fusion gene in the primary and scalp tumors. In addition, both malignancies presented a near triploid copy number with multiple chromosomal changes. No germline pathogenic variants in the gene panel described in **Study VII** were found.



Figure 18. Overview of Study IV. Summary of the main genetic and histopathology results from the case report of a boy with LGMDR1 who developed multiple desmoplastic small round cell tumors at the age of 17 years (y).

The co-occurrence of LGMD and a soft tissue sarcoma was interesting, as studies on mouse models of multiple muscular dystrophies show an increased sarcoma risk [118–121, 196, 197], including in a model for LGMDR1 [121]. However, based on the information from this case report, it is not possible to understand the sarcoma risk in individuals with LGMD. We further explore the cancer incidence in patients with muscular dystrophies in **Study V**.

In summary, **Study IV** presents the clinical and genetic findings of the first described patient with LGMD who developed a desmoplastic small round cell tumor. It is also the first case of cancer in LGMDR1 reported in the literature.

4.3.4 Muscular dystrophy and myotonic disorders

In **Study V**, we investigated the incidence of cancer in individuals with muscular dystrophy and myotonic dystrophy, through a literature review and information from the Swedish National Registries. 2355 patients with muscular dystrophy and 1968 with myotonic dystrophy were included in the epidemiological study, each matched to 50 comparisons by year of birth, sex, and birth county (Figure 19).

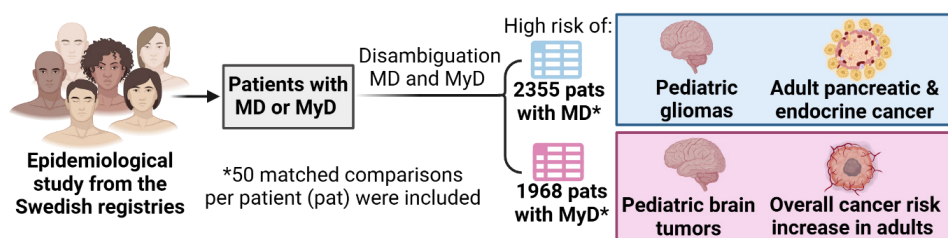


Figure 19. Outline of Study V. Information from the Swedish National Registries about individuals with muscular dystrophy (MD) and myotonic dystrophy (MyD) was used to evaluate their cancer risk spectrum. Main result in the panels to the right.

In muscular dystrophy, we did not find an overall cancer risk increase. However, patients presented a high risk for pediatric astrocytomas and other gliomas (Hazard Ratio [HR] 8.70, CI 3.57–21.20) as well as adult pancreatic cancer (HR 4.33, CI 1.55–12.11) and nonthyroid endocrine cancer (HR 2.35, CI 1.03–5.34). On the other hand, in the myotonic dystrophy cohort, we observed an overall cancer risk increase in adults (HR 2.26, CI 1.92–2.66), explained by a statistically significant increase of CNS tumors –34/39 (87.2%) of which were astrocytomas and other gliomas–, endocrine, endometrial, ovarian and nonmelanoma skin cancers. We also observed an increased risk of pediatric brain tumors (HR 3.23, CI 1.16–9.01).

Remarkably, the cancer risk spectrum in patients with muscular dystrophy and myotonic dystrophy is similar, including an increased risk of brain tumors and endocrine malignancies. Studies with stricter differential diagnosis between muscular dystrophy and myotonic dystrophy will be needed to strengthen this observation. Additionally, robust population-based studies are warranted to understand the associations between specific muscular dystrophies and cancer.

Our literature review included 121 articles on patients with cancer and muscular dystrophy or myotonic dystrophy. Interestingly, we observed a high number of soft tissue sarcomas in reported patients with DMD (9/23 [39%]). As previously mentioned, this is in line with results from mouse models [118–121, 196, 197]. Moreover, we did not observe an enrichment in sarcoma presentations in patients with muscular dystrophy in the Swedish National Registries, indicating that their overrepresentation in the literature may be a publication bias.

In conclusion, in **Study V** we carried out a literature review and an epidemiological study of cancer in muscular dystrophy and myotonic dystrophy. Although no overall cancer risk increase was observed in muscular dystrophy, a high number of pediatric astrocytomas and other gliomas and adult nonthyroid endocrine and pancreatic cancer was observed in this patient group. On the other hand, as in previous epidemiological studies, we observed an increased overall cancer risk in adults with myotonic dystrophy, as well as a high risk of brain tumors in children.

4.4 Discovery and characterization of a novel syndrome

4.4.1 Recessive *FLCN*-related disorder

In **Study VI**, we worked as detectives of the genome, seeking to find the cause of a novel disease in an 18-year-old boy. The patient had multiple symptoms, including intellectual disability, global developmental delay, immunodeficiency, and acute B-cell leukemia at 1-year of age. To solve this mystery, we carried out germline WGS of the patient and his parents. Sequencing analyses revealed a homozygous missense variant in the *FLCN* gene (p.G15S) inherited from both parents (Figure 20, upper panel). *In silico* tools predicted this variant as pathogenic, and computational modelling indicated a destabilization of the lysosomal FLCN complex.

As explained in the background, *FLCN* is the disease-causing gene for BHDS, an autosomal dominant disorder characterized by skin changes, lung collapse and kidney cancer predisposition [126]. Interestingly, neither the patient nor his parents present BHDS. However, animal models and patients with biallelic defects in *FLCN* or *FNIP1* –one of *FLCN*'s partners– develop symptoms resembling the clinical presentation of our patient, including immunodeficiency [198–205], metabolic defects [206, 207], and changes in brain development [208]. Further, patients with mosaic variants in the *TFE3* gene –a transcriptional activator inhibited by the FLCN complex– also present a similar, but more severe, phenotype to our patient [209, 210] (Figure 21).

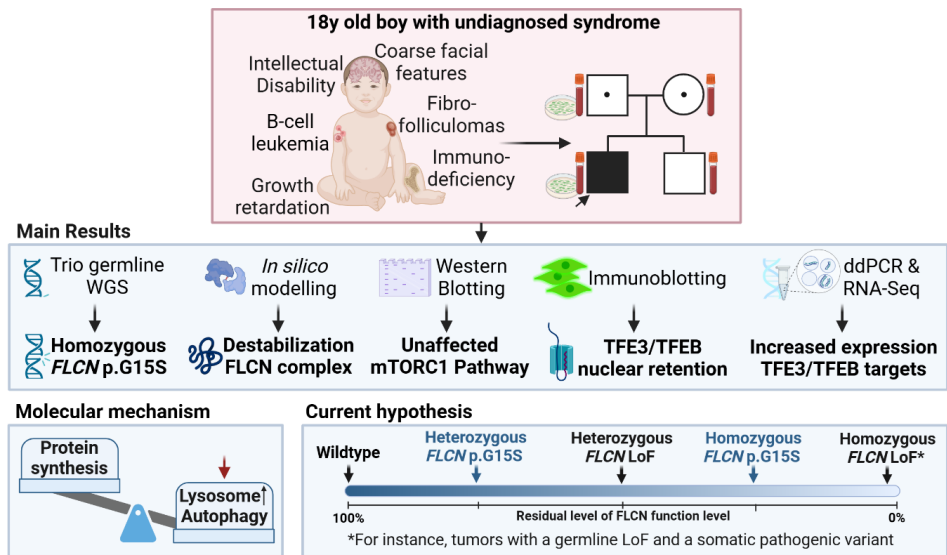


Figure 20. Main findings of Study VI. (Upper panel) Case presentation of an 18-year (y)-old boy with a novel syndrome. (Middle panel) Main results from the genetic, cellular, and molecular analyses carried out to understand the disease pathology. (Lower panel, left) Molecular mechanism of disease. Homozygous *FLCN* p.G15S leads to a metabolic imbalance by an increase in lysosomal biogenesis and autophagy. (Lower panel, right) Current hypothesis on *FLCN* residual function level. In the figure, the black arrow in the pedigree indicates the proband.

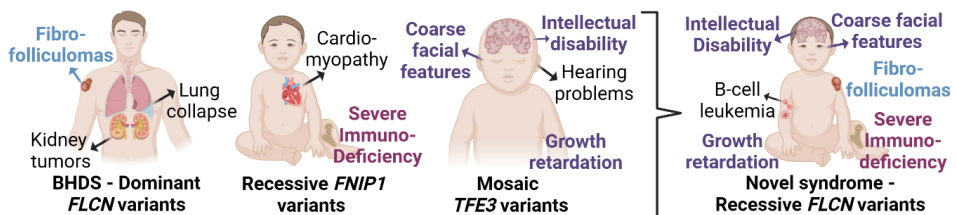


Figure 21. Clinical presentation of selected syndromes in *FLCN* related pathways. From left to right, Birt-Hogg-Dubé (BHDS), *FNIP1* autosomal recessive disorder, patients with *TFE3* mosaic variants, and novel *FLCN*-related disorder. Shared phenotypes in color.

The *FLCN* gene encodes a master modulator of metabolism with multiple cellular roles, including regulation of lysosomal biogenesis, autophagy, and mitochondrial synthesis [124]. We investigated the effect of *FLCN* p.G15S in two of the axes orchestrating these effects. Namely, the mTORC1 and TFE3/TFEB pathways. Analyses on skin fibroblasts did not show changes in the activation of downstream effectors of the canonical mTORC1 pathway. On the contrary, nuclear translocation of the transcriptional activators TFE3 and TFEB was increased in the patient. We also observed an overexpression of their target genes in patient

fibroblasts (Figure 20, middle panel). This increased activation possibly results in metabolic imbalances leading to lysosomal biogenesis and autophagy (Figure 20, lower panel left).

We hypothesized that *FLCN* p.G15S acts as a hypomorphic variant, meaning that the protein loses only part of its function (Figure 20, lower panel). The activity retained allows for homozygous variants to be viable, although *FLCN* autosomal recessive LoF variants are embryonic lethal. Furthermore, the remaining function possibly prevents the development of BHDS in heterozygous individuals.

In summary, in **Study VI** we characterized a novel autosomal recessive immunodeficiency syndrome with intellectual disability, short stature, dysmorphic features, and leukemia, caused by hypomorphic variants in the *FLCN* gene.

5 Ethical considerations

The rapid advancement in the field of genetic diagnosis carries with it important ethical concerns. The following paragraphs will briefly describe the main aspects of consideration. However, the ethical implications of CPS testing are various and complex. For sample case discussions, refer to [211].

One of the current areas of debate amongst childhood CPS diagnostic centers is which genes should be included in *in vitro* pediatric CPS gene panels. Although failing to detect important CPS variants may be detrimental for patients' clinical care, overdiagnosis can lead to significant psychological distress [212, 213], often without risk-reducing interventions. As previously mentioned, surveillance is recommended in Europe for a cumulative cancer risk over 5% by the age of 20. Moreover, it can be discussed with the family when the associated risk is 1–5% [214]. However, the penetrance of many CPS genes has not been calculated, and the benefits of surveillance remain to be assessed for most syndromes [215]. Weighing the benefits and burdens of genetic diagnosis is necessary for each CPS.

Access to genetic data must be accompanied by a clear understanding of the implications of the results obtained. Therefore, an increase in the availability of genetic tests must be paired with an increment in genetic counselling. CPS diagnoses impact not only the life of patients, but also of their relatives. Therefore, genetic counselling for the individual, parents, and family members at risk must be offered. Genetic counsellors and clinical geneticists accompany the family throughout the process of understanding the given genetic diagnosis and its implications. For instance, education about the genetic diagnosis, the risk of cancer occurrence and recurrence, what genetic tests can tell –and not–, how to make informed healthcare choices, among others [216].

