

From Department of Women's and Children's Health
Karolinska Institutet, Stockholm, Sweden

GENETIC STUDIES OF SUSCEPTIBILITY TO INFLAMMATION, AUTOIMMUNITY, AND HEMATOLOGICAL MALIGNANCY

Bianca Tesi



Stockholm 2017

The image shows a complex, chaotic attractor pattern, specifically the Lorenz Attractor, plotted in three dimensions. It consists of two interlocking, spiraling curves that form a butterfly-like shape, representing the sensitive dependence on initial conditions of the system.

The cover shows the Lorenz Attractor. In my research, I have witnessed a single nucleotide change produce a radical shift from health to disease. The Lorenz Attractor similarly exemplifies how non-linear dynamical systems evolve on vastly different trajectories in response to minute changes in initial conditions. Plotted in R version 3.3.1.

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

Published by Karolinska Institutet.

Printed by E-print AB

© Bianca Tesi 2017

ISBN 978-91-7676-774-0

To my parents Alberto and Nicoletta

Genetic Studies of Susceptibility to Inflammation, Autoimmunity, and Hematological Malignancy

THESIS FOR DOCTORAL DEGREE (Ph.D.)

The public defence of this thesis will be held at Skandiasalen, Q3:01, Astrid Lindgren Children's Hospital on Friday, November 24th, 2017, at 9.00 a.m.

By

Bianca Tesi

Principal Supervisor:

Professor Jan-Inge Henter
Karolinska Institutet
Department of Women's and Children's Health
Childhood Cancer Research Unit

Co-supervisor(s):

Assistant Professor Yenan T. Bryceson
Karolinska Institutet
Department of Medicine, Huddinge
Center for Hematology and Regenerative Medicine

Professor Magnus Nordenskjöld
Karolinska Institutet
Department of Molecular Medicine and Surgery

Marie Meeths, MD PhD

Karolinska Institutet
Department of Women's and Children's Health
Childhood Cancer Research Unit

Opponent:

Professor Cisca Wijmenga
University of Groningen
Faculty of Medical Sciences
Department of Genetics

Examination Board:

Professor David Gisselsson Nord
Lund University
Department of Laboratory Medicine
Section Clinical Genetics

Professor Richard Rosenquist Brandell
Karolinska Institutet
Department of Molecular Medicine and Surgery

Docent Jonas Abrahamsson

University of Gothenburg
Institute of Clinical Sciences
Department of Pediatrics

ABSTRACT

The immune system represents the body's defense against infectious organisms. Inborn defects of the immune system, called primary immunodeficiencies (PIDs), are a heterogeneous group of Mendelian disorders. Clinically, PIDs can cause isolated to broad susceptibility to pathogens, severe hyperinflammation, autoimmunity, allergy, and cancer. The studies in this thesis take advantage of the recent development in DNA-sequencing technologies and of our increased understanding of genetic variability to further explore the genetic architecture and phenotypic spectrum of hemophagocytic lymphohistiocytosis (HLH), an inborn error of lymphocyte cytotoxicity, and to elucidate the genetic factors behind autoimmunity and hematological malignancies in selected families.

Familial HLH (FHL) is a severe hyperinflammatory condition, genetically heterogeneous, caused by defective perforin-mediated lymphocyte cytotoxic activity. In **paper I** we use a high-throughput sequencing panel covering 12 HLH-related genes in 58 prospectively recruited patients with HLH and achieve a molecular diagnosis in 22 cases (38%). In **paper II** we show that perforin-deficiency due to biallelic *PRF1* missense variants is associated with broad intra-familial variability and clinical presentations seemingly unrelated to HLH, such as Hodgkin lymphoma. Using exome sequencing, in **paper III**, we identify biallelic pathogenic variants in *IFNGR1* and *INFGR2*, respectively, in two patients with HLH and disseminated mycobacterial infection. Previous studies have shown that HLH pathology is largely driven by IFN- γ . Instead, our findings suggest the existence of IFN- γ -independent mechanisms for the development of HLH. In **paper IV**, we uncover biallelic coding and non-coding variants in *RAB27A*, the gene responsible for Griscelli syndrome type 2 (GS2), in five unrelated patients with atypical HLH, normal pigmentation, and a functional defect suggestive of FHL. A complex structural variant disrupting the transcriptional start site (TSS) of one *RAB27A* transcript was shared among the patients. We show that the disrupted TSS is less predominantly used by melanocytes compared to lymphocytes, explaining the lack of hypopigmentation in these patients, otherwise present in GS2.

In **paper V** we report the beneficial effect of hematopoietic stem cell transplantation in a 14-year-old boy with LRBA deficiency and a seven-year history of severe autoimmune disorders.

In **paper VI** we uncover germline heterozygous missense variants in *SAMD9L*, a gene located on 7q21, in two families with multiple individuals affected by cytopenia, immunodeficiency, myelodysplastic syndrome (MDS) with cytogenetic aberrations of chromosome 7, and neurological disease. We show a gain-of-function (GOF) effect of the mutants, which inhibit cell proliferation. Germline *SAMD9L* GOF variants were lost in MDS cells and hematopoietic revertant mosaicism occurred frequently among less severely affected carriers. Our results indicate a strong selective advantage for hematopoietic cells that, through different somatic events, overcome the growth-inhibiting effect of germline *SAMD9L* GOF variants.

Taken together, these studies add to our understanding of the phenotypic and genetic spectrum of HLH, display the power of high-throughput sequencing in diagnostics of individuals affected by severe inflammation, autoimmunity, and hematological malignancies, and highlight *SAMD9L* as an important gene for regulation of hematopoietic cell proliferation.

LIST OF SCIENTIFIC PAPERS INCLUDED IN THE THESIS

- I. **Targeted high-throughput sequencing for genetic diagnostics of hemophagocytic lymphohistiocytosis.**
Tesi B, Lagerstedt-Robinson K, Chiang SC, Ben Bdira E, Abboud M, Belen B, Devecioglu O, Fadoo Z, Yeoh AE, Erichsen HC, Möttönen M, Akar HH, Hästbacka J, Kaya Z, Nunes S, Patiroglu T, Sabel M, Saribeyoglu ET, Tvedt TH, Unal E, Unal S, Unuvar A, Meeths M, Henter JI, Nordenskjöld M, Bryceson YT.
Genome Medicine (2015), 7,130.
- II. **Spectrum of atypical clinical presentations in patients with biallelic *PRF1* missense mutations.**
Tesi B, Chiang SC, El-Ghoneimy D, Hussein AA, Langenskiöld C, Wali R, Fadoo Z, Silva JP, Lecumberri R, Unal S, Nordenskjöld M, Bryceson YT, Henter JI, Meeths M.
Pediatr Blood Cancer. (2015) Dec;62(12):2094-100.
- III. **Hemophagocytic lymphohistiocytosis in 2 patients with underlying IFN- γ receptor deficiency.**
Tesi B^{*}, Sieni E^{*}, Neves C, Romano F, Cetica V, Cordeiro AI, Chiang S, Schlums H, Galli L, Avenali S, Tondo A, Canessa C, Henter JI, Nordenskjöld M, Hsu AP, Holland SM, Neves JF, Azzari C[#], Bryceson YT[#].
J Allergy Clin Immunol. (2015) Jun;135(6):1638-41.
- IV. **A *RAB27A* 5'UTR structural variant associated with late-onset hemophagocytic lymphohistiocytosis and normal pigmentation.**
Tesi B, Rascon J, Chiang SCC, Burnyte B, Löfstedt A, Fasth A, Heizmann M, Juozapaitė S, Kiudeliene R, Kvedaraite E, Miseviciene V, Muleviciene A, Müller ML, Nordenskjöld M, Samaitiene R, Speckmann C, Stankeviciene S, Zekas V, Voss M, Ehl S, Vaiciene-Magistris N, Henter J-I, Meeths M[#], Bryceson YT[#].
Manuscript (2017).
- V. **Successful hematopoietic stem cell transplantation in a patient with LPS-responsive beige-like anchor (LRBA) gene mutation.**
Tesi B, Priftakis P, Lindgren F, Chiang SC, Kartalis N, Löfstedt A, Lörinc E, Henter JI, Winiarski J, Bryceson YT, Meeths M.
J Clin Immunol. (2016) Jul;36(5):480-9.
- VI. **Gain-of-function *SAMD9L* mutations cause a syndrome of cytopenia, immunodeficiency, MDS, and neurological symptoms.**
Tesi B^{*}, Davidsson J^{*}, Voss M^{*}, Rahikkala E, Holmes TD, Chiang SCC, Komulainen-Ebrahim J, Gorcenco S, Rundberg Nilsson A, Ripperger T, Kokkonen H, Bryder D, Fioretos T, Henter JI, Möttönen M, Niinimäki R, Nilsson L, Pronk CJ, Puschmann A, Qian H, Uusimaa J, Moilanen J, Tedgård U, Cammenga J[#], Bryceson YT[#].
Blood. (2017) Apr 20;129(16):2266-2279.

^{*} shared first authorship

[#] shared senior authorship

RELATED SCIENTIFIC PAPERS

The following related papers are referred to in chapter 4 using Roman numerals.

VII. Partial oculocutaneous albinism and immunodeficiency syndromes: ten years experience from a single center in Turkey.

Patiroglu T, Akar HH, Unal E, Chiang SC, Schlums H, Tesi B, Ozkars MY, Karakukcu M.
Genet Couns. 2016;27(1):67-76.

VIII. The syndrome of hemophagocytic lymphohistiocytosis in primary immunodeficiencies: implications for differential diagnosis and pathogenesis.

Bode SF, Ammann S, Al-Herz W, Bataneant M, Dvorak CC, Gehring S, Gennery A, Gilmour KC, Gonzalez-Granado LI, Groß-Wieltsch U, Ifversen M, Lingman-Framme J, Matthes-Martin S, Mesters R, Meyts I, van Montfrans JM, Pachlopnik Schmid J, Pai SY, Soler-Palacin P, Schuermann U, Schuster V, Seidel MG, Speckmann C, Stepensky P, Sykora KW, Tesi B, Vraetz T, Waruiru C, Bryceson YT, Moshous D, Lehmberg K, Jordan MB, Ehl S; Inborn Errors Working Party of the EBMT.
Haematologica. 2015 Jul;100(7):978-88.

ADDITIONAL SCIENTIFIC PAPERS

Listed in chronological order:

- **A case of familial hemophagocytic lymphohistiocytosis type 4 with involvement of the central nervous system complicated with infarct.** Ciraci S, Ozcan A, Ozdemir MM, Chiang SCC, Tesi B, Ozdemir AM, Karakukcu M, Patiroglu T, Acipayam C, Doganay S, Gumus H, Unal E. *J Pediatr Hematol Oncol.* 2017 Aug;39(6):e321-e324.
- **Unperturbed cytotoxic lymphocyte phenotype and function in myalgic encephalomyelitis/chronic fatigue syndrome patients.** Theorell J, Bileviciute-Ljungar I, Tesi B, Schlums H, Johnsgaard MS, Asadi-Azarbaijani B, Bolle Strand E, Bryceson YT. *Front Immunol.* 2017 Jun 26;8:723.
- **Differences in granule morphology yet equally impaired exocytosis among cytotoxic T cells and NK cells from Chediak-Higashi syndrome patients.** Chiang SCC, Wood SM, Tesi B, Akar HH, Al-Herz W, Ammann S, Belen FB, Caliskan U, Kaya Z, Lehmberg K, Patiroglu T, Tokgoz H, Ünüvar A, Introne WJ, Henter JI, Nordenskjöld M, Ljunggren HG, Meeths M, Ehl S, Krzewski K, Bryceson YT. *Front Immunol.* 2017 Apr 18;8:426.
- **Targeting SAMHD1 with the Vpx protein to improve cytarabine therapy for hematological malignancies.** Herold N, Rudd SG, Ljungblad L, Sanjiv K, Myrberg IH, Paulin CB, Heshmati Y, Hagenkort A, Kutzner J, Page BD, Calderón-Montañ JM, Loseva O, Jemth AS, Bulli L, Axelsson H, Tesi B, Valerie NC, Höglund A, Bladh J, Wiita E, Sundin M, Uhlin M, Rassidakis G, Heyman M, Tamm KP, Warpman-Berglund U, Walfridsson J, Lehmann S, Grandér D, Lundbäck T, Kogner P, Henter JI, Helleday T, Schaller T. *Nat Med.* 2017 Feb;23(2):256-263.
- **Epigenetic regulation of adaptive NK Cell diversification.** Tesi B, Schlums H, Cichocki F, Bryceson YT. Epigenetic Regulation of Adaptive NK Cell Diversification. *Trends Immunol.* 2016 Jul;37(7):451-61.

- **Diversification and functional specialization of human NK cell subsets.** Cichocki F, Schlums H, Theorell J, Tesi B, Miller JS, Ljunggren HG, Bryceson YT. *Curr Top Microbiol Immunol.* 2016;395:63-94.
- **A hemophagocytic lymphohistiocytosis case with newly defined *UNC13D* (c.175G>C; p.Ala59Pro) mutation and a rare complication.** Balcı YI, Özgürler Akpinar F, Polat A, Kenar F, Tesi B, Greenwood T, Yalçın N, Koçyiğit A. *Turk J Haematol.* 2015 Dec;32(4):355-8.
- **Cytomegalovirus infection drives adaptive epigenetic diversification of NK cells with altered signaling and effector function.** Schlums H, Cichocki F, Tesi B, Theorell J, Beziat V, Holmes TD, Han H, Chiang SC, Foley B, Mattsson K, Larsson S, Schaffer M, Malmberg KJ, Ljunggren HG, Miller JS, Bryceson YT. *Immunity.* 2015 Mar 17;42(3):443-56.
- **Pathophysiology and spectrum of diseases caused by defects in lymphocyte cytotoxicity.** Meeths M, Chiang SC, Löfstedt A, Müller ML, Tesi B, Henter JI, Bryceson YT. *Exp Cell Res.* 2014 Jul 1;325(1):10-7.
- **Immunomodulatory activity of commonly used drugs on Fc-receptor-mediated human natural killer cell activation.** Theorell J, Gustavsson AL, Tesi B, Sigmundsson K, Ljunggren HG, Lundbäck T, Bryceson YT. *Cancer Immunol Immunother.* 2014 Jun;63(6):627-41.
- **Novel *STAT3* mutation causing hyper-IgE syndrome: studies of the clinical course and immunopathology.** Sundin M, Tesi B, Sund Böhme M, Bryceson YT, Pütsep K, Chiang SC, Thunberg S, Winiarski J, Wikström AC. *J Clin Immunol.* 2014 May;34(4):469-77.
- **An N-terminal missense mutation in *STX11* causative of FHL4 abrogates Syntaxin-11 binding to Munc18-2.** Müller ML, Chiang SC, Meeths M, Tesi B, Entesarian M, Nilsson D, Wood SM, Nordenskjöld M, Henter JI, Naqvi A, Bryceson YT. *sFront Immunol.* 2014 Jan 14;4:515.