Another aspect of concern, especially in the diagnosis of pediatric CPS, is the testing of underage siblings. For the patient, diagnosis is clinically relevant, and therefore genetic testing is justified. This is not always the case for family members at risk. Important aspects to be considered for cascade testing include the age of the individuals to be tested, the penetrance of the CPS, and the clinical actionability of a positive answer. Guidelines for genetic testing of asymptomatic minors from the European Society of Human Genetics are available for inherited syndromes [217]. Further, depending on the inheritance pattern of the disease

tested, some individuals may receive indirect genetic information about themselves. Genetic counselling should specify these possibilities to the family.

Finally, germline testing can result in secondary or incidental findings. A secondary finding refers to a result that is purposefully screened during genetic testing, but not the main indication for which the test was referred. Incidental findings are also unrelated to the disease screened but unintentionally found during the analysis. It is essential therefore to define which genetic variants will be reported back to the patient before germline testing is performed. Clear information about the intrinsic possibility of secondary and incidental findings should be included in the informed consent.

6 Conclusions

This thesis focused on increasing our knowledge about pediatric cancer predisposition. For this, different approaches were used, including germline and tumor WGS, epidemiological studies using the Swedish National Registries, literature reviews, molecular and genetic studies on patient's cells, and in-depth case presentations.

Initially, we developed a broad pediatric cancer predisposition research panel with 881 genes. We hope that this panel can be used for the discovery of known and novel childhood CPS genes and gene-disease associations, through pediatric cancer cohort analyses of germline WGS results. For instance, genome-wide or rare-variant association studies.

Further, we presented clinical, genetic and molecular information from patients with congenital disorders who developed pediatric malignancies. Specifically, we reported multiple ovarian tumors in a patient with PWS. We also described the first two reported patients with MFS who developed neuroblastoma and highlighted an early age at cancer diagnosis in published co-occurrences of MFS and cancer. Finally, we described a soft tissue sarcoma in a patient with LGMDR1. These reports highlight the importance of investigating cancer occurrences in patients with congenital disorders, to increase our understanding of their possible involvement in cancer development.

Based on the cases reported above, we decided to carry out register-based studies on the cancer risk in patients with PWS, muscular dystrophy and myotonic dystrophy. In PWS, we observed a high frequency of pediatric cancer. In muscular dystrophy, a high risk for pediatric astrocytomas and other gliomas was observed, while adults presented an increased incidence of pancreatic and nonthyroid endocrine cancer. Finally, individuals with myotonic dystrophy had a high risk of childhood brain tumors, while adults presented an increased overall cancer risk, explained by an overrepresentation of multiple malignancies. These results highlight the strengths of combining hypotheses derived from clinical case presentations with epidemiological studies.

Finally, we described a novel autosomal recessive intellectual disability syndrome with global developmental delay, immunodeficiency, facial dysmorphism, and leukemia, caused by variants in the *FLCN* gene. Analysis of the molecular mechanisms of disease led us to hypothesize that the *FLCN* p.G15S variant is

hypomorphic and results in TFE3/TFEB nuclear retention. This leads to upregulation of their target genes, affecting for instance the lysosomal biogenesis and autophagy pathways.

The studies presented in this thesis underscore the importance of tackling the study of pediatric CPS using multiple research tools. Initial hypotheses based on clinical presentations and candidate genetic variants must be confirmed with causation studies. Further, the conferred cancer risk can be estimated using register-based studies. Finding known and novel CPS and understanding their cancer risk spectrum is important for precision medicine; but it is also a complex task, best solved by incorporating information from different areas of biology, epidemiology, and medicine.

Overall, it was a pleasure to complete my PhD in such a growing field as the genetics of childhood cancer predisposition!

7 Points of perspective

“The problem [with genetic research] is, we’re just starting down this path, feeling our way in the dark. We have a small lantern in the form of a gene, but the lantern doesn’t penetrate more than a couple of hundred feet. We don’t know whether we’re going to encounter chasms, rock walls or mountain ranges along the way. We don’t even know how long the path is” — Francis S. Collins.

In the years to come, epidemiological and germline genetic studies on pediatric CPS will increase our knowledge of the incidence of cancer predisposition amongst oncology patients and in the general population. As the possibility of genetic testing expands in the clinic, thanks to an increase in the availability and affordability of sequencing tools, genetic testing will become more broadly used. Hopefully this will decrease the diagnostic gap amongst patients with pediatric CPS. Further, due to the growing number of genes associated with childhood cancer predisposition, gene panels will need to be periodically reviewed.

Better diagnostic pipelines for CPS allow for precision medicine, with benefits including access to available targeted treatments, prevention of treatment-related toxicity, surveillance initiation, and genetic counselling. As with rare diseases, however, there is an enormous need for better treatment options for patients with CPS. Increased understanding of cancer biology holds promise for the development of targeted treatments. Here, I believe that further advances in therapies based on monoclonal antibodies, small molecules and gene therapy will continue to be of great importance.

I anticipate that many undiscovered CPS are complex. Namely, caused by the interplay of multiple genes, each conferring a small increased cancer risk; or very rare, for instance, associated with congenital disorders. This, as CPS with a high penetrance and clear inheritance patterns have likely been described. Detecting novel CPS will therefore require large cohorts and smart pipelines, which account for the interaction of multiple variables in cancer development. Initial efforts in this direction include for instance rare-variant association studies [218], as in [219–223]. With the rapid progress of artificial intelligence, its use for the identification of enriched variants in pediatric patients with cancer is warranted. I expect that these studies will result in improved CPS diagnostic rates, as well as better stratification tools for cancer risk assessment, based on information on genetics, phenotypes, and family history of cancer.

Be it as it may, I am positive that the field of genetic diagnosis and treatment of rare diseases, including childhood CPS, will continue to grow rapidly in the years to come. Hopefully, these advances will translate into improved clinical outcomes for pediatric cancer patients in terms of treatment, prognosis, and surveillance.

8 Acknowledgements

Since I was in high school counting fruit flies, it was my life's goal to become a geneticist; a dream that grew bigger and bigger with time. Working at the Karolinska Institutet in Sweden was, however, way beyond my wildest fantasies. I was very fortunate to have so many people around to make this PhD possible. I would not be here without your encouragement, inspiration, and support.

My special thanks to:

My main supervisor **Ann Nordgren**, for your guidance during the duration of these studies. I have never met a person who is so curious, passionate, enthusiastic, and persistent about rare diseases and clinical research. You are brilliant, efficient, and always full of ideas. You look for ways to make other people grow, believe in themselves and shine. That makes you an excellent leader and a great supervisor. I am so thankful to have had the opportunity of working with you!

My co-supervisors, for you are the reason all projects have come to a successful end. **Fulya Taylan**, you are an amazing teacher, a great researcher, an extraordinary woman, and friend. When I grow up, I want to be like you. Thank you so much! **Jesper Eisfeldt**, you are an inexhaustible source of positive energy. Thank you for patiently trying to teach me to code –in Swedish– and for listening to my multiple tantrums about it. It is a treasure to have you as a supervisor and friend! **Bianca Tesi**, I appreciate you giving structure to my work, helping me become a better researcher, and teaching me about CPS. **David Gisselson**, thank you for joining our team and giving us input from a pathologist perspective.

My PhD mentor **Daniel Wilhelms**, for helping me plan, prioritize and solve all unexpected hurdles over the years. I am immensely grateful for your time! My previous mentors: **Lena Åslund**, **Ambrin Fatima** and **Neil Vásquez**. Thank you so much for believing in me, supporting me through the process of learning genetics, and sharing my love for science.

MMK administration, economy, and IT, especially **Ann-Britt**, **Susanne**, **Marietta** and **Jan-Erik**. Tack för att ni svarade på alla mina oändliga frågor. Ni gjorde min PhD-resa mycket smidigare!

All the **families** who participated in these studies. Thank you for your engagement!

Science is a collaborative process. All studies in this thesis are the result of the effort of multiple people. All **co-authors** and **collaborators**, thank you for your ideas, sharing of skills, and assistance. The **Nordic ChiCaP** group, for all joint efforts to study pediatric CPS. **Katja Ekholm**, tack för att du peppade mig igenom min PhD. Du är en fantastisk kollega och vän, kram! **Angélica**, compartir contigo es una recarga automática de energía. Tanto tu compromiso con los pacientes, como tu preocupación por los demás es admirable. **Dr. Fiona Poyer**, thank you for being an incredible colleague, friend, host and teacher. ¡Te quiero!

All colleagues at **BioClinicum J10:20**, who made this PhD a fun and fulfilling experience. **Jakob, Dominyka, Sailan, Isabel, Fulya** and **Marlene**, you were the greatest office companions. **Jesper, Kristin** and **Esmee**, having friends like you makes all the difference at work. I hope that we continue to plan adventures in the future. **Blaz**, I look up to you. You are a great scientist and friend –and probably the most organized person I know. Thank you for everything! **Nina**, tack så jättemycket för all hjälp på labbet. Du är en hjälte! Thanks to all colleagues at J10:20 over the years: **Alexandra, Anna, Anna Skarin, Anton, Behzad, Benedicte, Cecilia, Clara, Daniel, Emma, Fan, Fatemah, Giedre, Giorgio, Gleb, Hassan, Hillevi, Ida, John, José, Karl, Kicki, Larry, Mastoureh, Raquel, Richard, Sandra, Sofía, Teresita, and Zahra**; and at **SciLifeLab**: **Anders, Annick, Eva, Karl, Sebastián, and Vincent**, for adopting me in your lunchroom and group trips. You are great, guys!

Many people outside the lab also made this PhD possible:

Primero mis amigos, porque *“los amigos son la familia que escoges”*:

Team –Andi, Carlos, Diego, Herrón, Laura, Manu, Manu Pip, y Viole, poder contar con ustedes es inconmensurable; y que la mayoría esté en Europa, un sueño. ¡Los quiero muchísimo! **Verito**, ¿Qué puedo decir? Eres una hermosa persona y amiga. Aquí para lo que necesites, mi vecina favorita.

The wizard kings –Estefa, Endri, Santi y Ángela–, Dani, Roger y Sofia, Eugenia y Fernando. Gracias por estar conmigo tanto en los valles como en los picos. Ustedes se han comportado como una familia, y han hecho de Suecia mi segundo hogar ¡Los quiero hasta el cielo!

Uppsala social dance –Philip, Camilo, Max, Samuel, Paul, Emil, Chris & Sofia. **TKoeB** speciellt **Andrew, Åke, Johan, Lovisa, Sonja, och Ricarda**, tack för att ni gjorde alla dessa år roligare. Tack vare er kunde jag hålla en bra balans mellan

arbete, privatliv och mental hälsa. **Joni & Ola**, for keeping me sane during these years. You have been my most genuine confidants. Thank you for listening!

Gänghäng: Sam, Simon, Love, Matilda, Marcus, Malin, Linus, Joakim, Elvira, Henke, Cajsa, Björn, Gustav, Viktor, Linnea och, Diddis. Jag hade så roligt med er. Tack för att jag fick bli en del av gänget! **Carina, Steven, Stefan, Ida, Sanna, Sofia & Leon**, tack för alla minnen.

Edvard, Harriet, Sunne, John, Claire, och Benita, tack för att ni blev som min familj i Europa när jag flyttade till Sverige. Jag är jättetacksam för allt tålamod med att lära mig om kulturen och språket.

Y ahora mi familia:

Papá, me duele mucho que no puedas estar aquí para celebrar este logro conmigo. Te recuerdo y hoy te agradezco por todo tu esfuerzo para ayudarme a alcanzarlo. **Mamá**, gracias por siempre querer lo mejor para mí, por tu paciencia conmigo y tu amor. **Andre**, eres uno de los motores que me mantiene andando. ¡Te quiero mucho! **Norita y Nemo**, no tengo palabras para agradecer lo mucho que me han dado. En cada aspecto de mi personalidad hay una semillita sembrada por sus enseñanzas. ¡Gracias por tanto amor!

Las tías **Dorita, Lila, Norita, Nubia, Pili y Tis**, que me ayudaron a subir cada uno de los escalones necesarios para completar este logro. Su generosidad, cariño, y confianza en mis capacidades ha sido invaluable. **Alejo y Palis**, gracias por estar ahí para mí siempre, y ser una fuente de energía positiva y amor. Primos, **Alejo, Palis, Pipe, Linis, Ernes, Cami, Bibi, Juancho, Chago, Isa, Lesly, y Leti**, gracias por tanto cariño y enseñanzas. A todos, estoy muy feliz de que sean mi familia. ¡Tenerlos es un regalo de la vida!

Again, I am infinitely thankful to all the amazing people who were part of my PhD journey. I am also excited to see where the road takes me now, and I very much hope that most of you stay for the ride.

9 References

- 1 Liu Y, Li X. Darwin's Pangenesis and molecular medicine. *Trends Mol Med* 2012;**18**:506–8.
- 2 Darwin C, Murray J. *On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life*. London: : John Murray, Albemarle Street 1859. doi:10.5962/bhl.title.82303
- 3 Abbott S, Fairbanks DJ. Experiments on Plant Hybrids by Gregor Mendel. *Genetics* 2016;**204**:407–22.
- 4 Johann Friedrich Miescher. *Ueber die chemische Zusammensetzung der Eiterzellen*. 1871.
- 5 Garrod AE. The incidence of alkaptonuria: a study in chemical individuality. 1902. *Mol Med* 1996;**2**:274–82.
- 6 Morgan TH. The Application of the Conception of Pure Lines to Sex-Limited Inheritance and to Sexual Dimorphism. *Am Nat* 1911;**45**:65–78.
- 7 Morgan TH. Sex Limited Inheritance in *Drosophila*. *Science* (1979) 1910;**32**:120–2.
- 8 Sturtevant AH. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology* 1913;**14**:43–59.
- 9 Avery OT, MacLeod CM, McCarty M. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES. *Journal of Experimental Medicine* 1944;**79**:137–58.
- 10 Manchester KL. Historical Opinion: Erwin Chargaff and his 'rules' for the base composition of DNA: why did he fail to see the possibility of complementarity? *Trends Biochem Sci* 2008;**33**:65–70.
- 11 Chargaff Erwin, Lipshitz Rakoma, Green Charlotte, Hodes ME. The composition of the desoxyribonucleic acid of salmon sperm. *Journal of Biological Chemistry* 1951;**192**:223–30.
- 12 FRANKLIN RE, GOSLING RG. Molecular Configuration in Sodium Thymonucleate. *Nature* 1953;**171**:740–1.

- 13 WATSON JD, CRICK FHC. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 1953;**171**:737–8.
- 14 Sanger F, Tuppy H. The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates. *Biochemical Journal* 1951;**49**:463–81.
- 15 Sanger F. The terminal peptides of insulin. *Biochemical Journal* 1949;**45**:563–74.
- 16 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 1977;**74**:5463–7.
- 17 Pardue M Lou, Gall JG. MOLECULAR HYBRIDIZATION OF RADIOACTIVE DNA TO THE DNA OF CYTOLOGICAL PREPARATIONS. *Proceedings of the National Academy of Sciences* 1969;**64**:600–4.
- 18 JOHN HA, BIRNSTIEL ML, JONES KW. RNA-DNA Hybrids at the Cytological Level. *Nature* 1969;**223**:582–7.
- 19 Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harb Symp Quant Biol* 1986;**51**:263–73.
- 20 Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SBH, Hood LE. Fluorescence detection in automated DNA sequence analysis. *Nature* 1986;**321**:674–9.
- 21 DONISKELLER H. A genetic linkage map of the human genome. *Cell* 1987;**51**:319–37.
- 22 Antonarakis SE. History of the methodology of disease gene identification. *Am J Med Genet A* 2021;**185**:3266–75.
- 23 Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French

L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissole SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng J-F, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglu S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen H-C, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JGR, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin E V., Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran J V., Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AFA, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang S-P, Yeh R-F, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson M V., Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Patrinos A, Morgan MJ. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.

- 24 Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M,

Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Francesco V Di, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji R-R, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov G V., Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang ZY, Wang A, Wang X, Wang J, Wei M-H, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu SC, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratt S, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers Y-H, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander K V., Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang Y-H, Coyne M, Dahlke C, Mays AD, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The Sequence of the Human Genome. *Science* (1979) 2001;**291**:1304–51.

- 25 Milestones: Genomic Sequencing.
2021.[https://www.nature.com/collections/
genomic-sequencing-](https://www.nature.com/collections/genomic-sequencing-milestones)
milestones (accessed 12 Jan2024).
- 26 Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KHS. Viable offspring derived from fetal and adult mammalian cells. *Nature* 1997;**385**:810–3.
- 27 Campbell KHS, McWhir J, Ritchie WA, Wilmut I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* 1996;**380**:64–6.
- 28 Giani AM, Gallo GR, Gianfranceschi L, Formenti G. Long walk to genomics: History and current approaches to genome sequencing and assembly. *Comput Struct Biotechnol J* 2020;**18**:9–19.
- 29 Chen C, Wang J, Pan D, Wang X, Xu Y, Yan J, Wang L, Yang X, Yang M, Liu G. Applications of multi-omics analysis in human diseases. *MedComm (Beijing)* 2023;**4**. doi:10.1002/mco2.315
- 30 Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* 2012;**109**. doi:10.1073/pnas.1208507109
- 31 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (1979)* 2012;**337**:816–21.
- 32 Halbert CH. Equity in Genomic Medicine. *Annu Rev Genomics Hum Genet* 2022;**23**:613–25.
- 33 Patrushev LI, Kovalenko TF. Functions of noncoding sequences in mammalian genomes. *Biochemistry (Moscow)* 2014;**79**:1442–69.
- 34 Hombach S, Kretz M. Non-coding RNAs: Classification, Biology and Functioning. 2016. 3–17.
- 35 Eren K, Taktakoglu N, Pirim I. DNA Sequencing Methods: From Past to Present. *Eurasian J Med* 2023;**54**:S47–56.
- 36 Rhoads A, Au KF. PacBio Sequencing and its Applications. *Genomics Proteomics Bioinformatics* 2015;**13**:278–89.

- 37 Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* 2021;**39**:1348–65.
- 38 Zschocke J, Byers PH, Wilkie AOM. Mendelian inheritance revisited: dominance and recessiveness in medical genetics. *Nat Rev Genet* 2023;**24**:442–63.
- 39 Ferreira CR. The burden of rare diseases. *Am J Med Genet A* 2019;**179**:885–92.
- 40 Wang CM, Whiting AH, Rath A, Anido R, Ardigò D, Baynam G, Dawkins H, Hamosh A, Le Cam Y, Malherbe H, Molster CM, Monaco L, Padilla CD, Pariser AR, Robinson PN, Rodwell C, Schaefer F, Weber S, Macchia F. Operational description of rare diseases: a reference to improve the recognition and visibility of rare diseases. *Orphanet J Rare Dis* 2024;**19**:334.
- 41 Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y, Rath A. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics* 2019 28:2 2019;**28**:165–73.
- 42 Haendel M, Vasilevsky N, Unni D, Bologna C, Harris N, Rehm H, Hamosh A, Baynam G, Groza T, McMurry J, Dawkins H, Rath A, Thaxton C, Bocci G, Joachimiak MP, Köhler S, Robinson PN, Mungall C, Oprea TI. How many rare diseases are there? *Nat Rev Drug Discov* 2020;**19**:77–8.
- 43 Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y, Rath A. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics* 2020;**28**:165–73.
- 44 O'Connor DJ, Sheean ME, Hofer MP, Tsigkos S, Mariz S, Fregonese L, Larsson K, Hivert V, Westermarck K, Naumann-Winter F, Stoyanova-Beninska V, Barišić I, Capovilla G, Magrelli A, Sepodes B. Defining orphan conditions in the context of the European orphan regulation: challenges and evolution. *Nat Rev Drug Discov* 2019;**18**:479–80.
- 45 Casali PG, Trama A. Rationale of the rare cancer list: a consensus paper from the Joint Action on Rare Cancers (JARC) of the European Union (EU). *ESMO Open* 2020;**5**:e000666.

- 46 Bick D, Jones M, Taylor SL, Taft RJ, Belmont J. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet* 2019;**56**:783–91.
- 47 The Lancet Global Health. The landscape for rare diseases in 2024. *Lancet Glob Health* 2024;**12**:e341.
- 48 Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y, Rath A. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics* 2020;**28**:165–73.
- 49 Hanahan D, Weinberg RA. The Hallmarks of Cancer. *Cell* 2000;**100**:57–70.
- 50 Kontomanolis EN, Koutras A, Syllaios A, Schizas D, Mastoraki A, Garmpis N, Diakosavvas M, Angelou K, Tsatsaris G, Pagkalos A, Ntounis T, Fasoulakis Z. Role of Oncogenes and Tumor-suppressor Genes in Carcinogenesis: A Review. *Anticancer Res* 2020;**40**:6009–15.
- 51 Arem H, Loftfield E. Cancer Epidemiology: A Survey of Modifiable Risk Factors for Prevention and Survivorship. *Am J Lifestyle Med* 2018;**12**:200.
- 52 Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. 2020.<https://gco.iarc.fr/today> (accessed 29 Aug2022).
- 53 Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, Davies H, Stratton MR, Campbell PJ. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* 2017;**171**:1029–1041.e21.
- 54 Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* 2011;**144**:646–74.
- 55 Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* 2022;**12**:31–46.
- 56 Ladomery M. Aberrant alternative splicing is another hallmark of cancer. *Int J Cell Biol* 2013;**2013**:463786.
- 57 Swanton C, Bernard E, Abbosh C, André F, Auwerx J, Balmain A, Bar-Sagi D, Bernards R, Bullman S, DeGregori J, Elliott C, Erez A, Evan G, Febbraio MA,

- Hidalgo A, Jamal-Hanjani M, Joyce JA, Kaiser M, Lamia K, Locasale JW, Loi S, Malanchi I, Merad M, Musgrave K, Patel KJ, Quezada S, Wargo JA, Weeraratna A, White E, Winkler F, Wood JN, Vousden KH, Hanahan D. Embracing cancer complexity: Hallmarks of systemic disease. *Cell* 2024;**187**:1589–616.
- 58 Swedish Childhood Cancer Registry Annual Report, Barncancerregistret, 2022. Available from <https://sbcr.se/index.php/rapporter/> (accessed 17 Sep2024).
 - 59 Swedish Childhood Cancer Registry public statistics. Available at <https://statistik.incanet.se/sbcr/> (accessed 17 Sep2024).
 - 60 Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, Johann PD, Balasubramanian GP, Segura-Wang M, Brabetz S, Bender S, Hutter B, Sturm D, Pfaff E, Hübschmann D, Zipprich G, Heinold M, Eils J, Lawerenz C, Erkek S, Lambo S, Waszak S, Blattmann C, Borkhardt A, Kuhlen M, Eggert A, Fulda S, Gessler M, Wegert J, Kappler R, Baumhoer D, Burdach S, Kirschner-Schwabe R, Kontny U, Kulozik AE, Lohmann D, Hettmer S, Eckert C, Bielack S, Nathrath M, Niemeyer C, Richter GH, Schulte J, Siebert R, Westermann F, Molenaar JJ, Vassal G, Witt H, Zapatka M, Burkhardt B, Kratz CP, Witt O, Tilburg CMV, Kramm CM, Fleischhack G, Dirksen U, Rutkowski S, Frühwald M, Hoff K Von, Wolf S, Klingebiel T, Koscielniak E, Landgraf P, Koster J, Resnick AC, Zhang J, Liu Y, Zhou X, Waanders AJ, Zwijnenburg DA, Raman P, Brors B, Weber UD, Northcott PA, Pajtler KW, Kool M, Piro RM, Korbel JO, Schlesner M, Eils R, Jones DTW, Lichter P, Chavez L, Pfister SM. The landscape of genomic alterations across childhood cancers. *Nature* 2018;**555**:321–7.
 - 61 Kentsis A. Why do young people get cancer? *Pediatr Blood Cancer* 2020;**67**. doi:10.1002/pbc.28335
 - 62 Rahman N. Realizing the promise of cancer predisposition genes. *Nature* 2014 505:7483 2014;**505**:302–8.
 - 63 Jahn A, Rump A, Widmann TJ, Heining C, Horak P, Hutter B, Paramasivam N, Uhrig S, Gieldon L, Drukewitz S, Kübler A, Bermudez M, Hackmann K, Pormann J, Wagner J, Arlt M, Franke M, Fischer J, Kowalzyk Z, William D, Weth V, Oster S, Fröhlich M, Hüllein J, Valle González C, Kreutzfeldt S, Mock A, Heilig CE, Lipka DB, Möhrmann L, Hanf D, Oleś M, Teleanu V, Allgäuer M, Ruhnke L, Kutz O, Knurr A, Laßmann A, Endris V, Neumann O, Penzel R, Beck K, Richter D, Winter U, Wolf S, Pfütze K, Georg C, Meißburger B, Buchhalter I, Augustin M, Aulitzky