LIST OF ABBREVIATIONS

1000G	1000 Genomes project
AD	autosomal dominant
ADCC	antibody-dependent cellular cytotoxicity
ADEM	acute disseminated encephalomyelitis
AML	acute myeloid leukemia
APC	antigen-presenting cells
AR	autosomal recessive
ATXPC	ataxia-pancytopenia
BCR	B-cell receptor
CAGE	cap analysis gene expression
CHIP	clonal hematopoiesis of indeterminate potential
CHS	Chediak-Higashi syndrome
CMV	cytomegalovirus
CN-LOH	copy-neutral loss of heterozygosity
CNS	central nervous system
CNV	copy number variant
CTL	cytotoxic T-lymphocytes
CTLA-4	cytotoxic lymphocyte antigen-4
CVID	common variable immunodeficiency
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
dPCR	digital PCR
EBV	Epstein-Barr virus
ENCODE	Encyclopedia of DNA Elements
ES	exome sequencing
ExAC	Exome aggregation consortium
FHL	familial hemophagocytic lymphohistiocytosis
GATK	Genome Analysis Toolkit
gnomAD	Genome Aggregation Database
GOF	gain-of-function
GoNL	Genome of the Netherlands
GRC	Genome reference consortium
GS2	Grisicelli syndrome type 2
GWAS	genome-wide association studies
HGMD	Human Gene Mutation Database
HIT	hybrid immunotherapy
HLH	hemophagocytic lymphohistiocytosis
HSC	hematopoietic stem cell
HSCT	hematopoietic stem cell transplantation

HTS	high-throughput sequencing
IFN	interferon
IGV	Integrative Genomics Viewer
IL	interleukin
ILC	innate lymphoid cell
InDels	small insertions and deletions
IPEX	immunodeficiency polyendocrinopathy enteropathy X-linked
IUIS	International Union of Immunological societies
LCMV	lymphocytic choriomeningitis virus
LOF	loss-of-function
LOH	loss of heterozygosity
LU	lytic unit
MDS	myelodysplastic syndrome
MHC	major histocompatibility complex
NFTC	normophosphatemic familial tumoral calcinosis
NGS	next-generation sequencing
NK	natural killer
OMIM	Online Mendelian Inheritance in Man
PAMP	pathogen-associated molecular pattern
PBMC	peripheral blood mononuclear cell
PGM	Personal Genome Machine
PID	primary immunodeficiency
pLI	probability of LOF intolerance
PRR	pattern recognition receptor
qPCR	quantitative PCR
RNA	ribonucleic acid
RVIS	residual variation intolerance score
SCID	severe combined immunodeficiency
sIL2R	soluble IL-2 receptor
SLE	systemic lupus erythematosus
SNV	single nucleotide variant
TCR	T cell receptor
TFP	teal fluorescent protein
TLR	toll-like receptor
TREC	T cell receptor excision circles
Treg	T regulatory cells
TSS	transcriptional start site
UPD	uniparental disomy
VEP	Variant Effector Prediction
WGS	whole-genome sequencing

WHO	World Health Organization
XL	X-linked
XLP	X-linked lymphoproliferative syndrome

PREFACE

This thesis and its public defence represent my final steps towards fulfilling the requirements for a doctoral degree from the Karolinska Institutet. I started this journey in 2012 when I moved to Stockholm and joined the research group of Professor Jan-Inge Henter at the Department of Women's and Children's Health of Karolinska Institutet, Stockholm. I have been registered as PhD student at Karolinska Institutet since April 2013.

During the last five years, I have focused on the genetic studies of a group of disorders characterized by susceptibility to inflammation, autoimmunity, and hematological malignancies. Under the guidance of a fantastic team of supervisors and in collaboration with talented PhD-students, postdocs, researchers, and inspiring clinicians, my studies have been scientifically and personally very rewarding.

The thesis is divided into eight chapters. Chapter 1 is an introductory chapter. The first part of chapter 1 provides background knowledge on human genetics and immunology, while the second part introduces the genetic diseases that have been object of my studies. Chapter 2 states the aim of the thesis. Chapter 3 provides an overview of the methodological workflow used in studies of Mendelian diseases. Chapter 4 summarizes and discusses the results reported in the six original studies included in this thesis. Chapter 5 and 6 respectively describe conclusions and future perspectives. Chapter 7 acknowledges those who have been instrumental in my development as a researcher. Chapter 8 includes a list of all the cited literature. In the printed version, the six publications are also attached.

A handwritten signature in black ink, appearing to read "Björn Wennerberg".

Stockholm, October 16th, 2017.

TABLE OF CONTENTS

1 INTRODUCTION.....	3
1.1 HUMAN GENETICS.....	3
1.1.1 <i>The Human Genome</i>	3
1.1.2 <i>Genetic variation and population studies</i>	5
1.1.3 <i>Genetic variation and Mendelian diseases</i>	7
1.2 THE IMMUNE SYSTEM	9
1.2.1 <i>Innate immunity</i>	9
1.2.2 <i>Adaptive immunity</i>	11
1.2.3 <i>Inter-individual variation of the immune system</i>	12
1.3 PRIMARY IMMUNODEFICIENCIES (PIDS)	13
1.3.1 <i>General aspects of PIDs</i>	13
1.3.2 <i>Diseases of immune dysregulation</i>	16
1.3.2.1 Hemophagocytic lymphohistiocytosis	16
1.3.2.2 Other diseases of immune dysregulation	24
1.4 PREDISPOSITION TO HEMATOLOGICAL MALIGNANCIES.....	25
1.4.1 <i>Risk factors</i>	26
1.4.2 <i>Inherited myeloid malignancies</i>	26
1.5 A SNAPSHOT OF THE FIELD IN 2013	28
2 AIMS OF THE THESIS	30
3 METHODOLOGICAL WORKFLOW	31
3.1 PATIENTS STUDIED.....	31
3.2 GENERATION OF GENETIC DATA.....	32
3.2.1 <i>HTS technologies</i>	32
3.2.2 <i>HTS applications</i>	33
3.3 ANALYSIS AND INTERPRETATION OF GENETIC DATA	35
3.4 VALIDATION OF GENETIC FINDINGS	38
3.5 IMMUNOLOGICAL ASSAYS FOR DIAGNOSIS OF PRIMARY HLH	39
4 RESULTS AND DISCUSSION.....	41
4.1 GENETIC ARCHITECTURE OF HLH	41
4.1.1 <i>Variant spectrum of primary HLH</i>	41
4.1.2 <i>Non-coding pathogenic variants</i>	42
4.1.3 <i>HLH beyond defects of cytotoxic lymphocytes: HLH in other PIDs</i>	46
4.1.4 <i>Genetic variation in HLH genes in the population</i>	47
4.1.5 <i>Integrated genetic and functional diagnostics for HLH</i>	48
4.2 PHENOTYPIC SPECTRUM IN DISORDERS OF IMMUNE DYSREGULATION.....	50
4.2.1 <i>Atypical manifestations of defects in lymphocyte cytotoxicity</i>	50
4.2.1.1 Late-onset HLH	50
4.2.1.2 Cancer	50
4.2.1.3 Predominant neurological manifestations.....	51
4.2.1.4 Genotype-phenotype correlations in atypical primary HLH	52
4.2.2 <i>Severe autoimmunity in LRBA deficiency</i>	53

4.3 GAIN-OF-FUNCTION VARIANTS IN <i>SAMD9L</i> PREDISPOSE TO MDS/AML WITH MONOSOMY 7	54
4.3.1 <i>SAMD9L-related phenotypes</i>	54
4.3.2 <i>Effect of SAMD9L pathogenic variants</i>	56
4.3.3 <i>Predisposition to MDS/AML with monosomy 7</i>	56
4.3.4 <i>Frequent hematopoietic revertant mosaicism in carriers of SAMD9L GOF variants</i>	58
5 CONCLUSIONS.....	61
5.1 SNAPSHOT OF THE FIELD IN 2017	63
6 FUTURE PERSPECTIVES	65
7 ACKNOWLEDGEMENTS.....	67
8 REFERENCES	70

1 INTRODUCTION

1.1 HUMAN GENETICS

“It’s [the human genome] a history book – a narrative of the journey of our species through time. It’s a shop manual, with an incredibly detailed blueprint for building every human cell. And it’s a transformative textbook of medicine, with insights that will give health care providers immense new powers to treat, prevent and cure disease.” - Dr. Francis Collins, while announcing the completion of the Human Genome Project on June 26th, 2000.

Deoxyribonucleic acid (DNA) is the hereditary material of all living organisms. DNA is essentially a code sequence of four types of chemical bases, also called nucleotides: adenine (A), guanine (G), cytosine (C), and thymine (T). It is organized as a double helix with two complementary strands (with A complementary to T and C to G) running in opposite directions and is compactly stored inside the nucleus of each cell. DNA controls the production of proteins, and regulates biological processes. As a result of evolution, our DNA sequence is to a variable extent shared with all other species.

The complexity of living human beings, *i.e.* us, is not achieved merely thanks to the information coded into the linear sequence of the DNA. At the DNA level additional information, defined as *epigenetics*, is obtained through molecular modification of nucleotides (*e.g.* methylation of cytosine) and/or of the proteins involved in the folding and packaging of DNA into the nucleus. Further complexity is created when the DNA is transcribed into to ribonucleic acid (RNA), and when RNA is translated into proteins. Besides intrinsic factors, we are also shaped and challenged by environmental factors.

The recent development in DNA-sequencing technology has empowered large-scale studies of several thousands of individuals, boosting our understanding of the DNA’s linear sequence and of its functional complexity. This knowledge has enabled findings with clinical impact and promises to revolutionize the field of medicine. In the following sections, I will review the structure and function of the human genome understood from large-scale sequencing studies and discuss the inter-individual genetic variation and its relation to diseases.

1.1.1 The Human Genome

The human genome is the complete set of hereditary information contained in human cells, encompassing both the nuclear and the mitochondrial genome. The presence of 46 chromosomes in human cells was first discovered in 1956 by Tjio and Levan (Tjio and Levan, 1956). These are divided into 23 pairs of chromosomes, of which 22 are autosomes and 1 pair is composed of sex chromosomes, XX in females and XY in males. At the molecular level the genome contains *genes*. Genes are regions of DNA that constitute single units of transcription and inheritance, usually coding for a protein or a RNA molecule (Orgogozo et al., 2016).

A number of key methodological advances, including the discovery of a method for DNA-sequencing by Frederick Sanger in 1977 (Sanger et al., 1977) and the invention of recombinant DNA and cloning, paved the way to the conception of the Human Genome Project. Officially launched in 1990, the Human Genome Project aimed to obtain the sequence of the majority of the *euchromatic* DNA - the gene-rich and transcriptionally active portion of the human genome. A first draft of the human genome was published in 2001 by the Human Genome Project and by the private foundation Celera Genomics (Lander et al., 2001; Venter et al., 2001). Since then, the human genome assembly has undergone improvements and is currently curated by the Genome Reference Consortium (GRC). The human genome build GRCH37, released in February 2009, has dominated these recent years of large-scale sequencing both in research and clinical settings, and is also the build used in this thesis. Interestingly, this build is based on a mosaic haploid genome derived from up to 13 individuals, even though 66% of the total sequences are derived from a single male individual (E pluribus unum, 2010). In December 2013, the newest build - GRCH38 - has been announced. One of its main advantages is the removal of thousands of genetic variants that were found to be artifacts. Based on this and other improvements, it is compelling for the human genomics community to expedite a shift to this assembly.

The first draft of the human genome pushed forward efforts to understand its function. With approximately 20,500 protein-coding genes, accounting for only about 1.1% of the genome (Clamp et al., 2007), the function of most of our genome remained enigmatic. Repetitive and not conserved regions were simply labeled as “junk”. Sidney Brenner defined junk regions of DNA as “*junk, as the rubbish you keep, but not garbage, the stuff you throw out*” (Brenner, 1998). Several projects were launched aimed at understanding the function of our genome. Methodological advancements promoted by the Human Genome Project and in particular the development of *next-generation sequencing* or high-throughput sequencing (HTS) made these efforts feasible (Metzker, 2010). In 2012, the results from the ENyclopedia Of DNA Elements (ENCODE) project were published (Consortium, 2012). The project produced and analyzed 1640 datasets from 147 cell types using several techniques for the analysis of transcriptomes, histone modifications, and chromatin-accessibility. The project’s main conclusion was that 80% of the genome was *functional* in terms of either being involved in encoding for proteins or RNA, or displaying a reproducible signature (e.g. histone modification, protein-binding feature, etc.). ENCODE’s definition of *function* was strongly criticized, mainly from an evolutionary viewpoint where biological function is expected to be actively protected by selection (Eddy, 2012; Doolittle, 2013; Graur et al., 2013). In fact, only 5% of the genome appears to be under evolutionary constraint (Lindblad-Toh et al., 2011). Regardless, ENCODE’s contribution of refined methods to systematically analyze epigenetics data needs to be acknowledged. Similar endeavors are taken up by the FANTOM project (The FANTOM Consortium and the RIKEN PMI and Clst (dgt), 2014), the Roadmap Epigenomics (Roadmap Epigenomics Consortium et al., 2015), and the Blueprint Epigenome

project (Stunnenberg et al., 2016). The challenge of understanding the function of all 3 billion nucleotides that makes our genome is still largely unsolved.

1.1.2 Genetic variation and population studies

Our genome is *polymorphic*, meaning that the DNA sequence at specific positions will differ between individuals. Genetic variants originate from an initial mutation event that can be passed on to the next generation if occurring in gametes. Such mutation events can derive from errors in DNA replication or from the effect of chemicals. At a population level, genetic variants, if not under purifying selection, can persist and propagate through genetic recombination, gene flow, and sexual reproduction. As a result, each individual genome is unique. The study of the genetic basis of diseases is founded in the possibility to discern benign variants from potentially pathogenic variants. For effective discernment, a solid understanding of inter-individual genetic variation is essential.

Variation in our genome can be simplified into small and large-scale variants. To the category of small-scale variants belong single nucleotide variants (SNVs) and small insertions and deletions (InDels). SNVs involve the substitution of a single nucleotide. InDels include smaller deletions, duplications or insertions. To the group of large-scale variants belong structural variants (SVs). SVs can be balanced, like translocations and inversions, and unbalanced, like deletions or duplications. SVs that alter the dosage of a specific genomic region, usually larger than 1kb, are also defined as a copy number variant (CNV) (Redon et al., 2006).

A good way to deeply understand inter-individual genetic variation is to analyze individuals from different parts of the world. The HapMap project, which started after the completion of the Human Genome Project, was the first project to collect variants on a global scale, aiming to construct a haplotype map of the human genome (Frazer et al., 2007; Consortium, 2010). During this time, sequencing of individual genomes like those of Craig Venter (Levy et al., 2007) and James Watson (Wheeler et al., 2008) had also become possible, providing a first estimation of the inter-individual genetic variation. The 1000 Genomes Project (1000G) represents a milestone in this endeavor: 2504 individuals from 26 populations were sequenced to produce a catalog of genetic variants that enabled understanding of how a typical genome looks like (McVean et al., 2012; The 1000 Genomes Project Consortium, 2015) (Table 1). 1000G uncovered 88 million variants, of which most (62 million) were rare with a frequency of <0.5%. The frequency data for all the variants was made publicly available and has aided genetic studies of hereditary diseases by different communities.

Up to 7% of our genome is variable due to CNVs (Conrad et al., 2010; Sudmant et al., 2015), making this type of variation relevant in the context of diseases. CNVs are fewer in numbers but much larger in size compared to SNVs that only account for 1.1% of the human genome

variability. Interestingly, deletions seem to be under a stronger selective constraint than duplications (Sudmant et al., 2015).

Samples	503	
Mean coverage	7.4	
	Var. Sites	Singletons
SNPs	3.53M	11.4k
Indels	546k	-
Large deletions	939	5
CNVs	157	1
MEI (Alu)	919	0
MEI (L1)	123	0
MEI (SVA)	53	0
MEI (MT)	4	0
Inversions	9	0
Nonsynonymous	10.2K	116
Synonymous	11.2K	59
Intronic	1.68M	5.68k
Untranslated regions	30.0k	129
Promoter	82.2k	336
Insulator	57.7k	189
Enhancer	288k	1.02k
Transcription factor binding site	749	3
Filtered Loss-of-Function	149	3
HGMD-disease mutation	18	2
GWAS	2.08k	0
ClinVar	29	1

k, thousand; M, million; MEI, mobile element insertions.

Table 1. Median of autosomal variant sites per genome in Europeans. Adapted from 1000G (The 1000 Genomes Project Consortium, 2015).

Sequencing projects across many populations have produced an overview of global genetic variation, but are unsuited to identify population-specific rare variants. To address this problem, several countries have undertaken countrywide sequencing studies. Sequencing of 769 individuals of Dutch origin, within the Genome of the Netherlands (GoNL) project, revealed additional 7.6 million novel SNVs (The Genome of the Netherlands Consortium, 2014). Subsequently, 2636 Icelanders (Gudbjartsson et al., 2015), 1070 Japanese (Nagasaki et al., 2015), 1005 Qataris (Fakhro et al., 2016), around 4000 individuals from UK (The UK10K Consortium, 2015), and 1000 Swedes (Ameur et al., 2017) have been whole-genome sequenced, improving our understanding of how variation in the human genome contributes to human diversity. Focusing on the protein-coding portion of the genome, the Exome Aggregation Consortium (ExAC) project analyzed 60,706 individuals to produce the largest publicly available dataset of variants derived from sequencing of all exons *i.e.*, exome sequencing (Lek et al., 2016). With about one variant identified per every 8 base pairs, this dataset included 7.4 million high-quality variants, of which 99% had a frequency of <1% and 72% were novel compared to previous sequencing projects. Since its public release in 2014,

the ExAC dataset has had an unprecedented impact on clinical and research studies of genetic diseases. Its genome-based counterpart, the genome aggregation database (gnomAD), released in 2016, currently includes data from 123,136 exomes and 15,496 whole-genomes from unrelated individuals.

Altogether, these projects have revealed that humans harbor thousands of genetic variants, some of which are very rare. Since rare genetic variants are more common than genetic diseases, most rare variants are inevitably harmless. This is an important concept for the correct interpretation of findings from genetic studies, both in clinical and research settings.

1.1.3 Genetic variation and Mendelian diseases

Genetic variants can have very different phenotypic impact. Some can be embryonically lethal while others can have no obvious phenotypic effects. The studies included in this thesis focus on those highly penetrant genetic variants that are responsible for Mendelian disorders. In Mendelian disorders the phenotype is caused by pathogenic variants, also defined as *disease-causing variants*, in a single gene. The effect of such pathogenic variants is so large that they overcome most of the individual's background genetic variation. As a result, such variants are subjected to purifying selection and their frequency in the general population is low. On the other end of the spectrum are complex traits and diseases. Here, common genetic variants with modest effects contribute to what is known as a *polygenic* inheritance (Figure 1). The once strict distinction between Mendelian disorders and complex diseases is now fading. Some rare diseases can now be explained by a digenic or oligogenic inheritance, and the *mutational load* can modulate the phenotype in others (Katsanis, 2016). Moreover, it is known that about 20% of Mendelian genes contain variants associated to complex diseases by genome-wide association studies (GWAS) or are located in the proximity of such variants (Chong et al., 2015). Comorbidity between Mendelian disorders and complex diseases has also been observed (Blair et al., 2013). Intriguingly, the concept of “polygenic” has recently been challenged in favor of an “omnigenic” view, where potentially all variants can affect all phenotype in a cell (Boyle et al., 2017).

An ever-growing number of disease-causing genes have been elucidated since disease-causing variants in *CYBB* were identified as the cause for chronic granulomatous disease using positional cloning in 1986, incidentally the year I was born. By the time the Human Genome Project was completed, already 1000 Mendelian genes were known (Peltonen and McKusick, 2001). Mendelian disorders are cataloged in the “Online Mendelian Inheritance in Man” (OMIM) database (McKusick, 2007). As of September 20, 2017, the OMIM database contained above 5,000 Mendelian phenotypes caused by pathogenic variants in 3,425 different genes. The development seen in the field of genetics in the last 31 years is revolutionizing medicine and biology at an astonishing pace.

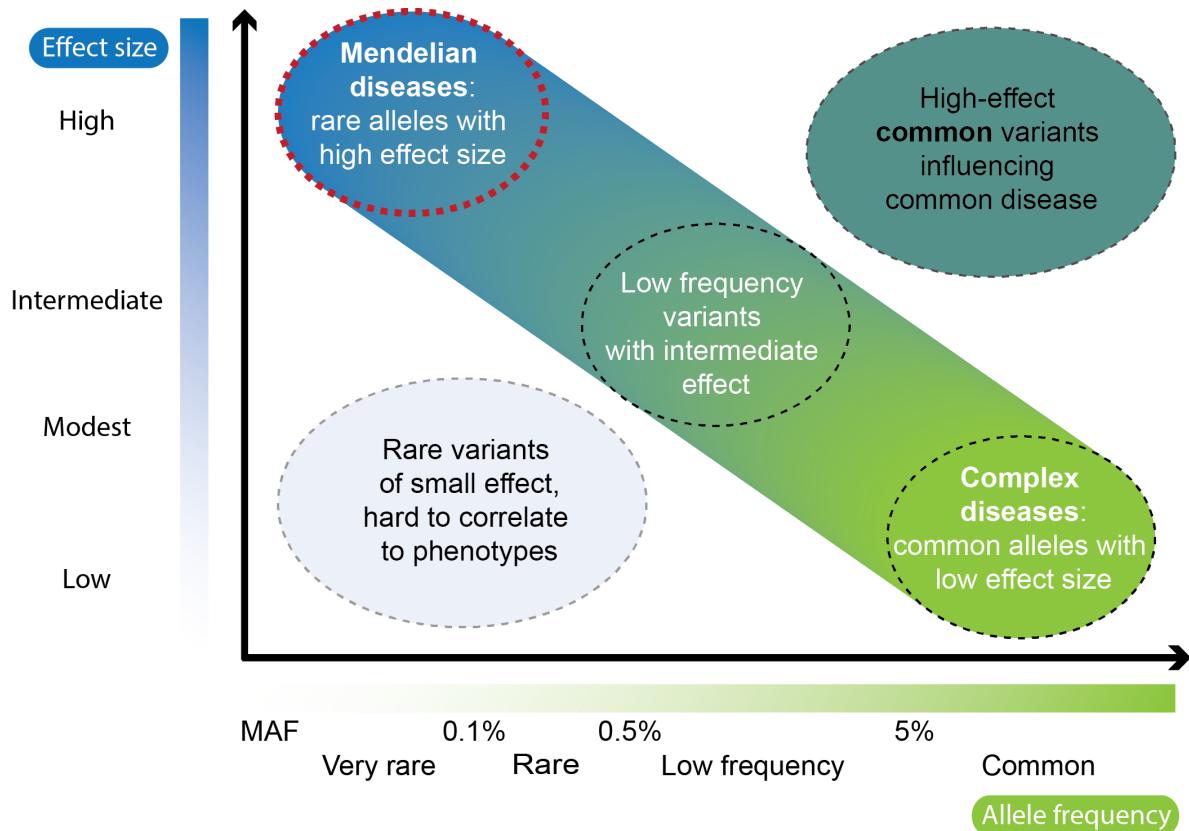


Figure 1. The spectrum of genetic variants and their relation to disease displayed according to effect size (how much a variant affects a phenotype) and allele frequency (frequency of the variant in the population). Figure is adapted from (Manolio et al., 2009). MAF= minor allele frequency.

Mendelian diseases are defined as *autosomal* if the mutated gene is located on an autosomal chromosome or *X-linked* if located on the X chromosome. Furthermore, they are classified as *dominant* if one mutated allele is enough to cause the phenotype, and *recessive* if both alleles need to harbor a pathogenic variant in order to cause a phenotype. Some Mendelian diseases occur *de novo*, meaning they are due to pathogenic variants only present in the affected individual but not in the germ line of the parents. In addition, mosaicism, incomplete penetrance, and variable expressivity may be present.

The confident assignment of pathogenicity to a genetic variant is key to the study of Mendelian diseases. Over the years all published disease-causing variants in Mendelian genes have been collected in the Human Gene Mutation Database (HGMD) (Stenson et al., 2017). Multiple large-scale sequencing studies have warned about the fact that several variants reported in the HGMD database as disease-causing are also present in reasonably healthy individuals. The number of such variants was, on average, 20 to 54 per individual, in the GoNL and the ExAC data, respectively (The Genome of the Netherlands Consortium, 2014; Lek et al., 2016). Carrier-ship for recessive disorders, incomplete penetrance, and variable expressivity can explain the detection of the red-flagged variants in healthy persons.

However, previous misclassification of genetic variants cannot be ruled out, which complicates the interpretation of personal genomes. Validation of HGMD data in populations with a high level of homozygosity due to inbreeding, such as the Saudi population, has indeed revealed the presence of individuals homozygous for variants red-flagged for autosomal dominant diseases. This increases the suspicion that some variants in HGMD are wrongly classified (Abouelhoda et al., 2016). Epistasis, *i.e.* genetic background, may also modify the pathogenicity of certain variants.

Scrutinizing presumably healthy adults for variants that truly cause severe childhood disorders is a good way to understand resilience to genetic disorders and possible buffering mechanisms (Chen et al., 2016b). A known buffering mechanism is somatic revertant mosaicism, which involves the *in vivo* correction of a pathogenic variant in a somatic cell. This mechanism can rescue or improve the phenotype and is most commonly seen in genetic diseases affecting tissues with fast cell turnover such as skin and blood (Hirschhorn, 2003). In a large study of over half a million individuals, 13 resilient individuals were detected (Chen et al., 2016b). Further studies of so-called “genetic superheroes” (MacArthur, 2016) could teach us important concepts for the treatment of genetic diseases.

1.2 THE IMMUNE SYSTEM

Our body is continuously exposed to potentially harmful microorganisms, yet we get ill relatively rarely. This is because the immune system, a set of tissues, cells, and molecules dedicated to defense from pathogens, is constantly working to maintain cellular and tissue homeostasis.

Traditionally, the immune system is divided into innate and adaptive immunity where the former represents an evolutionary conserved, initial, rapid, and general defense and the latter a more tailored and effective form of defense. Both arms of the immune system are capable of pathogen recognition, yet through different mechanisms. Innate immune cells are equipped with germline-encoded pattern recognition receptors, while adaptive immune cells can generate a vast repertoire of receptors through somatic recombination. Our immune system has evolved to a high level of complexity to ensure protection from pathogens as well as to avoid unnecessary response against “self”, our body components themselves. The following sections will give a brief overview of innate and adaptive immunity and of the inter-individual variation in immune response and its relation to disease.

1.2.1 Innate immunity

Innate immune cells include phagocytes, mast cells, monocytes, macrophages, dendritic cells, and innate lymphoid cells. Innate immune cells are responsible for removing microorganisms through phagocytosis, for inducing inflammation and antiviral response, as well as for supporting lymphocytes during adaptive response. They sense the extracellular and intracellular environment through germline-encoded pattern recognition receptors (PRRs),

which are able to recognize pathogen-associated molecular patterns (PAMPs) (Broz and Monack, 2013). One such class of PRRs is Toll-Like receptors (TLRs), of which to date 10 different types have been identified in humans. TLRs are located on the cell surface to recognize microbial proteins, lipids, and polysaccharides or inside endosomes to recognize nucleic acids from intracellular pathogens (Kawai and Akira, 2006). Stimulation of TLRs leads to the activation of the nuclear factor-kB (NF-kB) and of interferon-regulatory factors. End results of TLRs stimulation are the production of cytokines and interferons (IFNs) (Broz and Monack, 2013). Innate responses can also be triggered by damage-associated molecular patterns, endogenous molecules released from damaged or dying cells (Gallucci et al., 1999; Kono and Rock, 2008).

Natural killer (NK) cells were identified over 40 years ago as innate immune cells capable of cytotoxic activity against tumor cells without prior sensitization (Kiessling et al., 1975). During the last few years NK cells have been reclassified as belonging to the larger group of innate lymphoid cells (ILC), among which NK cells are the only cell type with cytotoxic activity (Artis and Spits, 2015). NK cells have germline-encoded activating receptors that bind ligands on the surface of target cells and, through synergy, can lead to cell activation (Bryceson et al., 2005, 2006, 2011). The activating receptor CD16 recognizes antibodies (immunoglobins G) on the surface of cells and thus mediates antibody-dependent cellular cytotoxicity (ADCC). Cytotoxicity towards normal cells is also restricted by the presence of germline-encoded inhibitory receptors that recognize major histocompatibility complex (MHC) class I molecules, present on all human cells, in order to elicit cytotoxicity only when self is “missing” (Ljunggren and Kärre, 1990; Bryceson et al., 2011). Once activated, NK cells have potent cytotoxic activity towards virus-infected cells as well as transformed cells and produce cytokines and chemokines. Thus, NK cells are important for proper antiviral defense and tumor immunosurveillance.

Innate immune cells cooperate to support inflammation and promote the adaptive response. Activated macrophages produce interleukin (IL)-15 which promotes proliferation of NK cells, and IL-12, which enhances the IFN- γ producing capacity of NK cells (Artis and Spits, 2015). In turn, the IFN- γ produced by NK cells activates macrophages. Macrophages and dendritic cells are antigen-presenting cells (APCs), and, together with B-cells, possess MHC class II molecules for the presentation of peptides to lymphocytes (Neefjes et al., 2011). Dendritic cells provide co-stimulatory signals to T cells necessary for their activation (Acuto and Michel, 2003), and promote T cell proliferation and differentiation through cytokines production. Also other ILCs besides NK cells support adaptive immune response through cytokine production and, vice versa, adaptive immune cells seem to regulate ILC responses (Gasteiger and Rudensky, 2014), demonstrating the extent of interplay and crosstalk between different immune cell types.

1.2.2 Adaptive immunity

Adaptive immune cells, such as B and T lymphocytes, are able to perform RAG-dependent recombination of antigen receptors (Mombaerts et al., 1992). This is a fascinating genetic mechanism, which emerged in jawed vertebrates about 500 million years ago (Agrawal et al., 1998). Both immunoglobulins (B cell receptor, BCR) and the T cell receptor (TCR) have a variable and a constant region in their chains (Davis and Bjorkman, 1988). The variable region is composed of different gene segments, called V, D, and J, located separately in the germline DNA. For each V, D, and J segments, several different functional segments exist and therefore different V(D)J combinations can be created (Jung et al., 2006). Specialized lymphocyte-specific enzymes, RAG1 and RAG2, realize this process (Mombaerts et al., 1992). Removing and adding a few nucleotides at the joining points introduce further diversity. In activated B cells further specificity is achieved by somatic hypermutation at the V region of an already rearranged immunoglobulin gene (Longerich et al., 2006). The RAG-dependent recombination of antigen receptors generates a receptor repertoire as vast as a 10^{11} order of magnitude, enabling an adaptive immune response against a great variety of antigens.

The rearrangement of antigen receptors is the first step of a tightly regulated developmental process, which takes place in the bone marrow and the thymus, for B and T cells, respectively. Essential during development is the selection of B and T cells without affinity for “self”, in order to avoid immune response against self-antigens and therefore autoimmunity. This process is known as acquisition of *tolerance* against self, further divided into central, acquired in bone marrow or thymus, and peripheral. Central tolerance is acquired through positive selection, where a weak interaction with self promotes survival. Maturing T cells are also subjected to negative selection, which induce cell death in maturing T cells with a strongly self-reactive TCR (Starr et al., 2003). Once mature B and T cells reach the peripheral lymphoid organs and encounter antigens matched to their specificity, stimulation of the receptors and the presence of co-stimulatory molecules promote their activation, clonal expansion, and acquisition of effector functions.