- WE, Hohenberger P, Kroiss M, Schirmacher P, Schlenk RF, Keilholz U, Klauschen F, Folprecht G, Bauer S, Siveke JT, Brandts CH, Kindler T, Boerries M, Illert AL, von Bubnoff N, Jost PJ, Metzeler KH, Bitzer M, Schulze-Osthoff K, von Kalle C, Brors B, Stenzinger A, Weichert W, Hübschmann D, Fröhling S, Glimm H, Schröck E, Klink B. Comprehensive cancer predisposition testing within the prospective MASTER trial identifies hereditary cancer patients and supports treatment decisions for rare cancers. *Annals of Oncology* 2022;**33**:1186–99.
- 64 Chavarri-Guerra Y, Slavin TP, Longoria-Lozano O, Weitzel JN. Genetic Cancer Predisposition Syndromes among Older Adults. *J Geriatr Oncol* 2020;**11**:1054.
- 65 Knudson AG. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences* 1971;**68**:820–3.
- 66 Kratz CP. Re-envisioning genetic predisposition to childhood and adolescent cancers. *Nat Rev Cancer* Published Online First: 3 December 2024. doi:10.1038/s41568-024-00775-7
- 67 Jongmans MCJ, Loeffen JLCM, Waanders E, Hoogerbrugge PM, Ligtenberg MJL, Kuiper RP, Hoogerbrugge N. Recognition of genetic predisposition in pediatric cancer patients: An easy-to-use selection tool. *Eur J Med Genet* 2016;**59**:116–25.
- 68 Hopman SMJ, Merks JHM, de Borgie CAJM, Aalfs CM, Biesecker LG, Cole T, Eng C, Legius E, Maher ER, van Noesel MM, Verloes A, Viskochil DH, Wagner A, Weksberg R, Caron HN, Hennekam RCM. The development of a clinical screening instrument for tumour predisposition syndromes in childhood cancer patients. *Eur J Cancer* 2013;**49**:3247–54.
- 69 Ripperger T, Bielack SS, Borkhardt A, Brecht IB, Burkhardt B, Calaminus G, Debatin K-M, Deubzer H, Dirksen U, Eckert C, Eggert A, Erlacher M, Fleischhack G, Frühwald MC, Gnekow A, Goehring G, Graf N, Hanenberg H, Hauer J, Hero B, Hettmer S, von Hoff K, Horstmann M, Hoyer J, Illig T, Kaatsch P, Kappler R, Kerl K, Klingebiel T, Kontny U, Kordes U, Körholz D, Koscielniak E, Kramm CM, Kühlen M, Kulozik AE, Lamottke B, Leuschner I, Lohmann DR, Meinhardt A, Metzler M, Meyer LH, Moser O, Nathrath M, Niemeyer CM, Nustede R, Pajtler KW, Paret C, Rasche M, Reinhardt D, Rieß O, Russo A, Rutkowski S, Schlegelberger B, Schneider D, Schneppenheim R, Schrappe M,

- Schroeder C, von Schweinitz D, Simon T, Sparber-Sauer M, Spix C, Stanulla M, Steinemann D, Strahm B, Temming P, Thomay K, von Bueren AO, Vorwerk P, Witt O, Wlodarski M, Wössmann W, Zenker M, Zimmermann S, Pfister SM, Kratz CP. Childhood cancer predisposition syndromes—A concise review and recommendations by the Cancer Predisposition Working Group of the Society for Pediatric Oncology and Hematology. *Am J Med Genet A* 2017;**173**:1017–37.
- 70 Goudie C, Coltin H, Witkowski L, Mourad S, Malkin D, Foulkes WD. The McGill Interactive Pediatric OncoGenetic Guidelines: An approach to identifying pediatric oncology patients most likely to benefit from a genetic evaluation. *Pediatr Blood Cancer* 2017;**64**. doi:10.1002/PBC.26441
- 71 Goudie C, Witkowski L, Cullinan N, Reichman L, Schiller I, Tachdjian M, Armstrong L, Blood KA, Brossard J, Brunga L, Cacciotti C, Caswell K, Cellot S, Clark ME, Clinton C, Coltin H, Felton K, Fernandez C v., Fleming AJ, Fuentes-Bolanos N, Gibson P, Grant R, Hammad R, Harrison LW, Irwin MS, Johnston DL, Kane S, Lafay-Cousin L, Lara-Corrales I, Larouche V, Mathews N, Meyn MS, Michaeli O, Perrier R, Pike M, Punnett A, Ramaswamy V, Say J, Somers G, Tabori U, Thibodeau ML, Toupin AK, Tucker KM, van Engelen K, Vairy S, Waespe N, Warby M, Wasserman JD, Whitlock JA, Sinnott D, Jabado N, Nathan PC, Shlien A, Kamihara J, Deyell RJ, Ziegler DS, Nichols KE, Dendukuri N, Malkin D, Villani A, Foulkes WD. Performance of the McGill Interactive Pediatric OncoGenetic Guidelines for Identifying Cancer Predisposition Syndromes. *JAMA Oncol* 2021;**7**:1806–14.
- 72 Schneider K, Zelle K, Nichols KE, Garber J. Li-Fraumeni Syndrome. *GeneReviews*(®) Published Online First: 21 November 2019.<https://www.ncbi.nlm.nih.gov/books/NBK1311/> (accessed 16 Sep2022).
- 73 Wimmer K, Kratz CP. Constitutional mismatch repair-deficiency syndrome. *Haematologica* 2010;**95**:699.
- 74 Hays L, Frohnmayr D, Frohnmayr L, Larsen K, Owen J. Fanconi Anemia. *Fanconi Anemia Research Fund, Inc* 2021;**1**–391.
- 75 Schultz KAP, Stewart DR, Kamihara J, Bauer AJ, Merideth MA, Stratton P, Huryn LA, Harris AK, Doros L, Field A, Carr AG, Dehner LP, Messinger Y, Hill DA. DICER1 Tumor Predisposition. *GeneReviews*® Published Online First: 30 April

2020.<https://www.ncbi.nlm.nih.gov/books/NBK196157/> (accessed 22 Dec2022).

- 76 Lohmann DR, Gallie BL. Retinoblastoma. *GeneReviews*® Published Online First: 21 November 2018.<https://www.ncbi.nlm.nih.gov/books/NBK1452/> (accessed 16 Sep2022).
- 77 Kraemer KH, DiGiovanna JJ, Tamura D. Xeroderma Pigmentosum. *GeneReviews*® Published Online First: 24 March 2022.<https://www.ncbi.nlm.nih.gov/books/NBK1397/> (accessed 16 Sep2022).
- 78 Uusitalo E, Rantanen M, Kallionpää RA, Pöyhönen M, Leppävirta J, Ylä-Outinen H, Riccardi VM, Pukkala E, Pitkäniemi J, Peltonen S, Peltonen J. Distinctive cancer associations in patients with neurofibromatosis type 1. *Journal of Clinical Oncology* 2016;**34**:1978–86.
- 79 Friedman JM. Neurofibromatosis 1. *GeneReviews* 2022;[no pagination].
- 80 Kratz CP, Rapisuwon S, Reed H, Hasle H, Rosenberg PS. Cancer in Noonan, Costello, Cardiofaciocutaneous and LEOPARD Syndromes. doi:10.1002/ajmg.c.30300
- 81 Smpokou P, Zand DJ, Rosenbaum KN, Summar ML. Malignancy in Noonan syndrome and related disorders. *Clin Genet* 2015;**88**:516–22.
- 82 Roberts AE. Noonan Syndrome. *GeneReviews*® Published Online First: 17 February 2022.<https://www.ncbi.nlm.nih.gov/books/NBK1124/> (accessed 16 Sep2022).
- 83 Lipska-Ziętkiewicz BS. WT1 Disorder. *GeneReviews*® Published Online First: 29 April 2021.<https://www.ncbi.nlm.nih.gov/books/NBK556455/> (accessed 16 Sep2022).
- 84 Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, Hedges D, Ma X, Zhou X, Yergeau DA, Wilkinson MR, Vadodaria B, Chen X, McGee RB, Hines-Dowell S, Nuccio R, Quinn E, Shurtleff SA, Rusch M, Patel A, Becksfort JB, Wang S, Weaver MS, Ding L, Mardis ER, Wilson RK, Gajjar A, Ellison DW, Pappo AS, Pui C-H, Nichols KE, Downing JR. Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med* 2015;**373**:2336–46.

- 85 Parsons DW, Roy A, Yang Y, Wang T, Scollon S, Bergstrom K, Kerstein RA, Gutierrez S, Petersen AK, Bavle A, Lin FY, López-Terrada DH, Monzon FA, Hicks MJ, Eldin KW, Quintanilla NM, Adesina AM, Mohila CA, Whitehead W, Jea A, Vasudevan SA, Nuchtern JG, Ramamurthy U, McGuire AL, Hilsenbeck SG, Reid JG, Muzny DM, Wheeler DA, Berg SL, Chintagumpala MM, Eng CM, Gibbs RA, Plon SE. Diagnostic Yield of Clinical Tumor and Germline Whole-Exome Sequencing for Children With Solid Tumors. *JAMA Oncol* 2016;**2**:616–24.
- 86 Oberg JA, Glade Bender JL, Sulis ML, Pendrick D, Sireci AN, Hsiao SJ, Turk AT, Dela Cruz FS, Hibshoosh H, Remotti H, Zylber RJ, Pang J, Diolaiti D, Koval C, Andrews SJ, Garvin JH, Yamashiro DJ, Chung WK, Emerson SG, Nagy PL, Mansukhani MM, Kung AL. Implementation of next generation sequencing into pediatric hematology-oncology practice: moving beyond actionable alterations. *Genome Med* 2016;**8**. doi:10.1186/S13073-016-0389-6
- 87 Wong M, Mayoh C, Lau LMS, Khuong-Quang DA, Pinese M, Kumar A, Barahona P, Wilkie EE, Sullivan P, Bowen-James R, Syed M, Martincorena I, Abascal F, Sherstyuk A, Bolanos NA, Baber J, Priestley P, Dolman MEM, Fleuren EDG, Gauthier ME, Mould EVA, Gayevskiy V, Gifford AJ, Grebert-Wade D, Strong PA, Manouvrier E, Warby M, Thomas DM, Kirk J, Tucker K, O'Brien T, Alvaro F, McCowage GB, Dalla-Pozza L, Gottardo NG, Tapp H, Wood P, Khaw SL, Hansford JR, Moore AS, Norris MD, Trahair TN, Lock RB, Tyrrell V, Haber M, Marshall GM, Ziegler DS, Ekert PG, Cowley MJ. Whole genome, transcriptome and methylome profiling enhances actionable target discovery in high-risk pediatric cancer. *Nat Med* 2020;**26**:1742–53.
- 88 Byrjalsen A, Hansen TVO, Stoltze UK, Mehrjouy MM, Barnkob NM, Hjalgrim LL, Mathiasen R, Lautrup CK, Gregersen PA, Hasle H, Wehner PS, Tuckuviene R, Sackett PW, Laspiur AO, Rossing M, Marvig RL, Tommerup N, Olsen TE, Scheie D, Gupta R, Gerdes AM, Schmiegelow K, Wadt K. Nationwide germline whole genome sequencing of 198 consecutive pediatric cancer patients reveals a high incidence of cancer prone syndromes. *PLoS Genet* 2020;**16**:e1009231.
- 89 Wagener R, Taeubner J, Walter C, Yasin L, Alzoubi D, Bartenhagen C, Attarbaschi A, Classen CF, Kontny U, Hauer J, Fischer U, Dugas M, Kuhlen M, Borkhardt A, Brozou T. Comprehensive germline-genomic and clinical profiling in 160 unselected children and adolescents with cancer. *European Journal of Human Genetics* 2021 29:8 2021;**29**:1301–11.