B cells are responsible of humoral immunity by producing antibodies while T cells are involved in cellular responses. CD4⁺ helper T cells, subdivided in Th1, Th2 and Th17, produce cytokines, which recruit and activate other leukocytes, promoting phagocytosis and killing of pathogens. CD8⁺ cytotoxic T lymphocytes (CTLs) kill virus-infected cells and malignant cells through their cytotoxic activity. For activation, CD8⁺ T cells require recognition of MHC class I molecules. CTLs share their cytotoxic mechanism with NK cells: upon contact with a target cell, NK and CTLs create an immunological synapse to which specialized lysosomes that contain perforin, called cytotoxic granules, are recruited. Once the cytotoxic granules are released, perforin creates pores on the membrane of the target cells thus facilitating entry of granzymes, which cause apoptosis (de Saint Basile et al., 2010). Another important group of T cells is the group of regulatory T cells (Treg), whose role in maintaining

self-tolerance and immune homeostasis is indispensable to protect us from autoimmunity (Sakaguchi et al., 2010).

Characteristic of adaptive immunity, together with the capacity to somatically recombine their antigen receptors, is the ability to generate long-lived memory cells towards specific antigens. Such cells will survive long after the infection and will mount a faster and stronger response during a repeated encounter with the same antigen. However, there is growing evidence that innate cells are also capable of immunological memory. For instance, adaptive features have also been recognized in NK cells (Schlums et al., 2015; Tesi et al., 2016).

1.2.3 Inter-individual variation of the immune system

Variability in individual responses to infections exists. As with many human traits, even variability in immune responses encompasses a large spectrum. At the population level humans are largely immunocompetent, yet at the individual level our immune system can fail. Individuals susceptible to very severe infections due to inborn errors of the immune system represent the extreme end of this spectrum (Casanova and Abel, 2005). A more detailed overview of this group of diseases, called primary immunodeficiencies (PID), will be provided in the next section. The study of individuals with rare PIDs has been instrumental to map down key genes for immune system function, yet more studies remain to be done to understand the impact of genetic and extrinsic factors on the individual's immunological make-up.

Technological development has been instrumental for exploring the diversity of immune cell repertoire in healthy adults. The advent of methods such as multi-parameter flow cytometry and mass cytometry has enabled large-scale analysis of immune cell repertoire in healthy adults at baseline and in response to stimuli (Tsang et al., 2014; Brodin et al., 2015; Carr et al., 2016). Because of its accessibility, most studies have been performed in blood. These studies have revealed marked inter-individual variation. Instead, intra-individual variation in the immunological landscape was minimal at steady state during longitudinal measurements in a time window of weeks to months. Moreover, acute perturbations, such as infections or vaccinations, introduced only a temporary change followed by a return to baseline, defined as the individual's "immunotype" (Tsang et al., 2014; Carr et al., 2016; Brodin and Davis, 2017). In the future, it will be interesting to see if these refined immunotypes may help identify groups at risk for specific infections or immune-mediated diseases.

Inter-individual variation in the immune system composition and function has been attributed to both heritable and non-heritable factors. Studies in twins have been fundamental to our understanding of the genetic contribution to a specific disease. This approach has also been used to dissect heritability of immune traits (Evans et al., 1999) and immune-mediated diseases (Cooper et al., 1999). Combined with advanced immunophenotyping, twin studies can provide a fine image of heritability of immune traits (Brodin et al., 2015; Roederer et al.,

2015). Brodin *et al.* studied 204 different parameters by mass cytometry in 105 twin pairs. Interestingly, 58% of all measurements, which included cell population frequencies, cytokine response, and serum proteins, were found to be largely non-heritable. Among the non-heritable factors, multiple studies have shown a large effect of age, characterized by an age-dependent reduction in lymphoid output (Brodin *et al.*, 2015; Aguirre-Gamboa *et al.*, 2016; Carr *et al.*, 2016). This is also reflected in younger twin pairs displaying more heritable traits (Evans *et al.*, 1999; Brodin *et al.*, 2015). Gender and seasonal effects have also been observed (Aguirre-Gamboa *et al.*, 2016; Brodin and Davis, 2017). Because of its function, it is intuitive that environmental factors such as the composition of the microbiota and exposition to certain pathogens can produce adaptive changes in the immune system (Brodin *et al.*, 2015; Schirmer *et al.*, 2016; Brodin and Davis, 2017). In the Brodin *et al.* study, monozygotic twins discordant for exposure to cytomegalovirus (CMV) infection displayed much less heritability of immune traits, showing that a single viral agent can leave a large imprint on the human immune system (Brodin *et al.*, 2015). GWAS studies have revealed several genomic loci associated with frequencies of specific immune cells or cytokine concentrations (Nalls *et al.*, 2011; Orrù *et al.*, 2013; Roederer *et al.*, 2015; Aguirre-Gamboa *et al.*, 2016; Li *et al.*, 2016). Several of the genomic loci highlighted by these studies overlap with hits from GWAS studies in immune-mediated diseases. Further studies of the shared genetic architecture of normal and diseased immune system might help pinpoint cell types and processes involved in disease pathogenesis and might suggest new therapeutic targets.

1.3 PRIMARY IMMUNODEFICIENCIES (PIDS)

Having appreciated the complexity of the human genome and of the immune system in the previous sections, we are now ready to dive into PIDs: the phenotypic results of highly damaging genetic variants in genes that are *essential* for immune system development and function.

The group of PIDs studied in this thesis concern diseases of immune dysregulation. We begin with an overview of general aspects of PIDs, followed by a more detailed description of diseases of immune dysregulation with a focus on *hemophagocytic lymphohistiocytosis*, a hyperinflammatory disease whose genetic architecture is the topic of most studies included in this thesis.

1.3.1 General aspects of PIDs

The knowledge of PIDs has evolved much over the last decades. PIDs are a heterogeneous group of inborn errors of immunity. About 300 genes responsible for PIDs are known today (Picard *et al.*, 2015). The phenotypic spectrum has recently expanded: traditionally associated with vulnerability to infections, PIDs are now linked also to autoimmunity, autoinflammation, allergy, and cancer (Notarangelo and Casanova, 2009).

The prevalence of PIDs has mostly been based on data from registries and is thought to be largely underestimated. Calculations suggest that 1/1,200 people worldwide might be affected by PID. PIDs are therefore a medical condition far from rare (Bousfiha et al., 2013).

PIDs have been classified by the International Union of Immunological Societies (IUIS) into the following 9 groups depending on their pathogenesis (Bousfiha et al., 2015; Picard et al., 2015): 1) immunodeficiencies affecting cellular and humoral immunity; 2) combined immunodeficiencies with associated or syndromic features; 3) predominantly antibody deficiencies, 4) diseases of immune dysregulation; 5) congenital defects of phagocyte number, function, or both; 6) defects in intrinsic and innate immunity; 7) autoinflammatory disorders; 8) complement deficiencies; and 9) phenocopies of PID. In the most recent update of the classification, each gene is arbitrarily assigned exclusively to one group (Picard et al., 2015). Nonetheless, in reality, *genetic pleiotropy*, the effect of one gene on multiple traits, and *genetic heterogeneity*, the existence of different genetic defects that cause the same disease, complicate the diagnostics of individuals with PIDs.

Since PIDs are monogenic disorders, they follow Mendelian inheritance patterns. Autosomal dominant (AD) and recessive (AR), as well as X-linked (XL) PIDs are known. Most PID-causing variants are inactivating or *loss-of-function* (LOF) variants that lead to diminished (hypomorph) or absent function (null) of the gene product. In AD disorders, LOF variants can cause disease if the gene dosage of only one healthy allele is insufficient (*haploinsufficiency*). Alternately, the mutated allele might interfere with the healthy one (*dominant-negative effect*). Interestingly, it has been recently recognized that also *gain-of-function* (GOF), *i.e.*, activating variants, can cause AD PIDs, often resulting in autoinflammation and autoimmunity (Boisson et al., 2015). Phenotypes are usually different when GOF and LOF disease-causing variants strike the same gene (Boisson et al., 2015), thus suggesting that both little and a lot of important immune proteins can lead to disease.

The introduction of HTS has changed the diagnostic approach to all Mendelian disorders including PIDs. A genotype-first approach, where genetic diagnostics precede other clinical tests, is becoming more and more common. Following the decline in costs for DNA-sequencing, gene panels and later exome sequencing (ES) and whole-genome sequencing (WGS) have made their entry in diagnostic and research settings (Meyts et al., 2016). Methodological considerations on the different strategies will be made in the methods section. Diagnostic rates of 15% to 25% have been reported for PIDs gene panels (Nijman et al., 2014; Al-Mousa et al., 2016). Stray-Pedersen *et al.* reported a diagnostic rate of 40% after studying 278 PID families by ES, including CNV analysis (Stray-Pedersen et al., 2017). However, the diagnostic rate varied largely across different PID subgroups with the least success achieved for autoinflammatory disorders. ES and WGS have also heavily contributed to the discovery of new PID genes during the last few years (Meyts et al., 2016).

Nonetheless, traditional diagnostics based on immunophenotyping and functional assays remains a very important tool in PID diagnostics. Immunological assays are still faster than most genetics approaches. They can aid interpretation and validation of genetic findings and may be the only choice in countries where DNA-sequencing is not readily available.

Early diagnosis is desirable for all PIDs. For PIDs that constitute real medical emergencies, quick diagnosis is mandatory. For instance, children born with severe combined immunodeficiency (SCID) have a much higher chance to survive if they undergo hematopoietic stem cell transplantation (HSCT) or gene therapy within the first months of life (Gennery et al., 2010). The impairment of T cell development characteristic of SCID can be measured through the quantification of T cell receptor excision circles (TRECs) produced during TCR rearrangement. Neonatal screening programs to measure TRECs in blood spots have been implemented in many states in the United States and are underway in several European countries (Buelow et al., 2014; Kwan et al., 2014). Rapid DNA-sequencing following a positive neonatal screening has been successful in identifying the underlying genetic defect (Yu et al., 2016).

Treatment choices for PIDs depend on the phenotypic manifestations of disease. Curative approaches are typically used for the more severe disorders and rely on HSCT and gene therapy. HSCT has been in practice for over 40 years and its success has relied on the availability of suitable donors, preferably HLA-matched siblings (Kang and Gennery, 2014). Gene therapy for PID can be seen as a form of autologous HSCT where hematopoietic stem cells are corrected *ex-vivo* and then reintroduced back into the affected individual. Initial efforts of gene therapy have been based on the gene addition method, where a normal copy of the gene is introduced through a viral vector. Discovery of gene editing tools such as CRISPR/Cas9 makes gene editing an appealing approach even in PIDs. Moreover, the increased molecular knowledge of PIDs has also empowered the use and development of biological drugs that specifically target the disease-causing mechanisms (Notarangelo and Fleisher, 2017).

Individuals with PIDs can be seen as “human knockouts” of key immune proteins. Thus, studies of PIDs can provide key insight into the functioning of the human immune system. Genes coding for essential proteins will be under stronger purifying selection compared to those for redundant proteins. Among the around 2000 genes associated with immune function by gene ontology, more than half are associated with innate immune response. Moreover, a higher number of adaptive immune genes are linked to PIDs compared to innate immune genes. This suggests that a high level of redundancy is present among innate host defense pathways (Fischer and Rausell, 2016). Interestingly, among GOF disorders, innate immune genes are more represented, suggesting that while redundancy may buffer the effect of LOF variants, an increased activity of innate immune genes can still lead to disease (Fischer and Rausell, 2016). Data from population genetics also confirms that innate immune genes show less depletion of LOF variants compared to adaptive immune genes (Fischer and Rausell,

2016). However, genetic differences exist even among innate immune genes. For example, intracellular TLRs that sense nucleic acid display a stronger purifying selection compared to cell-surface TLRs (Quintana-Murci and Clark, 2013).

1.3.2 Diseases of immune dysregulation

Diseases of immune dysregulation are a subgroup of PIDs characterized by severe autoimmunity or inflammation due to disrupted immune responses. In fact, a deregulated immune system can stem from impaired self/non-self discrimination and from the inability to correctly extinguish the immune response upon elimination of infections agents. Impaired tolerance leads to autoimmunity and improperly extinguished immune response to persistent inflammation and tissue damage. While autoimmunity and inflammation are key characteristics of this subgroup of PIDs, many other PIDs, especially T cell immunodeficiencies and common variable immunodeficiencies (CVID) also manifest with autoimmunity and inflammation. In general, individuals with PIDs have a higher risk to develop autoimmune cytopenias, inflammatory bowel diseases, and arthritis than the general population (Fischer et al., 2017).

Based on the 2015 update of the IUIS classification of PIDs, diseases of immune dysregulation are categorized into (Picard et al., 2015): 1) familial hemophagocytic lymphohistiocytosis syndromes; 2) T regulatory cells genetic defects; 3) autoimmunity with or without lymphoproliferation; 4) immune dysregulation with colitis; and 5) type I interferonopathies. An overview of these disorders is provided in Table 2 and 3.

1.3.2.1 Hemophagocytic lymphohistiocytosis

Hemophagocytic lymphohistiocytosis (HLH) is a severe hyperinflammatory syndrome due to a dysregulated immune response. HLH owes its name to two observations made early on histological specimens from individuals with HLH: *hemophagocytosis*, the engulfment of blood cells by activated macrophages in bone marrow, lymph node, liver, or spleen, and *lymphohistiocytosis*, i.e. lymphocyte infiltration in tissues. The first report of the disease in 1952 describes widespread histiocytic infiltration and hemophagocytosis in spleen, liver, and lymph nodes in autopsy material from two siblings who succumbed to a disease characterized by fever, hepatosplenomegaly, and cytopenias (Farquhar and Claireaux, 1952). From this first report, our understanding of pathogenic mechanisms and genetics of HLH has improved greatly, thanks to the commitment of dedicated physicians and scientists. Even today the occurrence of unremitting fever, hepatosplenomegaly and cytopenias in young children should trigger the clinical suspicion of HLH. However, better diagnostic and treatment strategies exist.

Disease	Gene	Protein role	Inheritance	Functional defect	Additional clinical features besides HLH	Phenotype OMIM
Perforin deficiency (FHL2)	<i>PRF1</i>	cytolytic protein, creates pore in target cell membrane	AR	Decreased to absent NK cell cytotoxicity and intracellular perforin expression		603553
Munc13-4 deficiency (FHL3)	<i>UNC13D</i>	prime vesicles for fusion	AR	Decreased to absent NK cell cytotoxicity and NK and CTL degranulation		608898
Syntaxin 11 deficiency (FHL4)	<i>STX11</i>	secretory vesicle fusion with the cell membrane	AR	Decreased to absent NK cell cytotoxicity and NK and CTL degranulation		603552
Munc18-2 deficiency (FHL5)	<i>STXBP2</i>	secretory vesicle fusion with the cell membrane	AR	Decreased to absent NK cell cytotoxicity and NK and CTL degranulation		613101
Chediak-Higashi syndrome (CHS)	<i>LYST</i>	lysosomal trafficking	AR	Decreased to absent NK cell cytotoxicity and NK and CTL degranulation	Partial albinism, giant inclusions in neutrophils on blood smear	214500
Griselli syndrome type 2 (GS2)	<i>RAB27A</i>	GTPase that promotes docking of secretory vesicles to the cell membrane	AR	Decreased to absent NK cell cytotoxicity and NK and CTL degranulation	Partial albinism	607624
SH2D1A deficiency (XLP1)	<i>SH2D1A</i>	adaptor protein regulating intracellular signaling	XL	No functional defect with HLH diagnostic assays	EBV infection, lymphoma	308240
XIAP deficiency (XLP2)	<i>XIAP</i>	inhibitor of apoptosis	XL	No functional defect with HLH diagnostic assays	EBV infection, inflammatory bowel disease	300635

Table 2. Overview of primary HLH disorders.

Disease	Gene	Clinical features	Phenotype OMIM
2. T regulatory cells genetic defects			
IPEX, immune dysregulation, polyendocrinopathy, enteropathy X-linked	<i>FOXP3</i>	Autoimmune enteropathy, early onset diabetes, thyroiditis hemolytic anemia, thrombocytopenia, eczema. Elevated IgE, IgA	304790
CD25 deficiency	<i>IL2RA</i>	Lymphoproliferation, autoimmunity. Impaired T cell proliferation.	606367
CTLA4 deficiency (ALPSV)	<i>CTLA4</i>	Autoimmune cytopenias, enteropathy, interstitial lung disease, extra-lymphoid lymphocytic infiltration recurrent infections.	616100
STAT3 GOF mutations	<i>STAT3</i>	Lymphoproliferation, solid organ autoimmunity, recurrent infections.	615952
3. Autoimmunity with or without lymphoproliferation			
APECED (APS-1), autoimmune polyendocrinopathy with candidiasis and ectodermal dystrophy	<i>AIRE</i>	Autoimmunity: hypoparathyroidism hypothyroidism, adrenal insufficiency, diabetes, gonadal dysfunction and other endocrine abnormalities, chronic mucocutaneous candidiasis, dental enamel hypoplasia, alopecia areata. Enteropathy, pernicious anemia.	240300
ITCH deficiency	<i>ITCH</i>	Early-onset chronic lung disease (interstitial pneumonitis). Autoimmune disorder (thyroiditis, type I diabetes, chronic diarrhea/enteropathy, and hepatitis). Failure to thrive, developmental delay, dysmorphic facial features.	613385
Tripeptidyl-Peptidase II Deficiency	<i>TPP2</i>	Variable lymphoproliferation, severe autoimmune cytopenias, hypergammaglobulinemia, recurrent infections.	Not yet assigned
3. Autoimmune lymphoproliferative syndrome (ALPS)			
ALPS-FAS	<i>TNFRSF6</i>	Splenomegaly, adenopathies, autoimmune cytopenias, increased lymphoma risk. IgG and A normal or increased. Elevated FasL and IL-10, vitamin B12.	601859
ALPS-FASLG	<i>TNFSF6</i>	Splenomegaly, adenopathies, autoimmune cytopenias, SLE. Soluble FasL is not elevated.	601859
ALPS-Caspase10	<i>CASP10</i>	Adenopathies, splenomegaly, autoimmunity.	603909
ALPS-Caspase 8	<i>CASP8</i>	Adenopathies, splenomegaly, bacterial and viral infections, hypogammaglobulinemia.	607271
FADD deficiency	<i>FADD</i>	Functional hypersplenism, bacterial and viral infections, recurrent episodes of encephalopathy and liver dysfunction.	613759
PRKC delta deficiency	<i>PRKCD</i>	Recurrent infections, EBV chronic infection, lymphoproliferation, SLE-like autoimmunity (nephrotic and antiphospholipid syndromes), hypogG.	615559
4. Immune dysregulation with colitis			
IL-10 deficiency	<i>IL10</i>	Inflammatory bowel disease (IBD), folliculitis, recurrent respiratory diseases, arthritis.	Not yet assigned
IL-10R α deficiency	<i>IL10RA</i>	IBD, folliculitis, recurrent respiratory diseases, arthritis, lymphoma.	613148
IL-10R β deficiency	<i>IL10RB</i>	IBD, folliculitis, recurrent respiratory diseases, arthritis, lymphoma.	612567
NFAT5 haploinsufficiency	<i>NFAT5</i>	IBD, recurrent sinopulmonary infections.	Not yet assigned
5. Type 1 Interferonopathies			
TREX1 deficiency, Aicardi-Goutieres syndrome 1 (AGS1)	<i>TREX1</i>	Progressive encephalopathy intracranial calcifications, cerebral atrophy, leukodystrophy, hepato and splenomegaly (HSMG), thrombocytopenia, elevated hepatic transaminases, chronic cerebrospinal fluid (CSF) lymphocytosis.	225750
RNASEH2B deficiency, AGS2	<i>RNASEH2B</i>	Progressive encephalopathy intracranial calcifications, cerebral atrophy, leukodystrophy, HSMG, thrombocytopenia, elevated hepatic transaminases, CSF lymphocytosis.	610181
RNASEH2C deficiency, AGS3	<i>RNASEH2C</i>	Progressive encephalopathy intracranial calcifications, cerebral atrophy, leukodystrophy, HSMG, thrombocytopenia, elevated hepatic transaminases, CSF lymphocytosis.	610329
RNASEH2A deficiency, AGS4	<i>RNASEH2A</i>	Progressive encephalopathy intracranial calcifications, cerebral atrophy, leukodystrophy, HSMG, thrombocytopenia, elevated hepatic transaminases, CSF lymphocytosis.	610333
SAMHD1 deficiency, AGS5	<i>SAMHD1</i>	Progressive encephalopathy intracranial calcifications, cerebral atrophy, leukodystrophy, HSMG, thrombocytopenia, anemia elevated lactates, chronic CSF lymphocytosis, skin vasculitis, mouth ulcers, arthropathy.	612952
ADAR1 deficiency, AGS6	<i>ADAR1</i>	Progressive encephalopathy intracranial calcification, severe developmental delay, leukodystrophy.	615010
Aicardi-Goutieres syndrome 7 (AGS7)	<i>IFIH1</i>	Progressive encephalopathy intracranial calcification, severe developmental delay, leukodystrophy.	615846
Spondyloenchondro-dysplasia with immune dysregulation	<i>ACP5</i>	Recurrent bacterial and viral infections, intracranial calcification, SLE-like autoimmunity (Sjögren's syndrome, hypothyroidism, inflammatory myositis, Raynaud's disease and vitiligo), hemolytic anemia, thrombocytopenia, skeletal dysplasia, short stature.	607944
STING-associated vasculopathy, infantile-onset	<i>TMEM173</i>	Severe infantile-onset autoinflammatory vasculopathy.	615934
ADA2 deficiency	<i>CECR1</i>	Polyarteritis nodosa, childhood-onset, early-onset recurrent ischemic stroke and fever.	615688

Table 3. Overview of other diseases of immune dysregulation. The table is adapted from (Picard et al., 2015).

Etiologically, HLH is divided in primary HLH, also called familial HLH (FHL), and secondary HLH. Secondary HLH develops secondary to pre-existing infections, malignancy, and autoimmune diseases in the absence of a known genetic defect. A broad range of conditions can promote development of secondary HLH, but herpes virus infections, especially Epstein-Barr virus (EBV), and hematological malignancies are among the most common triggers (Ramos-Casals et al., 2014). Primary and secondary HLH overlap clinically. Since the onset of primary HLH is also often triggered by infections, it is difficult to distinguish primary from secondary forms of HLH (Henter et al., 1993; Janka, 2012).

Initially regarded as a pediatric disease, HLH is increasingly recognized as a disease in adults as well, especially in its secondary form. The incidence of primary HLH in the pediatric population is estimated to be 1 in 50,000 live births in Sweden (Henter et al., 1991a; Meeths et al., 2015). This is similar to the incidence of SCID as detected by neonatal screening (Kwan et al., 2014), suggesting that primary HLH might be even more common than SCID. In the majority of FHL cases, HLH will develop within the first year of age (Trottestam et al., 2011). Although the incidence of primary and secondary HLH in adults is unknown, about 1% of individuals with hematological malignancies will develop HLH (Machaczka et al., 2011).

Genetics of primary HLH

The first gene associated to familial HLH was *PRF1*, which codes for perforin (Stepp et al., 1999). Perforin is a protein contained in secretory lysosomes, also known as cytotoxic granules, of NK cells and CTLs. Upon exocytosis, perforin creates pores in the target cell membrane, thus enabling entry of granzymes (Law et al., 2010). The discovery of *PRF1* pathogenic variants in FHL patients confirmed the hypothesis of Fadeel *et al.* that FHL might be due to a defect in the perforin-granzyme system (Fadeel et al., 1999), and provided the explanation for the previously observed defective lymphocyte cytotoxicity in FHL patients (Egeler et al., 1996). Subsequent genetic studies identified biallelic pathogenic variants in *UNC13D*, *STX11*, and *STXBP2* as additional causes of FHL (Feldmann et al., 2003; zur Stadt et al., 2005; Côte et al., 2009; zur Stadt et al., 2009). These genes encode, respectively, for Munc13-4, syntaxin-11, and Munc18-2. While perforin is a cytotoxic granule constituent, the other FHL proteins are involved at different steps of cytotoxic granule exocytosis (Figure 2) (de Saint Basile et al., 2010). Therefore, genetic findings have defined FHL as an autosomal recessive defect of perforin-mediated lymphocyte cytotoxicity.

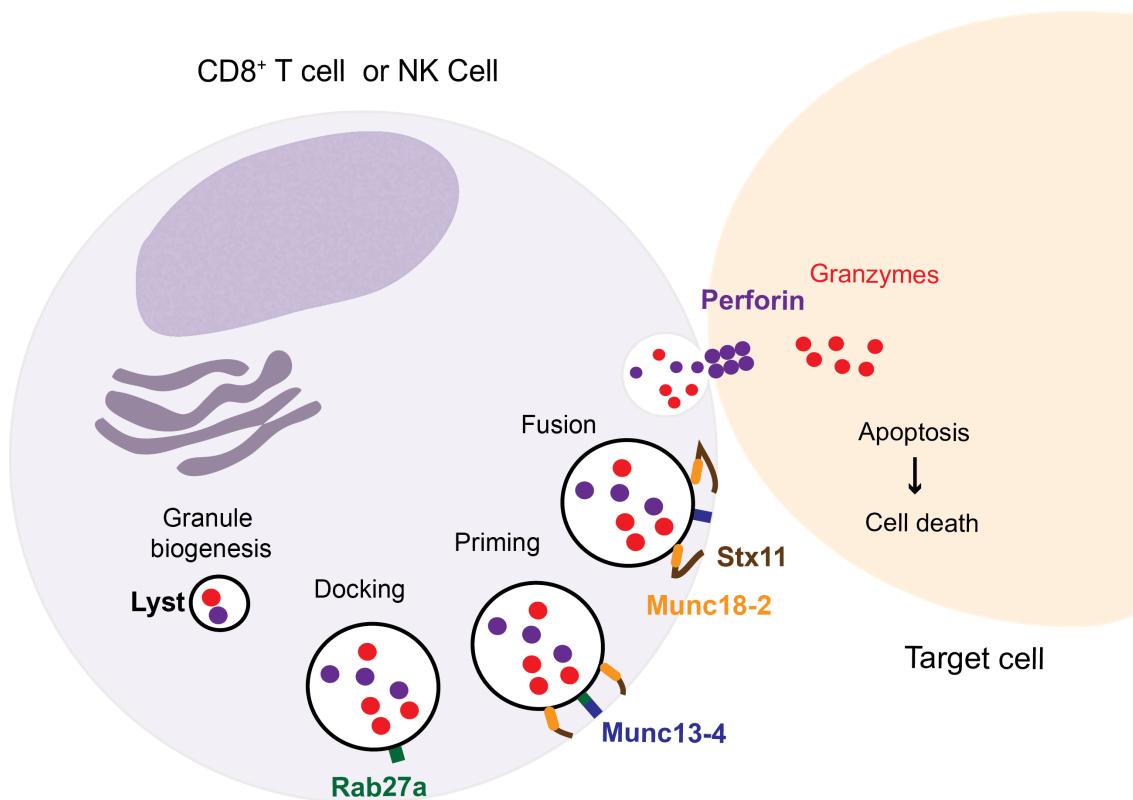


Figure 2. Simplified representation of the steps required for proper exocytosis of cytotoxic granules by cytotoxic lymphocytes with focus on the proteins that can be defective in primary HLH.

PRF1 and *UNC13D*, causative of FHL2 and FHL3, respectively, are the most commonly mutated genes in primary HLH, but geographical and ethnic differences exist. HLH-causing variants in *STX11* are usually confined to Turkish patients (Sieni et al., 2014). Among 171 Italian FHL patients, 37%, 33%, and 5% respectively carried biallelic disease-causing variants in *PRF1*, *UNC13D*, and *STXBP2* (Cetica et al., 2016). Disease-causing variants in FHL genes are LOF and this is reflected in a mutation spectrum that includes nonsense, frameshift, splice-site, and missense variants. While a growing number of different disease-causing variants are being reported, some are more common due to population founder effects. For instance, in Sweden, a 253kb inversion in *UNC13D* is the single most common FHL-causing aberration, especially in patients from Northern Sweden (Meeths et al., 2011). Disease-causing variants in an intronic, evolutionarily conserved region of *UNC13D* have also been identified (Meeths et al., 2011; Entesarian et al., 2013; Seo et al., 2013). This region represents a gene enhancer and an alternative transcriptional start site (Cichocki et al., 2014).

Primary HLH is also observed in two syndromes, Griscelli syndrome type 2 (GS2) and Chediak-Higashi syndrome (CHS), which present with partial albinism in addition to HLH. GS2 and CHS are caused by biallelic pathogenic variants in the genes *RAB27A* and *LYST* respectively. Even GS2 and CHS patients display defective lymphocyte cytotoxicity since Rab27a and LYST are also involved in trafficking and exocytosis of cytotoxic granules (Figure 2) (de Saint Basile et al., 2010). The co-occurrence of partial albinism is due to a

shared lysosomal machinery between melanocytes and cytotoxic lymphocytes. Furthermore, patients with X-linked lymphoproliferative syndromes type 1 (XLP1) and type 2 (XLP2) caused by pathogenic variants in *SH2D1A* and *XIAP* respectively, may present with HLH, especially after EBV infection. Besides severe EBV infection and HLH, XLP2 patients are at risk of inflammatory bowel disease, while a third of individuals with XLP1 develop lymphoma (Booth et al., 2010; Speckmann et al., 2013). Patients with other primary immunodeficiencies may also present with HLH (Bode et al., 2015).

Pathogenesis and clinical manifestations of HLH

Defective lymphocyte cytotoxicity is a hallmark of HLH. CD8⁺ T cells and NK cells kill virus-infected cells and are important to turn off the immune response by killing activated immune cells. Defective lymphocyte cytotoxicity leads therefore to poor control of viral infections, since infected cells cannot be eradicated. Simultaneously, cytotoxic lymphocytes continue to proliferate and produce cytokines, stimulating inflammatory response and failing to kill activated immune cells. Mouse models of primary HLH have suggested a crucial role of IFN- γ in the development of HLH (Jordan, 2004). *Prf1* knockout mice challenged with lymphocytic choriomeningitis virus (LCMV) infection develop fatal HLH-like disease that can be rescued by antibody-mediated IFN- γ neutralization (Schmid et al., 2009). The production of IFN- γ by CD8⁺ T cells and NK cells sustains macrophage activation and fuels inflammation. Sepulveda et al. showed that CD8⁺ T cells and NK cells have non-redundant roles in the pathogenesis of HLH in mice: CD8⁺ T cells are devoted to viral clearance and NK cells limit hyperactivation of CTLs and macrophage infiltration (Sepulveda et al., 2015).