- 90 Fiala EM, Jayakumaran G, Mauguen A, Kennedy JA, Bouvier N, Kemel Y, Fleischut MH, Maio A, Salo-Mullen EE, Sheehan M, Arnold AG, Latham A, Carlo MI, Cadoo K, Murkherjee S, Slotkin EK, Trippett T, Glade Bender J, Meyers PA, Wexler L, Dela Cruz FS, Cheung NK, Basu E, Kentsis A, Ortiz M, Francis JH, Dunkel IJ, Khakoo Y, Gilheeney S, Farouk Sait S, Forlenza CJ, Sulis M, Karajannis M, Modak S, Gerstle JT, Heaton TE, Roberts S, Yang C, Jairam S, Vijai J, Topka S, Friedman DN, Stadler ZK, Robson M, Berger MF, Schultz N, Ladanyi M, O'Reilly RJ, Abramson DH, Ceyhan-Birsoy O, Zhang L, Mandelker D, Shukla NN, Kung AL, Offit K, Zehir A, Walsh MF. Prospective pan-cancer germline testing using MSK-IMPACT informs clinical translation in 751 patients with pediatric solid tumors. *Nature Cancer* 2021 2:3 2021;2:357–65.
- 91 Tesi B, Robinson KL, Abel F, Díaz de Ståhl T, Orrsjö S, Poluha A, Hellberg M, Wessman S, Samuelsson S, Frisk T, Vogt H, Henning K, Sabel M, Ek T, Pal N, Nyman P, Giraud G, Wille J, Pronk CJ, Norén-Nyström U, Borssén M, Fili M, Stålhammar G, Herold N, Tettamanti G, Maya-Gonzalez C, Arvidsson L, Rosén A, Ekholm K, Kuchinskaya E, Hallbeck A-L, Nordling M, Palmebäck P, Kogner P, Smoler GK, Lähteenmäki P, Fransson S, Martinsson T, Shamik A, Mertens F, Rosenquist R, Wirta V, Tham E, Grillner P, Sandgren J, Ljungman G, Gisselsson D, Taylan F, Nordgren A. Diagnostic yield and clinical impact of germline sequencing in children with CNS and extracranial solid tumors—a nationwide, prospective Swedish study. *The Lancet Regional Health - Europe* 2024;:100881.
- 92 Newman S, Nakitandwe J, Kesserwan CA, Azzato EM, Wheeler DA, Rusch M, Shurtleff S, Hedges DJ, Hamilton K V., Foy SG, Edmonson MN, Thrasher A, Bahrami A, Orr BA, Klco JM, Gu J, Harrison LW, Wang L, Clay MR, Ouma A, Silkov A, Liu Y, Zhang Z, Liu Y, Brady SW, Zhou X, Chang TC, Pande M, Davis E, Becksfort J, Patel A, Wilkinson MR, Rahbarinia D, Kubal M, Maciaszek JL, Pastor V, Knight J, Gout AM, Wang J, Gu Z, Mullighan CG, McGee RB, Quinn EA, Nuccio R, Mostafavi R, Gerhardt EL, Taylor LM, Valdez JM, Hines-Dowell SJ, Pappo AS, Robinson G, Johnson LM, Pui CH, Ellison DW, Downing JR, Zhang J, Nichols KE. Genomes for Kids: The Scope of Pathogenic Mutations in Pediatric Cancer Revealed by Comprehensive DNA and RNA Sequencing. *Cancer Discov* 2021;11:3008–27.
- 93 Hasle H, Haunstrup Clemmensen I, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. *Lancet* 2000;355:165–9.

- 94 Turner JT, Brzezinski J, Dome JS. Wilms Tumor Predisposition. *GeneReviews*®
Published Online First: 24 March 2022.<https://www.ncbi.nlm.nih.gov/books/NBK1294/> (accessed 15 Sep2022).
- 95 Mahamdallie S, Yost S, Poyastro-Pearson E, Holt E, Zachariou A, Seal S, Elliott A, Clarke M, Warren-Perry M, Hanks S, Anderson J, Bomken S, Cole T, Farah R, Furtwaengler R, Glaser A, Grundy R, Hayden J, Lewis S, Millot F, Nicholson J, Ronghe M, Skeen J, Williams D, Yeomanson D, Ruark E, Rahman N. Identification of new Wilms tumour predisposition genes: an exome sequencing study. *Lancet Child Adolesc Health* 2019;**3**:322–31.
- 96 Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader-Willi syndrome. Published Online First: 2012. doi:10.1038/gim.Ob013e31822bead0
- 97 Davies HD, Leusink GL, McConnell A, Deyell M, Cassidy SB, Fick GH, Coppes MJ. Myeloid leukemia in Prader-Willi syndrome. *J Pediatr* 2003;**142**:174–8.
- 98 Patja K, Sund R, Kaski M, Pukkala E. Cancer incidence among persons with Prader-Willi syndrome in Finland. *International Journal on Disability and Human Development* 2008;**7**:69–72.
- 99 Coppes MJ, Rose S, Cassidy S, McConnell A, Deyell M, Davies D. The Risk of Testicular Cancer in Males with Prader-Willi Syndrome (PWS) Registered with the US PWS Association. *Pediatr Res* 1999;**45**. doi:10.1203/00006450-199904020-00723
- 100 Reik W, Walter J. Imprinting mechanisms in mammals. *Curr Opin Genet Dev* 1998;**8**. doi:10.1016/S0959-437X(98)80136-6
- 101 Eggermann T, Perez De Nanclares G, Maher ER, Karen Temple I, Tümer Z, Monk D, Mackay DJG, Grønskov K, Riccio A. Imprinting disorders: a group of congenital disorders with overlapping patterns of molecular changes affecting imprinted loci. Published Online First: 2011. doi:10.1186/s13148-015-0143-8
- 102 Lobo I. Genomic Imprinting and Patterns of Disease Inheritance. *Nature Education* 2008;**1(1)**:66.

- 103 Moulton T, Crenshaw T, Hao Y, Moosikasuwan J, Lin N, Dembitzer F, Hensle T, Weiss L, McMorro L, Loew T, Kraus W, Gerald W, Tycko B. Epigenetic lesions at the H19 locus in Wilms' tumour patients. *Nat Genet* 1994;**7**:440–7.
- 104 Steenman MJC, Rainier S, Dobry CJ, Grundy P, Horon IL, Feinberg AP. Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nat Genet* 1994;**7**:433–9.
- 105 Eldar–Geva T, Gross–Tsur V, Hirsch HJ, Altarescu G, Segal R, Zeligson S, Golomb E, Epsztejn–Litman S, Eiges R. Incomplete methylation of a germ cell tumor (Seminoma) in a Prader–Willi male. *Mol Genet Genomic Med* 2018;**6**. doi:10.1002/mgg3.448
- 106 Socialstyrelsen – Sällsynta hälsotillstånd – csdsamverkan. <https://csdsamverkan.se/diagnosdatabaser/socialstyrelsen---sallsynta-halsotillstand?url=-233391368%2Fkunksapsstod-och-regler%2Fomraden%2Fsallsynta-halsotillstand%2Fmarfans-syndrom%2F&sv.url=12.5ad2782217c981018891393d#main-content> (accessed 21 Dec2022).
- 107 Milewicz DM, Braverman AC, De Backer J, Morris SA, Boileau C, Maumenee IH, Jondeau G, Evangelista A, Pyeritz RE, De Backer J. Marfan syndrome. *Nature Reviews Disease Primers* 2021 **7**:1 2021;**7**:1–24.
- 108 Dietz H. FBN1–Related Marfan Syndrome. *GeneReviews*® Published Online First: 17 February 2022.<https://www.ncbi.nlm.nih.gov/books/NBK1335/> (accessed 21 Dec2022).
- 109 LY S, DR K, E E. Fibrillin, a new 350–kD glycoprotein, is a component of extracellular microfibrils. *J Cell Biol* 1986;**103**:2499–509.
- 110 Pyeritz RE. Marfan Syndrome and Related Disorders. *Emery and Rimoin's Principles and Practice of Medical Genetics* 2013;**1**–52.
- 111 Hsu C–W, Wang J–C, Liao W–I, Chien W–C, Chung C–H, Tsao C–H, Wu Y–F, Liao M–T, Tsai S–H. Association between malignancies and Marfan syndrome: a population–based, nested case–control study in Taiwan. *BMJ Open* 2017;**7**:e017243.
- 112 Gadalla SM, Lund M, Pfeiffer RM, Gørtz S, Mueller CM, Moxley RT, Kristinsson SY, Björkholm M, Shebl FM, Hilbert JE, Landgren O, Wohlfahrt J, Melbye M,

- Greene MH. Cancer Risk Among Patients With Myotonic Muscular Dystrophy. *JAMA* 2011;**306**. doi:10.1001/jama.2011.1796
- 113 Win AK, Perattur PG, Pulido JS, Pulido CM, Lindor NM. Increased Cancer Risks in Myotonic Dystrophy. *Mayo Clin Proc* 2012;**87**:130–5.
 - 114 Abbott D, Johnson NE, Cannon–Albright LA. A population–based survey of risk for cancer in individuals diagnosed with myotonic dystrophy. *Muscle Nerve* 2016;**54**:783–5.
 - 115 Wang Y, Pfeiffer RM, Alsaggaf R, Meeraus W, Gage JC, Anderson LA, Bremer RC, Nikolenko N, Lochmuller H, Greene MH, Gadalla SM. Risk of skin cancer among patients with myotonic dystrophy type 1 based on primary care physician data from the United Kingdom Clinical Practice Research Datalink. *Int J Cancer* 2018;**142**:1174–81.
 - 116 Alsaggaf R, St. George DMM, Zhan M, Pfeiffer RM, Wang Y, Wagner KR, Greene MH, Amr S, Gadalla SM. Cancer Risk in Myotonic Dystrophy Type I: Evidence of a Role for Disease Severity. *JNCI Cancer Spectr* 2018;**2**. doi:10.1093/jncics/pky052
 - 117 Fernández–Torrón R, García–Puga M, Emparanza J–I, Maneiro M, Cobo A–M, Poza J–J, Espinal J–B, Zulaica M, Ruiz I, Martorell L, Otaegui D, Matheu A, López de Munain A. Cancer risk in DM1 is sex–related and linked to miRNA–200/141 downregulation. *Neurology* 2016;**87**:1250–7.
 - 118 Chamberlain JS, Metzger J, Reyes M, Townsend D, Faulkner JA. Dystrophin–deficient *mdx* mice display a reduced life span and are susceptible to spontaneous rhabdomyosarcoma. *The FASEB Journal* 2007;**21**:2195–204.
 - 119 Fernandez K, Serinagaoglu Y, Hammond S, Martin LT, Martin PT. Mice Lacking Dystrophin or α Sarcoglycan Spontaneously Develop Embryonal Rhabdomyosarcoma with Cancer–Associated p53 Mutations and Alternatively Spliced or Mutant Mdm2 Transcripts. *Am J Pathol* 2010;**176**:416–34.
 - 120 Sher RB, Cox GA, Mills KD, Sundberg JP. Rhabdomyosarcomas in Aging A/J Mice. *PLoS One* 2011;**6**:e23498.
 - 121 Schmidt WM, Uddin MH, Dysek S, Moser–Thier K, Pirker C, Höger H, Ambros IM, Ambros PF, Berger W, Bittner RE. DNA Damage, Somatic Aneuploidy, and

Malignant Sarcoma Susceptibility in Muscular Dystrophies. *PLoS Genet* 2011;**7**:e1002042.