The persistent activation and proliferation of CTLs, macrophages and histiocytes lead to the clinical picture of HLH. The clinical manifestations of HLH include non-remitting fever, hepatosplenomegaly, and cytopenias (Henter et al., 2007). Fever is caused by the heavy production of IL-1, IL-6, and tumor necrosis factor (TNF). TNF, IFN- γ , and the heavy subunit of ferritin suppress hematopoiesis, leading to cytopenia. Other abnormalities present in HLH are low fibrinogen and high levels of ferritin due, respectively, to macrophage secretion of plasminogen activator and ferritin. Inhibition of lipoprotein lipase stimulates triglyceride synthesis and causes hypertriglyceridemia (Henter et al., 1991c). Hemophagocytosis and tissue infiltration result from persistent macrophage activation (Henter et al., 1991b; Janka, 2012). Further immunological features include elevated IL-2 receptor (sIL2R or soluble-CD25) and impaired lymphocyte cytotoxicity. Neurological involvement is reported in 37–63% of HLH patients and can lead to severe long-term sequelae in HLH survivors (Horne et al., 2008; Deiva et al., 2012).

Immune dysregulation can also be caused by infections, malignant and autoimmune diseases without an underlying genetic defect, thus leading to secondary HLH, whose pathogenesis is less well understood. EBV-triggered HLH is thought to occur because of prolonged antigen stimulation which leads to a massive expansion of EBV-specific CTLs. EBV-specific CTLs

will eventually clear the infection, but the unavoidable immunopathology caused by CTLs might clinically manifest as HLH. This also explains why genetic susceptibility to EBV infection, such as in XLP, ITK deficiency (Huck et al., 2009), and CD27 deficiency (van Montfrans et al., 2012) is a risk factor for HLH (Parvaneh et al., 2013).

Diagnostics of HLH

Clinical and laboratory findings in HLH patients result from severe hyperinflammation and are therefore common in other hyperinflammatory states such as sepsis. To facilitate recognition of disease, diagnostic criteria exist (Table 4)(Henter et al., 2007). A diagnosis of HLH is established upon fulfillment of 5 out of 8 diagnostic criteria. Neutropenia occurs usually later and hemophagocytosis is not always present. Instead, the level of ferritin and thrombocytes can be used as prognostic markers and to monitor therapy (Janka and Lehmberg, 2014). It is difficult to distinguish primary and secondary HLH based on conventional clinical parameters, yet primary cases may display higher total lymphocytes and a high sIL2R/ferritin ratio (Yasumi et al., 2015).

Defective NK cell cytotoxicity is also one of the diagnostic criteria (Henter et al., 2007). Although widely used, this assay is sensitive to amounts of circulating NK cells. Since the amount of NK cells is often low in secondary HLH the NK cell cytotoxicity assay cannot readily distinguish primary from secondary HLH (Janka and Lehmberg, 2014). A more refined laboratory diagnostic algorithm has been proposed for HLH patients (Bryceson et al., 2012). This algorithm includes assessment of intracellular expression of perforin and NK cell degranulation in addition to NK cell cytotoxicity, thus being able to reliably distinguish primary from secondary forms and to guide genetic analysis for primary HLH. These methods will be discussed further in the methods section of this thesis. A CTL degranulation assay in response to CD3, based on the finding that CD8⁺CD57⁺ T cells contain high amounts of cytotoxic granules (Chiang et al., 2013), has also been introduced in the diagnostics of HLH. The CD8⁺CD57⁺ T cells degranulation assay showed very good signal-to-noise ratio and better sensitivity and specificity than the NK cell degranulation assay (Chiang, 2015). The improvement achieved in diagnostic immunological assays for HLH patients may call for a revision of the diagnostic criteria.

A molecular diagnosis consistent with HLH is an independent criterion to establish a diagnosis of primary HLH. Since treatment recommendations may differ between primary and secondary HLH, rapid molecular studies can complement immunological investigations and have a direct impact on clinical management. A molecular diagnosis enables also a more appropriate genetic counseling, the identification of individuals at risk, and prenatal and preimplantation genetic diagnostics. The suspicion of primary HLH is strong in case of a positive family history, parental consanguinity, recurrent disease, the co-occurrence with partial albinism, and other clinical manifestations specific to certain HLH syndromes and, as previously discussed, the detection of an immunological defect consistent with primary HLH.

However, the genetic heterogeneity of HLH makes the search for a molecular defect difficult and time-consuming. This thesis addresses this challenge by leveraging HTS for the diagnostics of patients with HLH.

A diagnosis of HLH can be established upon:
(A) a molecular finding consistent with HLH
(B) Fulfillment of five out of the eight criteria below:
1. Unremitting fever
2. Splenomegaly
3. Cytopenias (affecting ≥ 2 of 3 lineages in the peripheral blood):
• Hemoglobin <90 g/L
• Platelets $<100 \times 10^9$ /L
• Neutrophils $<1.0 \times 10^9$ /L
4. Hypertriglyceridemia and/or hypofibrinogenemia:
• Fasting triglycerides ≥ 3.0 mmol/L (i.e., ≥ 265 mg/dl)
• Fibrinogen ≤ 1.5 g/L
5. Hemophagocytosis in bone marrow or spleen or lymph nodes
No evidence of malignancy
6. Ferritin ≥ 500 mg/L
7. Soluble CD25 (i.e., soluble IL-2 receptor) $\geq 2,400$ U/ml
8. Low or absent NK-cell activity (according to local laboratory reference)

Table 4. HLH-2004 diagnostic criteria (Henter et al., 2007).

Treatment of HLH

Untreated FHL patients survive only about 2 months on average (Henter et al., 1991a). A prompt diagnosis and initiation of therapy is therefore important. The first international protocol for HLH treatment was developed by the Histiocyte Society in 1994 and called HLH-94 (Henter et al., 1997). HLH-94 was based on an induction regimen with etoposide and dexamethasone, followed by a maintenance therapy that would bridge into HSCT for primary cases. Usually, patients with acquired disease can be put off treatment if in remission after the 8 weeks of induction therapy. Depending on the severity and underlying cause of the secondary HLH treatment can often be modified and made less intensive or shorter. HLH-94 managed to raise the overall 5-year survival from 4% (Janka, 1983) to 54% (Trottestam et al., 2011). Etoposide, a topoisomerase II inhibitor widely used as chemotherapeutic drug, was shown to promote depletion of activated T cells and suppress inflammatory cytokine production in perforin-knockout mice (Johnson et al., 2014). In the subsequent protocol, HLH-2004, cyclosporin A was included in the induction regimen. Results from this study confirm the effectiveness of the etoposide/dexamethasone combination, while the addition of cyclosporin A upfront did not significantly improve outcome (Bergsten et al., 2017). Similar results have been obtained from a single center study based on a combination of antithymocyte globulins (ATG), corticosteroids, cyclosporin A, and intrathecal methotrexate

(Mahlaoui et al., 2007). To further assess the benefit of a combined immuno-chemotherapy, two hybrid trials, hybrid immunotherapy for HLH (HIT)-HLH and Euro-HIT, which utilize ATG, etoposide, and dexamethasone, have been carried out and results are pending. One ongoing trial aims to evaluate the use of alemtuzumab, a monoclonal antibody against CD52, as first-line therapy, after promising results as salvage therapy (Marsh et al., 2013). Based on mouse models of HLH, a clinical trial is open for the evaluation of an anti-IFN γ monoclonal antibody (NI-0501). The JAK1/2 inhibitor ruxolitinib has shown promising results in mice (Das et al., 2016; Maschalidi et al., 2016), and there is currently a clinical trial open for treatment of secondary HLH. As of today, also based on the recently published data from the HLH-2004 study, the recommendation of the Histiocytose Society is to use the HLH-94 as first line treatment in patients fulfilling HLH-2004 diagnostic criteria (Henter et al., 1997, 2007).

Neurologic disease in HLH is a major concern for long-term sequelae. Currently, intrathecal methotrexate and corticosteroids are widely used in patients with neurologic involvement (Horne et al., 2017).

HSCT is currently the only cure for primary HLH. In the HLH-94 study, around half of the patients underwent HSCT (Trottestam et al., 2011). In this cohort, the 3-year overall survival was 64%, with no significant difference between matched related and matched unrelated donors (Horne et al., 2005). A reduced intensity conditioning carries less toxicity and better survival (Marsh et al., 2010), at the cost of mixed chimerism and higher risk of reactivation. Interestingly, both in perforin-knockout mice and in humans, donor chimerism >20-30% was enough to protect against reactivation (Terrell and Jordan, 2013; Hartz et al., 2016). Another curative option might be gene therapy. Used in other PIDs, it would be exciting to see the current development in gene editing translating into an opportunity for gene therapy for this group of patients.

1.3.2.2 Other diseases of immune dysregulation

A summary of all diseases of immune dysregulation besides primary HLH according to IUIS 2015 classification of PIDs is found in Table 3. Here below, I briefly introduce defects in *CTLA4/LRBA* and type I interferonopathies which are related to some of the diseases studied in this thesis.

Cytotoxic lymphocyte antigen-4 (CTLA-4) is an immune checkpoint receptor that down-regulates T cell activation (Tivol et al., 1995; Krummel and Allison, 1996). It is expressed constitutively by Treg-cells and transiently on activated T cells. It competes with CD28, a co-stimulatory protein, for binding to B7 molecules, and is important for maintenance of peripheral tolerance and prevention of autoimmunity. Recently, heterozygous pathogenic variants in *CTLA4* leading to haploinsufficiency of the receptor have been reported in individuals with multiple autoimmune manifestations, such as immune thrombocytopenic purpura, autoimmune hemolytic anemia, enteropathy, interstitial lung disease and others

(Kuehn et al., 2014; Schubert et al., 2014). The clinical presentation of CTLA4 deficiency resembles the one of LPS-responsive and beige-like anchor (LRBA) deficiency, an autosomal recessive PID first described in 2012 among young patients with severe CVID and multiple autoimmune manifestations (Lopez-Herrera et al., 2012). LRBA deficiency is classified as a combined immunodeficiency according to IUIS, yet the pathogenic mechanisms overlap with CTLA4 deficiency. In fact, LRBA co-localizes with CTLA4 in recycling endosomes and regulates the lysosomal degradation of CTLA4 (Lo et al., 2015). Lack of LRBA due to biallelic LOF variants leads to degradation of CTLA4. The reduced expression of CTLA4 on Treg and activated T cells leads in turn to autoimmunity (Lo et al., 2015). Almost all patients with LRBA deficiency display autoimmune manifestations, sometimes even resembling immunedysregulation polyendocrinopathy enteropathy X-linked (IPEX) syndrome, which is due to pathogenic variants in the transcription factor *FOXP3* and causes lack of Treg cells (Sakaguchi et al., 2010).

The concept of type I interferonopathies as a new subgroup of inborn errors of immunity was introduced in 2011 by Yanick Crow (Crow, 2011). Mendelian type I interferonopathies are defined as diseases caused by an improper stimulation of type I IFN response, or by a defect in its negative regulation (Crow, 2011). IFNs are proteins produced in response to nucleic acids detection by infected cells, through cytoplasmatic PRRs, and by innate immune cells, through endosomal TLRs. IFNs induce an antiviral state and also stimulate innate and adaptive immune responses (Ivashkiv and Donlin, 2014). Despite clinical and genetic heterogeneity, central nervous system (CNS) and skin involvement are common. Autoimmunity is also a common presentation, for example in the form of a systemic lupus erythematosus (SLE) phenotype. A specific “IFN signature” characterized by increased expression of interferon stimulated genes can be found in the blood of these patients (Crow and Manel, 2015).

1.4 PREDISPOSITION TO HEMATOLOGICAL MALIGNANCIES

Cancer is the most common genetic disease. Every cancer originates from genetic alterations, defined as *somatic*, *i.e.* acquired in non-germ cells. Most of the acquired genetic alterations do not cause any cellular changes. A somatic mutation in a tumor suppressor gene or an oncogene can instead initiate the process of clonal evolution. During this process cells iteratively acquire mutations and expand clonally, which may result into the development of clinically recognizable cancers. Cell and DNA replication invariably constitute a risk for our genome to develop cancer mutations. Not surprisingly, a high positive correlation has been shown between the lifetime risk for a cancer type and the total number of cell stem division in the original cell type (Tomasetti and Vogelstein, 2015; Tomasetti et al., 2017). This suggests a rather debated strong contribution of simple “bad luck” to the occurrence of most cancers. However, the effect of extrinsic factors on cancer development is large and makes prevention of certain cancer forms to a large extent possible (Wu et al., 2016).

Among all cancers, about 8% originate from cells of the bone marrow and lymphatic system. These, called hematological malignancies, include various forms of leukemia, lymphoma, and myeloma (Rodriguez-Abreu et al., 2007). In this section, the risk factors for hematological malignancies, and in particular myeloid neoplasms, will be introduced.

1.4.1 Risk factors

A combination of genetic and environmental factors contributes to the individual's risk to develop cancer. Certain extrinsic factors predisposing to hematological malignancies exist. For instance, exposure to ionizing radiation predisposes to acute leukemia and to chronic myeloid leukemia. Moreover, cancer patients treated with chemotherapy and radiotherapy are at increased risk for "secondary" leukemia, especially myelodysplasia and acute myeloid leukemia (AML). Certain infections, like from human T cell leukemia virus (HTLV) type-1 and type-2 and EBV among others, are associated with specific hematological malignancies (Rodriguez-Abreu et al., 2007).

Age also has profound effects on hematopoiesis, and as previously discussed, on immunity. The production of mature blood cells during a lifetime requires continuous cell divisions, with the inevitable risk that somatic mutations might occur in hematopoietic progenitor and stem cells. This is also reflected in a higher incidence of some hematological malignancies, especially of myeloid origin, in older individuals (Pang et al., 2017). With age, the number of hematopoietic stem cells (HSCs) increases at the cost of a reduced diversity and of a myeloid bias. The occurrence of somatic mutations in genes important for HSCs can lead to the clonal expansion of mutated HSCs, a phenomenon known as *clonal hematopoiesis*.

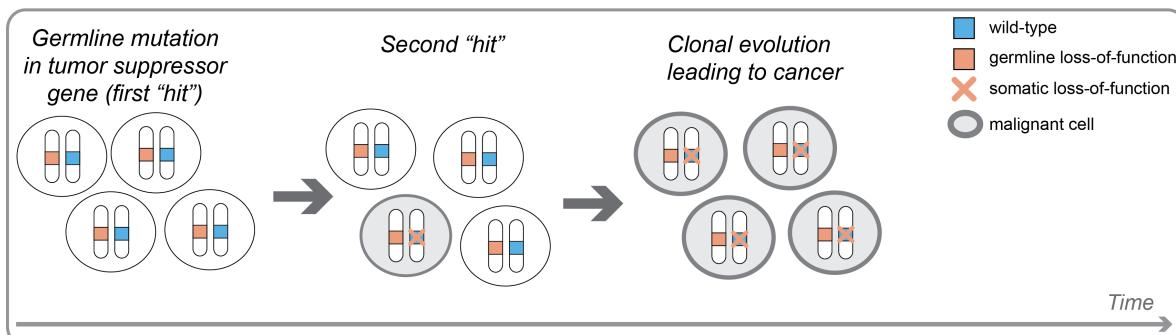
Clonal hematopoiesis was first identified by analysis of X-chromosome inactivation (Busque et al., 1996). Recently, genetic studies of large populations have revealed that somatic chromosomal aberrations and mutations in hematopoietic cells leading to clonal hematopoiesis are commonly acquired with age (Jacobs et al., 2012; Laurie et al., 2012; Genovese et al., 2014; Jaiswal et al., 2014). The term clonal hematopoiesis of indeterminate potential (CHIP) was proposed to describe the occurrence of clonal hematopoiesis in individuals who do not fulfill other diagnostic criteria for a hematological malignancy (Steensma et al., 2015). The majority of individuals with CHIP will not develop a hematological malignancy. However, CHIP confers a 10-fold increased risk for hematological malignancy (Jaiswal et al., 2014). Notably, CHIP is also associated with increased risk of coronary heart disease and all-cause mortality (Jaiswal et al., 2014, 2017).

1.4.2 Inherited myeloid malignancies

It is estimated that about 5-10% of all cancers are due to genetic predisposition. In fact, about 10% of cancer patients harbor germline pathogenic variants compatible with a cancer predisposition syndrome (Zhang et al., 2015; Schrader et al., 2016). Genetic predisposition to cancer most commonly implies the presence of a heterozygous germline pathogenic variant in

a tumor suppressor gene, *e.g.* *RB1* in retinoblastoma and *TP53* in Li-Fraumeni syndrome (Knudson, 1971; Malkin et al., 1990). Germline predisposing variants represent a “shortcut” to the process of clonal evolution, as captured by Knudson in 1971 in his *two-hit hypothesis* of cancer predisposition (Figure 3), based on observation from hereditary and sporadic retinoblastoma (Knudson, 1971). This “shortcut” effect is usually reflected in a younger age at cancer diagnosis for individuals with a genetic predisposition to cancer.

A) Hereditary cancer



B) Sporadic cancer

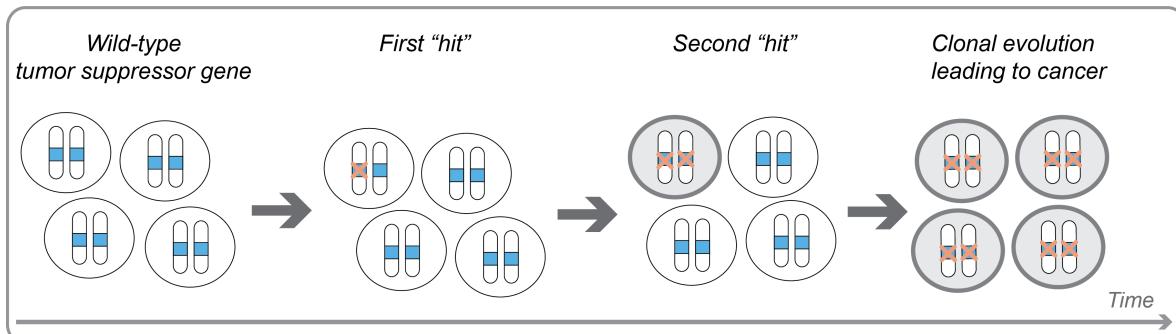


Figure 3. The “two hit” hypothesis. Individuals with inherited predisposition to cancer are born with a pathogenic variant in, usually, a tumor suppressor gene. A second somatic hit on the other allele can lead to clonal expansion and cancer development. In sporadic cancers, two somatic hit in the same gene in the same cell need to occur for cancer to develop.

Hematological malignancies can be associated with genetic syndromes predisposing to a broader group of cancers, as in Li-Fraumeni syndrome, but they may also be the primary oncologic manifestation of a genetic condition. Moreover, specific genetic syndromes, such as Down syndrome, Fanconi anemia, and others, display an increased risk for certain hematological malignancies (McGee and Nichols, 2016). Hematological malignancies, in particular lymphomas, are common in patients with PIDs (Leechawengwongs and Shearer, 2012), which confirms an important role of the immune system in recognizing and eradicating neoplastic and malignant cells (Hanahan and Weinberg, 2011).

Myelodysplastic syndrome (MDS) is a group of heterogeneous clonal disorders of hematopoietic stem and progenitor cells. MDS is characterized by inefficient hematopoiesis

and risk of progression to AML. The incidence of MDS increases with age, representing only 9% of pediatric hematological malignancies (Hasle et al., 1995; Adès et al., 2014). Even though more rare in younger age, the role for genetic predisposition in this group of patients might be larger. One recent study identified predisposing germline variants in about 5% and 13% of children and young adults that underwent HSCT for aplastic anemia or myelodysplastic syndrome, respectively (Keel et al., 2016). Inherited MDS may progress from pre-existent bone marrow failure syndromes caused by pathogenic variants in genes required for DNA-repair, chromosomal stability, and telomere elongation. Furthermore, several rare genetic causes of familial MDS/AML have recently been described, including autosomal-dominant *CEBPA*, *DDX41*, *RUNX1*, *ANKRD26*, *ETV6*, *SRP72*, and *GATA2* pathogenic variants (Godley and Shimamura, 2017). Disease-causing variants in *GATA2* cause a syndrome characterized by variable expressivity of immunodeficiency, susceptibility to mycobacterial infections, lymphedema as well as MDS/AML (Spinner et al., 2014). Notably, germline *GATA2* pathogenic variants explain 6% of pediatric MDS cases and up to 37% of pediatric MDS cases with monosomy 7, a cytogenetic aberration commonly associated with *GATA2*-related MDS (Włodarski et al., 2016). Altogether, pathogenic variants in known genes predisposing to MDS/AML explain only about 30% of familial cases (Churpek et al., 2015). Additional yet undiscovered genetic factors predisposing to MDS/AML may therefore exist.

In 2016, germline predisposition to myeloid malignancies, such as MDS and AML, was added to the World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues (Arber et al., 2016). This attests the recent development in the field of inherited myeloid malignancies and the need to incorporate this knowledge into clinical practice. In fact, the more and more widespread use of HTS in the clinical evaluation of patients with hematological malignancies might detect variants in genes such as *RUNX1*, *CEBPA*, and *GATA2*, that could be either somatic or germline and require follow-up (DiNardo et al., 2016). The detection of germline pathogenic variants usually triggers genetic counseling and family investigation. Yet, prospective data is sparse and currently only recommendations from experts are available for the management and follow-up of patients and carriers with such pathogenic variants (Godley and Shimamura, 2017).

1.5 A SNAPSHOT OF THE FIELD IN 2013

In 2013, when I started work on my thesis, the core genetic architecture of HLH had been defined. As reviewed in the introduction, pathogenic variants in *PRF1*, *UNC13D*, *STX11*, and *STXBP2* were implicated in FHL, and the genetic bases for XLP and syndromes with albinism and HLH were known (Janka, 2012). In 2011, all Swedish children diagnosed with HLH between December 2005 and January 2011 and with a functional defect consistent with primary HLH received a molecular diagnosis (Meeths et al., 2011). The focus of HLH research was shifting from classical forms to more atypical ones (Rohr et al., 2010; Zhang et al., 2011). The genetic bases of secondary HLH remained largely unknown. At the same

time, HTS had recently made its entry into the field of Mendelian disorders (Ng et al., 2010), and its application to diagnostics of HLH and in general to PIDs seemed promising to improve the diagnostics of these patients and aid gene discovery. In 2013, about 250 PID genes were known. GATA2 syndrome, characterized by PID and familial MDS, was one of the first PID to be discovered by ES (Dickinson et al., 2011; Hahn et al., 2011). LRBA-deficiency was reported first in 2012 in 5 patients with severe, early-onset B-cell deficiency and autoimmunity, which received symptomatic treatment (Lopez-Herrera et al., 2012).

2 AIMS OF THE THESIS

The overall aim of this thesis is to increase understanding of the genetic architecture of HLH and its phenotypic spectrum and to elucidate new genetic factors behind autoimmunity and hematological malignancies by studying families with rare diseases.

The specific aims are:

- to improve genetic diagnostics of HLH (**Papers I-IV**)
- to elucidate the phenotypic spectrum of HLH and the genetic mechanisms in patients with atypical and/or secondary HLH (**Papers II-IV**)
- to identify the genetic defect in a patient with severe autoimmune manifestations (**Paper V**)
- to describe, understand, and characterize the genetic mechanisms behind a familial form of MDS (**Paper VI**)
- to describe clinical and laboratory characteristics of all patients included in the thesis (**Papers I-VI**)

Most of the studies in this thesis have been performed prospectively. Genetic and immunological analyses of patients with HLH were carried out within the diagnostic algorithm for primary HLH. The results have been reported back to clinicians and were used and appreciated in the clinical care of these patients.

3 METHODOLOGICAL WORKFLOW

The journey that has led to the realization of this thesis began with the clinical suspicion of a genetic disease of the immune system by astute clinicians working in pediatric and adult hospital wards all over the world. Our part in this journey starts with the reception of a blood sample for genetic and immunological analyses, accompanied by consent for inclusion in research studies.

The genetic studies included in this thesis share a common methodological workflow based on the generation of genetic data, their analysis and interpretation, and the subsequent validation of findings. For **papers I-IV** that focus on HLH, immunological assays for diagnostics of primary HLH are an essential part of this workflow. This section will present the methods used in this thesis by grouping them according to the different steps of the methodological workflow (Figure 4). Immunological assays for HLH diagnostics are introduced at the end of this section. Further details on the methods used are available in the methods section of the respective papers.

The Regional Ethical Review Board in Stockholm, Sweden has approved all studies included in this thesis.

3.1 PATIENTS STUDIED

In total, 86 individuals from 75 unrelated families are included in this thesis. Individuals in **papers I-IV** were included based on fulfillment of at least five out of eight HLH-2004 diagnostic criteria (Henter et al., 2007). If clinical and laboratory criteria were not fulfilled at the time of sampling, a result from immunological assays suggesting primary HLH was considered sufficient for inclusion.

In **paper II** we performed a retrospective analysis of individuals previously diagnosed at our center with biallelic *PRF1* pathogenic variants associated with atypical phenotypes or intra-familial variable expressivity.

The families studied in **papers V-VI** were recruited because of an unexplained immunodeficiency, also associated with familial MDS in **paper VI**.

There is no overlap of study subjects between the different studies. The functional data of two patients with biallelic pathogenic variants in *PRF1* detected in **paper I** was used as the control cohort of typical FHL2 cases in **paper II**.

Genetic analyses were performed mostly on genomic DNA isolated from peripheral blood according to standard procedures. In **paper VI** other DNA sources were used, such as fibroblasts, dried blood spot, bone marrow, and sorted subpopulations of white blood cells.

3.2 GENERATION OF GENETIC DATA

Upon suspicion of a genetic disease, a genetic test was performed. In **papers I** and **III-VI** different HTS technologies and applications have been used to achieve a molecular diagnosis in the individuals studied.

	Generation of genetic data	Analysis and interpretation of genetic data	Validation of genetic findings
Paper I	<ul style="list-style-type: none"> ▪ Gene panel 12 HLH-related genes Ion Proton (PGM) 	<ul style="list-style-type: none"> ▪ Ion Torrent Suite for mapping & variant calling ▪ Custom variant filtering 	<ul style="list-style-type: none"> ▪ Segregation studies ▪ Review of phenotype ▪ HLH diagnostics
Paper II	<ul style="list-style-type: none"> ▪ Sanger sequencing of <i>PRF1</i> 	<ul style="list-style-type: none"> ▪ Retrospective study ▪ Genotype-phenotype correlation 	<ul style="list-style-type: none"> ▪ Segregation studies ▪ Review of phenotype ▪ HLH diagnostics
Paper III	<ul style="list-style-type: none"> ▪ ES of patient A (paired end Illumina seq) ▪ Sanger sequencing of HLH genes and <i>IFNGR1/2</i> in patient B 	<ul style="list-style-type: none"> ▪ ES analysis according to GATK best practices ▪ Custom variant filtering (AR, rare variants) 	<ul style="list-style-type: none"> ▪ Segregation studies ▪ Review of phenotype ▪ HLH diagnostics ▪ RNA studies and pSTAT1 analysis in patient A
Paper IV	<ul style="list-style-type: none"> ▪ WGS of P1 and P2 (paired end Illumina seq) 	<ul style="list-style-type: none"> ▪ WGS analysis with SpeedSeq ▪ Analysis of <i>RAB27A</i> SVs in IGV 	<ul style="list-style-type: none"> ▪ Segregation studies ▪ Review of phenotype ▪ HLH diagnostics ▪ Breakpoint PCR, MLPA ▪ qPCR <i>RAB27A/PIGB</i> ▪ Western blot RAB27A ▪ FANTOM CAGE data
Paper V	<ul style="list-style-type: none"> ▪ ES (paired end Illumina seq) 	<ul style="list-style-type: none"> ▪ ES analysis according to GATK best practices ▪ Custom variant filtering (AR, rare variants) 	<ul style="list-style-type: none"> ▪ Segregation studies ▪ Review of phenotype
Paper VI	<ul style="list-style-type: none"> ▪ ES on individuals from three generations (F1) (paired end Illumina seq) 	<ul style="list-style-type: none"> ▪ ES analysis according to GATK best practices ▪ Custom variant filtering (AD/XL, rare variants) 	<ul style="list-style-type: none"> ▪ Segregation studies ▪ Review of phenotype ▪ Sanger sequencing ▪ SNP arrays ▪ Cell sorting & dPCR ▪ Overexpression of wild-type and mutant SAMD9L in HEK 293T cells

Figure 4. Summary of methodological workflow for each paper. *HLH diagnostics* refers to the immunological assays described in section **3.5**.

3.2.1 HTS technologies

HTS, also known as *massively parallel sequencing*, encompasses different sequencing technologies whose common characteristic is the ability to simultaneously - *in parallel* - sequence millions of DNA fragments (Metzker, 2010). Although different in several aspects, HTS technologies share a common workflow based on library preparation, sequencing, and data processing.

Paper I employs Ion Torrent technology, while **papers III-VI** employ Illumina technology, which, as of today, dominates the scene of HTS. Both methods can be defined as sequencing-by-synthesis methods since they are DNA-polymerase dependent, yet they differ largely (Goodwin et al., 2016). They can be further distinguished in cyclic reversible termination (Illumina) and single-nucleotide addition (Ion Torrent) sequencing-by-synthesis methods.

After generation of DNA templates (libraries) by different means, Illumina instruments produce sequencing data through solid-phase bridge amplification, sequencing by cyclic reversible termination, and imaging. Briefly, the DNA templates are distributed and attached to a flow-cell on which they are amplified to obtain clusters containing copies of the original DNA templates. After amplification, as many sequencing cycles as the selected read length occur (from 50 to 300 bp depending on the sequencers). At each cycle all the differently fluorescently labeled reversible terminator-bound deoxyribonucleotides (dNTPs) are present and compete with each other for incorporation in the growing DNA strand by DNA polymerase. After incorporation of one nucleotide, unbound nucleotides are washed away. At this point, a total internal reflection fluorescence microscopy is used to record which base is incorporated at each cluster (Goodwin et al., 2016). The terminator molecule is then removed and a new cycle can begin. **Papers III-VI** are based on paired-end sequence data produced by Illumina sequencers HiSeq 2000, 2500, and X (Illumina Inc, San Diego, CA, USA).

Ion Torrent technology is described as a single-nucleotide addition method since each of the four nucleotides is added iteratively. In fact, this technology does not rely on optical detection, but it employs a semiconductor device that detects pH changes caused by the release of a H⁺ ion when a dNTP is incorporated. At homopolymeric sequences, multiple dNTPs can be incorporated during a cycle (Goodwin et al., 2016). The change in pH will be larger, but the accuracy in discriminating the number of dNTPs incorporated diminishes with the length of the homopolymer, creating a distinctive error profile for this technology. In **paper I** the sequencer Ion Personal Genome Machine (PGM) (Ion Torrent, Thermo Fisher Scientific) was used to generate the data.