- 122 Sattler EC, Reithmair M, Steinlein OK. Kidney cancer characteristics and genotype-phenotype-correlations in Birt-Hogg-Dubé syndrome. *PLoS One* 2018;**13**. doi:10.1371/JOURNAL.PONE.0209504
- 123 Furuya M, Yao M, Tanaka R, Nagashima Y, Kuroda N, Hasumi H, Baba M, Matsushima J, Nomura F, Nakatani Y. Genetic, epidemiologic and clinicopathologic studies of Japanese Asian patients with Birt-Hogg-Dubé syndrome. *Clin Genet* 2016;**90**:403–12.
- 124 Schmidt LS, Linehan WM. FLCN: The causative gene for Birt-Hogg-Dubé syndrome. *Gene*. 2018;**640**:28–42.
- 125 Hasumi H, Furuya M, Tatsuno K, Yamamoto S, Baba M, Hasumi Y, Isono Y, Suzuki K, Jikuya R, Otake S, Muraoka K, Osaka K, Hayashi N, Makiyama K, Miyoshi Y, Kondo K, Nakaigawa N, Kawahara T, Izumi K, Teranishi J, Yumura Y, Uemura H, Nagashima Y, Metwalli AR, Schmidt LS, Aburatani H, Linehan WM, Yao M. BHD-associated kidney cancer exhibits unique molecular characteristics and a wide variety of variants in chromatin remodeling genes. *Hum Mol Genet* 2018;**27**:2712.
- 126 Nickerson ML, Warren MB, Toro JR, Matrosova V, Glenn G, Turner ML, Duray P, Merino M, Choyke P, Pavlovich CP, Sharma N, Walther M, Munroe D, Hill R, Maher E, Greenberg C, Lerman MI, Linehan WM, Zbar B, Schmidt LS. Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dubé syndrome. *Cancer Cell* 2002;**2**:157–64.
- 127 Palmirotta R, Savonarola A, Ludovici G, Donati P, Cavaliere F, DE Marchis ML, Ferroni P, Guadagni F. Association between Birt Hogg Dube syndrome and cancer predisposition. *Anticancer Res* 2010;**30**:751–7.
- 128 Ramirez Reyes JM, Cuesta R, Pause A. Folliculin: A Regulator of Transcription Through AMPK and mTOR Signaling Pathways. *Front Cell Dev Biol* 2021;**9**. doi:10.3389/fcell.2021.667311
- 129 Tsaousis GN, Papadopoulou E, Apessos A, Agiannitopoulos K, Pepe G, Kampouri S, Diamantopoulos N, Floros T, Iosifidou R, Katopodi O, Koumarianou A, Markopoulos C, Papazisis K, Venizelos V, Xanthakis I,

- Xepapadakis G, Banu E, Eniu DT, Negru S, Stanculeanu DL, Ungureanu A, Ozmen V, Tansan S, Tekinel M, Yalcin S, Nasioulas G. Analysis of hereditary cancer syndromes by using a panel of genes: novel and multiple pathogenic mutations. doi:10.1186/s12885-019-5756-4
- 130 Ross JP, Dion PA, Rouleau GA. Exome sequencing in genetic disease: recent advances and considerations. *FlOOORes* 2020;**9**:336.
 - 131 Hiemenz MC, Ostrow DG, Busse TM, Buckley J, Maglinte DT, Bootwalla M, Done J, Ji J, Raca G, Ryutov A, Xu X, Zhen CJ, Conroy JM, Hazard FK, Deignan JL, Rogers BB, Treece AL, Parham DM, Gai X, Judkins AR, Triche TJ, Biegel JA. OncoKids: A Comprehensive Next-Generation Sequencing Panel for Pediatric Malignancies. *J Mol Diagn* 2018;**20**:765–76.
 - 132 Byrjalsen A, Diets IJ, Bakhuizen J, Hansen T van O, Schmiegelow K, Gerdes AM, Stoltze U, Kuiper RP, Merks JHM, Wadt K, Jongmans M. Selection criteria for assembling a pediatric cancer predisposition syndrome gene panel. *Fam Cancer* 2021;**20**:279.
 - 133 Glaire MA, Brown M, Church DN, Tomlinson I. Cancer predisposition syndromes: lessons for truly precision medicine. *J Pathol* 2017;**241**:226–35.
 - 134 Hasin Y, Seldin M, Lusi A. Multi-omics approaches to disease. *Genome Biology* 2017 18:1 2017;**18**:1–15.
 - 135 Tidball AM. Disease in a Dish: Cellular Models to Understand Human Conditions. In: *Cellular and Animal Models in Human Genomics Research*. Elsevier 2019. 19–47.
 - 136 Reza Khorramizadeh M, Saadat F. Animal models for human disease. In: *Animal Biotechnology*. Elsevier 2020. 153–71.
 - 137 Li Z, Zheng W, Wang H, Cheng Y, Fang Y, Wu F, Sun G, Sun G, Lv C, Hui B. Application of Animal Models in Cancer Research: Recent Progress and Future Prospects. *Cancer Manag Res* 2021;**Volume 13**:2455–75.
 - 138 Richter M, Piwocka O, Musielak M, Piotrowski I, Suchorska WM, Trzeciak T. From Donor to the Lab: A Fascinating Journey of Primary Cell Lines. *Front Cell Dev Biol* 2021;**9**:1869.
 - 139 Hason, Bartůněk. Zebrafish Models of Cancer—New Insights on Modeling Human Cancer in a Non-Mammalian Vertebrate. *Genes (Basel)* 2019;**10**:935.

- 140 Yamamura R, Ooshio T, Sonoshita M. Tiny *Drosophila* makes giant strides in cancer research. *Cancer Sci* 2021;**112**:505–14.
- 141 Cheon D–J, Orsulic S. Mouse Models of Cancer. *Annual Review of Pathology: Mechanisms of Disease* 2011;**6**:95–119.
- 142 Kobar K, Collett K, Prykhozhij S v., Berman JN. Zebrafish Cancer Predisposition Models. *Front Cell Dev Biol* 2021;**9**. doi:10.3389/fcell.2021.660069
- 143 Jahid S, Lipkin S. Mouse Models of Inherited Cancer Syndromes. *Hematol Oncol Clin North Am* 2010;**24**:1205–28.
- 144 Hunter AJ. Have animal models of disease helped or hindered the drug discovery process? *Ann N Y Acad Sci* 2011;**1245**:1–2.
- 145 Uhl EW, Warner NJ. Mouse Models as Predictors of Human Responses: Evolutionary Medicine. *Curr Pathobiol Rep* 2015;**3**:219–23.
- 146 Jongmans MCJ, Van Der Burgt I, Hoogerbrugge PM, Noordam K, Yntema HG, Nillesen WM, Kuiper RP, Ligtenberg MJL, Van Kessel AG, Van Krieken JHJM, Kiemeny LALM, Hoogerbrugge N. Cancer risk in patients with Noonan syndrome carrying a PTPN11 mutation. *Eur J Hum Genet* 2011;**19**:870–4.
- 147 Rosenquist R, Cuppen E, Buettner R, Caldas C, Dreau H, Elemento O, Frederix G, Grimmond S, Haerlach T, Jobanputra V, Meggendorfer M, Mullighan CG, Wordsworth S, Schuh A. Clinical utility of whole–genome sequencing in precision oncology. *Semin Cancer Biol* 2022;**84**:32–9.
- 148 Luthra R, Chen H, Roy–Chowdhuri S, Rajesh Singh R, Farah S, Chi–shing Cho W. Next–Generation Sequencing in Clinical Molecular Diagnostics of Cancer: Advantages and Challenges. Published Online First: 2015. doi:10.3390/cancers7040874
- 149 Wetterstrand KA. The Cost of Sequencing a Human Genome (GSP). <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost> (accessed 28 Sep2022).
- 150 Ultima Genomics Announces End of Early Access and Upcoming Launch of UG 100™. 2024.<https://www.businesswire.com/news/home/20240130828872/en/Ultima-Genomics-Announces-End-of-Early-Access-and-Upcoming-Launch-of-UG-100%E2%84%A2> (accessed 13 Dec2024).

- 151 Almogly G, Pratt M, Oberstrass F, Lee L, Mazur D, Beckett N, Barad O, Soifer I, Perelman E, Etzioni Y, Sosa M, Jung A, Clark T, Trepagnier E, Lithwick-Yanai G, Pollock S, Hornung G, Levy M, Coole M, Howd T, Shand M, Farjoun Y, Emery J, Hall G, Lee S, Sato T, Magner R, Low S, Bernier A, Gandi B, Stohlman J, Nolet C, Donovan S, Blumenstiel B, Cipicchio M, Dodge S, Banks E, Lennon N, Gabriel S, Lipson D. Cost-efficient whole genome-sequencing using novel mostly natural sequencing-by-synthesis chemistry and open fluidics platform. 2022. doi:10.1101/2022.05.29.493900
- 152 Kernohan KD, Boycott KM. The expanding diagnostic toolbox for rare genetic diseases. *Nat Rev Genet* 2024;**25**:401–15.
- 153 Erdmann F, Frederiksen LE, Bonaventure A, Mader L, Hasle H, Robison LL, Winther JF. Childhood cancer: Survival, treatment modalities, late effects and improvements over time. *Cancer Epidemiol* 2021;**71**:101733.
- 154 Gatta G, Botta L, Rossi S, Aareleid T, Bielska-Lasota M, Clavel J, Dimitrova N, Jakab Z, Kaatsch P, Lacour B, Mallone S, Marcos-Gragera R, Minicozzi P, Sánchez-Pérez M-J, Sant M, Santaquilani M, Stiller C, Tavilla A, Trama A, Visser O, Peris-Bonet R. Childhood cancer survival in Europe 1999–2007: results of EUROCare-5—a population-based study. *Lancet Oncol* 2014;**15**:35–47.
- 155 Huang J, Chan SC, Ngai CH, Lok V, Zhang L, Lucero-Prisno DE, Xu W, Zheng Z, Elcarte E, Withers M, Wong MCS. Global incidence, mortality and temporal trends of cancer in children: A joinpoint regression analysis. *Cancer Med* 2023;**12**:1903–11.
- 156 The Lancet Regional Health – Europe. Reduce cancer inequity and inequality to reduce cancer mortality. *The Lancet Regional Health – Europe* 2023;**25**:100591.
- 157 Green AL, Furutani E, Ribeiro KB, Galindo CR. JOURNAL OF CLINICAL ONCOLOGY Death Within 1 Month of Diagnosis in Childhood Cancer: An Analysis of Risk Factors and Scope of the Problem. *J Clin Oncol* 2017;**35**:1320–7.
- 158 Loeffen EAH, Knops RRG, Boerhof J, Feijen EAM (Lieke), Merks JHM, Reedijk AMJ, Lieverst JA, Pieters R, Boezen HM, Kremer LCM, Tissing WJE. Treatment-related mortality in children with cancer: Prevalence and risk factors. *Eur J Cancer* 2019;**121**:113–22.

- 159 Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, Friedman DL, Marina N, Hobbie W, Kadan-Lottick NS, Schwartz CL, Leisenring W, Robison LL. Chronic Health Conditions in Adult Survivors of Childhood Cancer. *New England Journal of Medicine* 2006;**355**:1572–82.
- 160 Haslam A, Kim MS, Prasad V. Updated estimates of eligibility for and response to genome-targeted oncology drugs among US cancer patients, 2006–2020. *Ann Oncol* 2021;**32**:926–32.
- 161 Lee YT, Tan YJ, Oon CE. Molecular targeted therapy: Treating cancer with specificity. *Eur J Pharmacol* 2018;**834**:188–96.
- 162 Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science (1979)* 2018;**359**:1350–5.
- 163 Miliotou AN, Papadopoulou LC. CAR T-cell Therapy: A New Era in Cancer Immunotherapy. *Curr Pharm Biotechnol* 2018;**19**:5–18.
- 164 DeMaria PJ, Bilusic M. Cancer Vaccines. *Hematol Oncol Clin North Am* 2019;**33**:199–214.
- 165 Brodeur GM, Nichols KE, Plon SE, Schiffman JD, Malkin D. Pediatric Cancer Predisposition and Surveillance: An Overview, and a Tribute to Alfred G. Knudson, Jr HHS Public Access. *Clin Cancer Res* 2017;**23**:1–5.
- 166 ERN GENTURIS. Clinical practice guidelines. <https://www.genturis.eu/l=eng/Guidelines-and-pathways/Clinical-practice-guidelines.html> (accessed 16 Sep2022).
- 167 For scientists and health professionals – FIT. <http://www.krebspraedisposition.de/en/for-scientists-and-health-professionals/> (accessed 22 Dec2022).
- 168 Pediatric Cancer Working Group of the American Association for Cancer Research (AACR). Pediatric Oncology Series | Clinical Cancer Research | American Association for Cancer Research. <https://aacrjournals.org/clincancerres/collection/57/Pediatric-Oncology-Series> (accessed 29 Sep2022).
- 169 Ludvigsson JF, Andersson E, Ekblom A, Feychting M, Kim JL, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the Swedish national inpatient register. *BMC Public Health* 2011;**11**:1–16.

- 170 Cnattingius S, Källén K, Sandström A, Rydberg H, Månsson H, Stephansson O, Frisell T, Ludvigsson JF. The Swedish medical birth register during five decades: documentation of the content and quality of the register. *Eur J Epidemiol* 2023;**38**:109–20.
- 171 Barlow L, Westergren K, Holmberg L, Tälback M. The completeness of the Swedish Cancer Register – a sample survey for year 1998. <https://doi.org/10.1080/02841860802247664> 2009;**48**:27–33.
- 172 Ludvigsson JF, Almqvist C, Bonamy AKE, Ljung R, Michaëlsson K, Neovius M, Stephansson O, Ye W. Registers of the Swedish total population and their use in medical research. *Eur J Epidemiol* 2016;**31**:125–36.
- 173 Ludvigsson JF, Svedberg P, Olén O, Bruze G, Neovius M. The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. *Eur J Epidemiol* 2019;**34**. doi:10.1007/S10654-019-00511-8
- 174 Ekblom A. The Swedish Multi-generation Register. *Methods Mol Biol* 2011;**675**:215–20.
- 175 Genomics England PanelApp. <https://panelapp.genomicsengland.co.uk/> (accessed 22 Aug 2024).
- 176 Byrjalsen A, Hansen TVO, Stoltze UK, Mehrjouy MM, Barnkob NM, Hjalgrim LL, Mathiasen R, Lautrup CK, Gregersen PA, Hasle H, Wehner PS, Tuckuviene R, Sackett PW, Laspiur AO, Rossing M, Marvig RL, Tommerup N, Olsen TE, Scheie D, Gupta R, Gerdes A, Schmiegelow K, Wadt K. Nationwide germline whole genome sequencing of 198 consecutive pediatric cancer patients reveals a high incidence of cancer prone syndromes. *PLoS Genet* 2020;**16**:e1009231.
- 177 Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nissan MH, Chang MT, Chandarlapaty S, Traina TA, Paik PK, Ho AL, Hantash FM, Grupe A, Baxi SS, Callahan MK, Snyder A, Chi P, Danila DC, Gounder M, Harding JJ, Hellmann MD, Iyer G, Janjigian YY, Kaley T, Levine DA, Lowery M, Omuro A, Postow MA, Rathkopf D, Shoushtari AN, Shukla N, Voss MH, Paraiso E, Zehir A, Berger MF, Taylor BS, Saltz LB, Riely GJ, Ladanyi M, Hyman DM, Baselga J, Sabbatini P, Solit DB, Schultz N. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* 2017;**1**:1–16.

- 178 Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ, Forbes SA. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* 2019;**47**:D941–7.
- 179 Harrison PW, Amode MR, Austine–Orimoloye O, Azov AG, Barba M, Barnes I, Becker A, Bennett R, Berry A, Bhai J, Bhurji SK, Boddu S, Branco Lins PR, Brooks L, Ramaraju SB, Campbell LI, Martinez MC, Charkhchi M, Chougule K, Cockburn A, Davidson C, De Silva NH, Dodiya K, Donaldson S, El Houdaigui B, Naboulsi TE, Fatima R, Giron CG, Genez T, Grigoriadis D, Ghattaoraya GS, Martinez JG, Gurbich TA, Hardy M, Hollis Z, Hourlier T, Hunt T, Kay M, Kaykala V, Le T, Lemos D, Lodha D, Marques–Coelho D, Maslen G, Merino GA, Mirabueno LP, Mushtaq A, Hossain SN, Ogeh DN, Sakthivel MP, Parker A, Perry M, Piližota I, Poppleton D, Prosovetskaia I, Raj S, Pérez–Silva JG, Salam AIA, Saraf S, Saraiva–Agostinho N, Sheppard D, Sinha S, Sipos B, Sitnik V, Stark W, Steed E, Suner M–M, Surapaneni L, Sutinen K, Tricomi FF, Urbina–Gómez D, Veidenberg A, Walsh TA, Ware D, Wass E, Willhoft NL, Allen J, Alvarez–Jarreta J, Chakiachvili M, Flint B, Giorgetti S, Haggerty L, Ilsley GR, Keatley J, Loveland JE, Moore B, Mudge JM, Naamati G, Tate J, Trevanion SJ, Winterbottom A, Frankish A, Hunt SE, Cunningham F, Dyer S, Finn RD, Martin FJ, Yates AD. Ensembl 2024. *Nucleic Acids Res* 2024;**52**:D891–9.
- 180 Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, Kasprzyk A. BioMart – biological queries made easy. *BMC Genomics* 2009;**10**:22.
- 181 Haas OA. Primary Immunodeficiency and Cancer Predisposition Revisited: Embedding Two Closely Related Concepts Into an Integrative Conceptual Framework. *Front Immunol* 2019;**9**. doi:10.3389/fimmu.2018.03136
- 182 Banda A, Naaldenberg J, Timen A, van Eeghen A, Leusink G, Cuypers M. Cancer risks related to intellectual disabilities: A systematic review. *Cancer Med* 2024;**13**. doi:10.1002/cam4.7210
- 183 Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer–Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin

- N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Neale BM, Daly MJ, MacArthur DG. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020 581:7809 2020;**581**:434–43.
- 184 Stranneheim H, Lagerstedt-Robinson K, Magnusson M, Kvarnung M, Nilsson D, Lesko N, Engvall M, Anderlid B-M, Arnell H, Johansson CB, Barbaro M, Björck E, Bruhn H, Eisfeldt J, Freyer C, Grigelioniene G, Gustavsson P, Hammarsjö A, Hellström-Pigg M, Iwarsson E, Jemt A, Laaksonen M, Enoksson SL, Malmgren H, Naess K, Nordenskjöld M, Oscarson M, Pettersson M, Rasi C, Rosenbaum A, Sahlin E, Sardh E, Stödberg T, Tesi B, Tham E, Thonberg H, Töhönen V, von Döbeln U, Vassiliou D, Vonlanthen S, Wikström A-C, Wincent J, Winqvist O, Wredenberg A, Ygberg S, Zetterström RH, Marits P, Soller MJ, Nordgren A, Wirta V, Lindstrand A, Wedell A. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med* 2021;**13**:40.
- 185 Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013;**14**:178–92.
- 186 Driscoll DJ, Miller JL, Schwartz S, Cassidy SB. Prader-Willi Syndrome. *Neuropediatrics* 2017;**29**:124–6.
- 187 Stoltze UK, Foss-Skiftesvik J, Hansen T van O, Rasmussen S, Karczewski KJ, Wadt KAW, Schmiegelow K. The evolutionary impact of childhood cancer on the human gene pool. *Nat Commun* 2024;**15**:1881.
- 188 Jaffray B, Moore L, Dickson AP. Brief report: Prader-Willi syndrome and intratubular germ cell neoplasia. *Med Pediatr Oncol* 1999;**32**. doi:10.1002/(SICI)1096-911X(199901)32:1<73::AID-MPO19>3.0.CO;2-X
- 189 Crinò A, Schiaffini R, Ciampalini P, Spera S, Beccaria L, Benzi F, Bosio L, Corrias A, Gargantini L, Salvatoni A, Tonini G, Trifirò G, Livieri C. Hypogonadism and pubertal development in Prader-Willi syndrome. *Eur J Pediatr* 2003;**162**:327–33.

- 190 Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, Smith PJ, Reeve AE. Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. *Nature* 1993;**362**:749–51.
- 191 Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP. Relaxation of imprinted genes in human cancer. *Nature* 1993;**362**:747–9.
- 192 Jelinic P, Shaw P. Loss of imprinting and cancer. *J Pathol* 2007;**211**:261–8.
- 193 McLeod C, Gout AM, Zhou X, Thrasher A, Rahbarinia D, Brady SW, Macias M, Birch K, Finkelstein D, Sunny J, Mudunuri R, Orr BA, Treadway M, Davidson B, Ard TK, Chiao A, Swistak A, Wiggins S, Foy S, Wang J, Sioson E, Wang S, Michael JR, Liu Y, Ma X, Patel A, Edmonson MN, Wilkinson MR, Frantz AM, Chang T-C, Tian L, Lei S, Islam SMA, Meyer C, Thangaraj N, Tater P, Kandali V, Ma S, Nguyen T, Serang O, McGuire I, Robison N, Gentry D, Tang X, Palmer LE, Wu G, Suh E, Tanner L, McMurphy J, Lear M, Pappo AS, Wang Z, Wilson CL, Cheng Y, Meshinchi S, Alexandrov LB, Weiss MJ, Armstrong GT, Robison LL, Yasui Y, Nichols KE, Ellison DW, Bangur C, Mullighan CG, Baker SJ, Dyer MA, Miller G, Newman S, Rusch M, Daly R, Perry K, Downing JR, Zhang J. St. Jude Cloud: A Pediatric Cancer Genomic Data-Sharing Ecosystem. *Cancer Discov* 2021;**11**:1082–99.
- 194 Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* 2012;**2**:401–4.
- 195 Collod-Bérout G, Bourdelles S Le, Ades L, Ala-Kokko L, Booms P, Boxer M, Child A, Comeglio P, Paepe A De, Hyland JC, Holman K, Kaitila I, Loeys B, Matyas G, Nuytinck L, Peltonen L, Rantamaki T, Robinson P, Steinmann B, Junien C, Bérout C, Boileau C. Update of the UMD-FBN1 mutation database and creation of an FBN1 polymorphism database. *Hum Mutat* 2003;**22**:199–208.
- 196 Hosur V, Kavirayani A, Riefler J, Carney LMB, Lyons B, Gott B, Cox GA, Shultz LD. Dystrophin and dysferlin double mutant mice: a novel model for rhabdomyosarcoma. *Cancer Genet* 2012;**205**:232–41.

- 197 Boscolo Sesillo F, Fox D, Sacco A. Muscle Stem Cells Give Rise to Rhabdomyosarcomas in a Severe Mouse Model of Duchenne Muscular Dystrophy. *Cell Rep* 2019;**26**:689–701.e6.
- 198 Saettini F, Poli C, Vengoechea J, Bonanomi S, Orellana JC, Fazio G, Rodriguez FH, Noguera LP, Booth C, Jarur-Chamy V, Shams M, Iascone M, Vukic M, Gasperini S, Quadri M, Barroeta Seijas A, Rivers E, Mauri M, Badolato R, Cazzaniga G, Bugarin C, Gaipa G, Kroes WGM, Moratto D, van Oostaijen-ten Dam MM, Baas F, van der Maarel S, Piazza R, Coban-Akdemir ZH, Lupski JR, Yuan B, Chinn IK, Daxinger L, Biondi A. Absent B cells, agammaglobulinemia, and hypertrophic cardiomyopathy in folliculin-interacting protein 1 deficiency. *Blood* 2021;**137**:493–9.
- 199 Niehues T, Özgür TT, Bickes M, Waldmann R, Schöning J, Bräsen J, Hagel C, Ballmaier M, Klusmann J, Niedermayer A, Pannicke U, Enders A, Dückers G, Siepermann K, Hempel J, Schwarz K, Viemann D. Mutations of the gene *FNIP1* associated with a syndromic autosomal recessive immunodeficiency with cardiomyopathy and pre-excitation syndrome. *Eur J Immunol* 2020;**50**:1078–80.
- 200 Baba M, Toyama H, Sun L, Takubo K, Suh H-C, Hasumi H, Nakamura-Ishizu A, Hasumi Y, Klarmann KD, Nakagata N, Schmidt LS, Linehan WM, Suda T, Keller JR. Loss of *Folliculin* Disrupts Hematopoietic Stem Cell Quiescence and Homeostasis Resulting in Bone Marrow Failure. *Stem Cells* 2016;**34**:1068–82.
- 201 Park H, Staehling K, Tsang M, Appleby MW, Brunkow ME, Margineantu D, Hockenbery DM, Habib T, Liggitt HD, Carlson G, Iritani BM. Disruption of *Fnip1* Reveals a Metabolic Checkpoint Controlling B Lymphocyte Development. *Immunity* 2012;**36**:769–81.
- 202 Siggs OM, Stockenhuber A, Deobagkar-Lele M, Bull KR, Crockford TL, Kingston BL, Crawford G, Anzilotti C, Steeples V, Ghaffari S, Czibik G, Bellahcene M, Watkins H, Ashrafian H, Davies B, Woods A, Carling D, Yavari A, Beutler B, Cornall RJ. Mutation of *Fnip1* is associated with B-cell deficiency, cardiomyopathy, and elevated AMPK activity. *Proc Natl Acad Sci U S A* 2016;**113**:E3706–15.
- 203 Baba M, Keller JR, Sun H-W, Resch W, Kuchen S, Suh HC, Hasumi H, Hasumi Y, Kieffer-Kwon K-R, Gonzalez CG, Hughes RM, Klein ME, Oh HF, Bible P, Southon E, Tessarollo L, Schmidt LS, Linehan WM, Casellas R. The folliculin-FNIP1

- pathway deleted in human Birt-Hogg-Dubé syndrome is required for murine B-cell development. *Blood* 2012;**120**:1254–61.
- 204 Park H, Tsang M, Iritani BM, Bevan MJ. Metabolic regulator Flnp1 is crucial for iNKT lymphocyte development. *Proceedings of the National Academy of Sciences* 2014;**111**:7066–71.
 - 205 Ramírez JA, Iwata T, Park H, Tsang M, Kang J, Cui K, Kwong W, James RG, Baba M, Schmidt LS, Iritani BM. Folliculin Interacting Protein 1 Maintains Metabolic Homeostasis during B Cell Development by Modulating AMPK, mTORC1, and TFE3. *The Journal of Immunology* 2019;**203**:2899–908.
 - 206 Baba M, Endoh M, Ma W, Toyama H, Hirayama A, Nishikawa K, Takubo K, Hano H, Hasumi H, Umemoto T, Hashimoto M, Irie N, Esumi C, Kataoka M, Nakagata N, Soga T, Yao M, Kamba T, Minami T, Ishii M, Suda T. Folliculin Regulates Osteoclastogenesis Through Metabolic Regulation. *Journal of Bone and Mineral Research* 2018;**33**:1785–98.
 - 207 Liu W, Chen Z, Ma Y, Wu X, Jin Y, Hou S. Genetic Characterization of the Drosophila Birt-Hogg-Dubé Syndrome Gene. *PLoS One* 2013;**8**:e65869.
 - 208 Kenyon EJ, Luijten MNH, Gill H, Li N, Rawlings M, Bull JC, Hadzhiev Y, van Steensel MAM, Maher E, Mueller F. Expression and knockdown of zebrafish folliculin suggests requirement for embryonic brain morphogenesis. *BMC Dev Biol* 2016;**16**:23.
 - 209 Diaz J, Berger S, Leon E. TFE3-associated neurodevelopmental disorder: A distinct recognizable syndrome. *Am J Med Genet A* 2020;**182**:584–90.
 - 210 Lehalle D, Vabres P, Sorlin A, Bierhals T, Avila M, Carmignac V, Chevarin M, Torti E, Abe Y, Bartolomaeus T, Clayton-Smith J, Cogné B, Cusco I, Duplomb L, De Bont E, Duffourd Y, Duijkers F, Elpeleg O, Fattal A, Geneviève D, Guillen Sacoto MJ, Guimier A, Harris DJ, Hempel M, Isidor B, Jouan T, Kuentz P, Koshimizu E, Lichtenbelt K, Loik Ramey V, Maik M, Miyakate S, Murakami Y, Pasquier L, Pedro H, Simone L, Sondergaard-Schatz K, St-Onge J, Thevenon J, Valenzuela I, Abou Jamra R, van Gassen K, van Haelst MM, van Koningsbruggen S, Verdura E, Whelan Habela C, Zacher P, Rivière J-B, Thauvin-Robinet C, Betschinger J, Faivre L. De novo mutations in the X-linked TFE3 gene cause intellectual disability with pigmentary mosaicism and storage disorder-like features. *J Med Genet* 2020;**57**:808–19.

- 211 Kenny J, Burcher S, Kohut K, Eastman N. Ethical Issues in Genetic Testing for Inherited Cancer Predisposition Syndromes: the Potentially Conflicting Interests of Patients and Their Relatives. *Curr Genet Med Rep* 2020;**8**:72–7.
- 212 van Engelen K, Barrera M, Wasserman JD, Armel SR, Chitayat D, Druker H, Gallinger B, Malkin D, Villani A. Tumor surveillance for children and adolescents with cancer predisposition syndromes: The psychosocial impact reported by adolescents and caregivers. *Pediatr Blood Cancer* 2021;**68**. doi:10.1002/pbc.29021
- 213 Gopie JP, Vasen HFA, Tibben A. Surveillance for hereditary cancer: Does the benefit outweigh the psychological burden?—A systematic review. *Crit Rev Oncol Hematol* 2012;**83**:329–40.
- 214 Kratz CP, Jongmans MC, Cavé H, Wimmer K, Behjati S, Guerrini-Rousseau L, Milde T, Pajtler KW, Golmard L, Gauthier-Villars M, Jewell R, Duncan C, Maher ER, Brugieres L, Pritchard-Jones K, Bourdeaut F. Predisposition to cancer in children and adolescents. *Lancet Child Adolesc Health* 2021;**5**:142–54.
- 215 Bakhuizen JJ, Bourdeaut F, Wadt KAW, Kratz CP, Jongmans MCJ, Waespe N. Genetic testing for childhood cancer predisposition syndromes: Controversies and recommendations from the SIOPE Host Genome Working Group meeting 2022. *EJC Paediatric Oncology* 2024;**4**:100176.
- 216 Resta R, Biesecker BB, Bennett RL, Blum S, Hahn SE, Strecker MN, Williams JL. A New Definition of Genetic Counseling: National Society of Genetic Counselors' Task Force Report. *J Genet Couns* 2006;**15**:77–83.
- 217 Recommendations of the European Society of Human Genetics. Genetic testing in asymptomatic minors: recommendations of the European Society of Human Genetics. *European Journal of Human Genetics* 2009;**17**:720–1.
- 218 Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, Daly MJ, Neale BM, Sunyaev SR, Lander ES. Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences* 2014;**111**. doi:10.1073/pnas.1322563111
- 219 Zhang Y, Long J, Lu W, Shu X-O, Cai Q, Zheng Y, Li C, Li B, Gao Y-T, Zheng W. Rare Coding Variants and Breast Cancer Risk: Evaluation of Susceptibility Loci Identified in Genome-Wide Association Studies. *Cancer Epidemiology, Biomarkers & Prevention* 2014;**23**:622–8.

- 220 Zheng J, Wang X, Li J, Wu Y, Chang J, Xin J, Wang M, Wang T, Wei Q, Wang M, Zhang R. Rare variants confer shared susceptibility to gastrointestinal tract cancer risk. *Front Oncol* 2023;**13**. doi:10.3389/fonc.2023.1161639
- 221 Artomov M, Stratigos AJ, Kim I, Kumar R, Lauss M, Reddy BY, Miao B, Daniela Robles-Espinoza C, Sankar A, Njauw C-N, Shannon K, Gragoudas ES, Marie Lane A, Iyer V, Newton-Bishop JA, Timothy Bishop D, Holland EA, Mann GJ, Singh T, Barrett J, Adams DJ, Jönsson G, Daly MJ, Tsao H. Rare Variant, Gene-Based Association Study of Hereditary Melanoma Using Whole-Exome Sequencing. *JNCI: Journal of the National Cancer Institute* 2017;**109**. doi:10.1093/jnci/djx083
- 222 Esai Selvan M, Onel K, Gnjjatic S, Klein RJ, Gümüş ZH. Germline rare deleterious variant load alters cancer risk, age of onset and tumor characteristics. *NPJ Precis Oncol* 2023;**7**:13.
- 223 Mitchell J, Camacho N, Shea P, Stopsack KH, Joseph V, Burren O, Dhindsa R, Nag A, Berchuck JE, O'Neill A, Abbasi A, Zoghbi AW, Alegre-Díaz J, Kuri-Morales P, Berumen J, Tapia-Conyer R, Emberson J, Torres JM, Collins R, Wang Q, Goldstein D, Matakidou A, Haefliger C, Anderson-Dring L, March R, Jobanputra V, Dougherty B, Carss K, Petrovski S, Kantoff PW, Offit K, Mucci LA, Pomerantz M, Fabre MA. Characterising the contribution of rare protein-coding germline variants to prostate cancer risk and severity in 37,184 cases. 2024. doi:10.1101/2024.05.10.24307164