3.2.2 HTS applications

HTS can be applied in different ways: to explore DNA sequence, RNA variety and expression, and even epigenetic markers. While the sequencing step is the same for most applications, a wide range of methods for library preparation exists depending on the question that is being addressed. For studies of Mendelian diseases, the three main applications are gene panels, ES, and WGS. All three applications were used in this thesis. HTS data used in this thesis was generated at Beijing Genomics Institute (Illumina), at the Clinical Genomics facility of Science for Life Laboratory in Stockholm (Illumina) and at the Department of Clinical Genetics of the Karolinska University Hospital, Stockholm (Ion Torrent).

In **paper I** a custom-designed gene panel was used. The panel targeted the coding regions (+25 base pairs at exon-intron boundaries) of 12 genes associated to HLH or susceptibility to EBV infection (*PRFI*, *UNC13D*, *STX11*, *STXBP2*, *SH2D1A*, *XIAP*, *RAB27A*, *LYST*, *AP3B1*, *BLOCIS6*, *MAGT1*, and *ITK*). Evolutionarily conserved regions of FHL genes were also included, since intronic disease-causing variants in *UNC13D* were previously reported (Meeths et al., 2011; Entesarian et al., 2013; Seo et al., 2013). Totally, the size of the target was about 40kb. The panel was designed according to Ion Ampliseq technology (Ion Torrent, Thermo Fisher Scientific), which relies on multiplex-PCR for amplification of the regions of interest. The libraries were subsequently sequenced on an Ion PGM sequencer (Ion Torrent, Thermo Fisher Scientific). An overview of the method is shown in Figure 5A.

In **papers III, V, and VI** ES was used. ES has been a very successful HTS strategy to discover variants underlying Mendelian phenotypes. In ES, only the coding-part of the genome, about 1% of the genome, is captured and sequenced on a high-throughput sequencer. Even though only a minimal fraction of the genome is analyzed, most variants causative of Mendelian disorders disrupt protein-coding sequences (Stenson et al., 2017), thus making ES data a highly informative dataset to look for variants with large effect sizes. The feasibility of ES for the identification of disease-causing variants was first demonstrated in 2009 (Ng et al., 2009). Shortly after ES made its first gene discovery (Ng et al., 2010), and was also used in a clinical setting to diagnose a child with a severe and intractable inflammatory bowel disease. Interestingly, that child was diagnosed with XLP2 due to a pathogenic variant in *XIAP* (Worthey et al., 2011). Due to ES and most recently also to WGS the rate of gene discovery in the last few years has been three times faster than when conventional methods, such as positional cloning, were used (Chong et al., 2015). ES/WGS has also enabled many gene discoveries in the field of PIDs (Meyts et al., 2016).

In **papers III, V, and VI** we used Agilent SureSelect (Agilent Technologies, Santa Clara, CA, USA) as capture strategy, which is illustrated in Figure 5B. In **paper VI** ES data of individual F1:I-3 was produced on an Ion Proton machine after enrichment with Ion Ampliseq Exome kit (Ion Torrent, Thermo Fisher Scientific).

In **paper IV** WGS was used on patients with monoallelic pathogenic variants in *RAB27A* detected by exome sequencing or yet unexplained cases. In WGS, as shown in Figure 5B, the fragmented DNA is directly prepared for sequencing without PCR amplification (PCR-free) which results in a more even coverage. Data from PCR-free WGS is also very suitable for detection of SVs. In **paper IV** fragmented DNA from P1 and P2 was prepared for sequencing using TruSeq DNA PCR-free HT Sample preparation kit and sequenced on a Illumina X sequencer to achieve a 30X sequencing depth (Illumina Inc, San Diego, CA, USA).

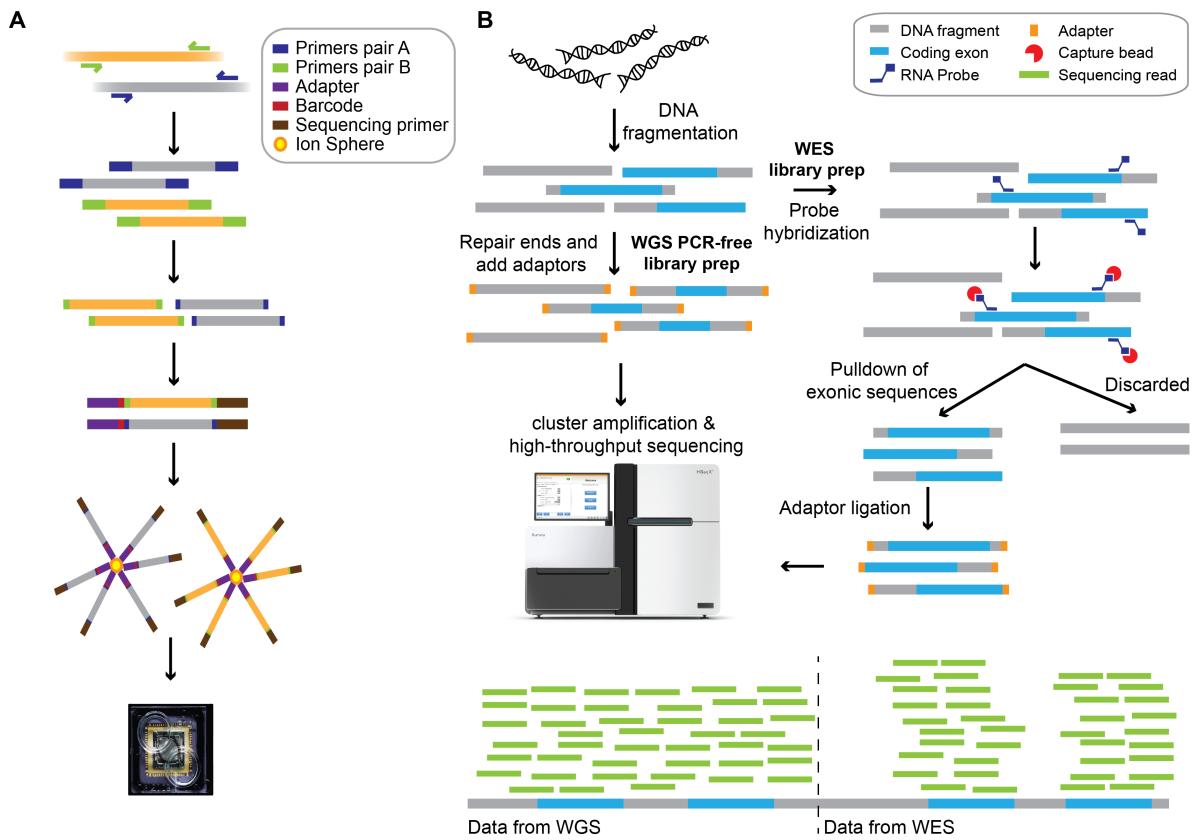


Figure 5. A) Overview of the experimental workflow from library preparation to sequencing according to Ion Ampliseq technology. Briefly, genomic DNA is enriched by multiplex-PCR. After partial digestions of primers, adapters with barcode sequences and sequencing primers are ligated to the PCR product. Libraries are purified and pooled. Thereafter, the DNA templates are attached to ion spheres, amplified through emulsion-PCR, and then loaded into a sequencing chip. B) Overview of ES and WGS experimental workflow. In ES, fragmented DNA is hybridized to biotinylated RNA-probes complementary to human exons. Fragments bound to the RNA-probes are then captured through streptavidin-coated magnetic beads. In WGS, fragmented DNA is directly prepared for sequencing. The image of Illumina HiSeq X Ten is a courtesy of Illumina, Inc.

Gene panels, ES, and WGS all have their strengths and weaknesses. For instance, strategies that rely on capture and amplification, like gene panels and ES, will suffer from non-uniform coverage. In fact, WGS outperforms ES for coverage of coding sequences (Lelieveld et al., 2015). On the other hand, both gene panels and ES might allow for higher sequencing depth, which is required, for example, to properly capture low-frequency mosaicism. Depending on the clinical or scientific question, a selection of the most appropriate method and study design should be made.

3.3 ANALYSIS AND INTERPRETATION OF GENETIC DATA

High-throughput sequencers produce millions of short reads, about 60 million for an exome and 400 million for a whole-genome. Going from a series of disorganized DNA fragments to a list of candidate variants is a multi-step process. During recent years, the field of

bioinformatics has also developed and improved to support the increasing use of HTS. The text below gives a short overview of my bioinformatics pipeline as it looked during the last period of my PhD studies. The specific versions and software used in each paper are indicated in the respective methods section. The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC b2012204 and b2015280.

The first step in the analysis of HTS data, besides standard quality control, consists in the assignment of each sequencing read to a specific location in the genome, a process called *mapping*. The starting point for the mapping are *fastq* files, which contain all the sequencing reads and their per-base quality scores. In this thesis the genome build *hg19* was used for mapping. Aligned reads are collected in *BAM* files, which can be used for visualization of sequencing reads in genome browsers, such as the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al., 2013). *BAM* files are also the starting point for *variant calling*, i.e. the comparison between the sequencing reads produced by the HTS experiment and the reference genome. Called variants are stored in variant call format (*VCF*) files.

The Ion Torrent data produced for **paper I** and **VI** was analyzed with the Ion Torrent Suite Software (versions 4.0.2 and 4.0.3, Ion Torrent, Thermo Fisher Scientific), while Illumina ES data was analyzed according to the Genome Analysis Toolkit (GATK) best practices for germline SNP and InDels discovery in ES (DePristo et al., 2011; Van der Auwera et al., 2013). The GATK best practices recommend the use of the mapper BWA (Li and Durbin, 2009) and of the variant caller GATK Haplotype Caller, together with a series of intermediate steps in order to reduce false positive and false negative calls. WGS data in **paper IV** was analyzed with the SpeedSeq pipeline (Chiang et al., 2015). For SV analysis, ExomeDepth was used in **paper III** (Plagnol et al., 2012), and LUMPY in **paper IV** (Chiang et al., 2015). In **paper IV**, visual inspection of the *RAB27A* loci in IGV was also used to identify the breakpoints of the SVs and split reads were mapped using the BLAT function of the UCSC genome browser.

Having obtained a list of genetic variants, *annotation* is the next step. The Variant Effector Prediction (VEP) software (McLaren et al., 2010), among others, annotates variants with several variant-level information, e.g. the gene affected if any, the kind of variant (e.g. exonic/intronic or missense/synonymous, etc.), the frequency of the variant in large datasets such as 1000G (The 1000 Genomes Project Consortium, 2015) and ExAC (Lek et al., 2016), *in-silico* predictions of pathogenicity, evolutionary conservation score, and much more.

The HTS data in this thesis has been filtered based on inheritance models, frequency of the variants in large datasets, and variants' effect. For instance, in **paper III** and **V** an autosomal recessive disorder was suspected based on parental consanguinity. In **paper VI** an X-linked or an autosomal dominant inheritance patterns was suspected. Disease-causing variants for

rare Mendelian disorders are expected to be rare. In the case of severe pediatric diseases like most PIDs are, pathogenic variants are expected to be nearly absent from large datasets, or to appear at a frequency compatible with carrier-ship for autosomal recessive disorders. The release of the ExAC data in 2014 has been a game changer for frequency-based filtering: for example, the use of a “popmax” frequency filter based on the ExAC dataset produces a 7-fold reduction in number of variants compared to only using data from the Exome Sequencing Project (Lek et al., 2016). As for the effect of the variants, nonsynonymous, nonsense, splice-site, and frameshift variants were analyzed further. In **paper I**, rare non-coding variants in evolutionarily conserved intronic sequences were also retained for downstream analysis.

The GEMINI software has been very valuable to the analysis of genomic data in this thesis (Paila et al., 2013). GEMINI allows for annotation and storage of genetic variants in a database that can be queried according to inheritance models, as well as through custom variant- and gene-level filters. While HTS data in **paper III-V** was produced only for the affected individuals, ES and WGS are especially powerful with a trio design, where the DNA from the affected individual and his/her parents are sequenced (Bamshad et al., 2011). This way, filtering on putative inheritance models becomes much more effective, especially for the identification of compound heterozygous and *de novo* variants. Indeed, three generational ES data was used in **paper VI** to identify the disease-causing variant in family 1.

Having applied the hard filters described above, variants were then reviewed manually. Manual curation weighted in both variant- and gene-level information. *In-silico* predictions such as PolyPhen-2 (Adzhubei et al., 2010), SIFT (Ng and Henikoff, 2003), CADD scores (Kircher et al., 2014), and NNSPLICE 0.9 (Reese et al., 1997) were used for variant-level analysis. When analyzing small gene panels for a specific phenotype, like in **paper I**, variant-level evidence of pathogenicity might be enough to classify a variant as pathogenic or likely pathogenic, and therefore to establish a molecular diagnosis (Richards et al., 2015). When analyzing ES data, such as in **papers III-V**, candidate variants in known-genes, especially if missense variants as in **paper III**, should be evaluated further through review of the clinical phenotype, segregation studies, and, if possible, design of functional experiments to confirm pathogenicity. When reviewing variants in genes not previously associated with Mendelian diseases, several variant- and gene-level lines of evidence are required to classify a variant as disease-causing (Casanova et al., 2014; MacArthur et al., 2014). Some lines of evidence, such as expression of the gene in the disease-relevant cell type (*e.g.* immune cells) can be obtained through publicly available data, for others experimental validation is required. Generally, despite the use of experimental validation, a novel gene can be indisputably associated with a Mendelian phenotype only when multiple phenotypically similar, unrelated patients are shown to carry pathogenic variants in the same gene (MacArthur et al., 2014). However, in the field of PIDs several novel genes have been discovered through studies of single patients, mainly due to a better accessibility of blood and immune cells for functional validation (Casanova et al., 2014).

3.4 VALIDATION OF GENETIC FINDINGS

Different methods can be used to validate genetic findings. Segregation study of a putative disease-causing variant among family members is important to confirm segregation of the candidate variant with the disease and confirm the inheritance pattern. Family segregation studies were performed in **all papers** when possible. In **papers I and III-VI**, Sanger-based verification of findings from HTS was also performed. For individuals diagnosed with primary HLH in **papers I-IV**, immunological assays for HLH diagnostics were performed when possible.

In **paper III**, the effect of the exonic deletion detected in *IFNGR2* was evaluated in RNA isolated from patient A. To prove a dysfunction in IFN- γ signaling, phosphorylation of STAT1 upon stimulation with IFN- γ was evaluated in primary white blood cells obtained from patient A.

In **paper V**, the SVs detected in *RAB27A* were validated by breakpoint PCR and by a custom-designed multiplex ligation-dependent probe amplification (MLPA) assay. MLPA is a multiplex-PCR based method for the detection of CNVs in specific regions of interest. MLPA was also used to study segregation of the SVs among family members. The effect of the SVs on *RAB27A* expression was evaluated by quantitative PCR (qPCR) and western blot analysis. The qPCR assay was designed to be able to distinguish between the *RAB27A* isoform NM_183235.2 and the others. Expression of *PIGB* upstream and downstream of the breakpoint was also evaluated by qPCR.

In **paper VI** several genetic methods, such as Sanger sequencing, SNP-arrays, and digital-PCR (dPCR) were used to understand the genetic mechanisms of *SAMD9L* disease-causing variants. In particular, SNP-arrays were used to detect copy-neutral loss of heterozygosity (CN-LOH), and dPCR was used to quantify the presence of germline and somatic *SAMD9L* variants in different cell types isolated through cell sorting. Functional assessment of *SAMD9L* variants was performed in HEK 293FT cells. Expression vectors containing wild type and mutant *SAMD9L* were transfected into HEK 293FT cells and their effect was evaluated through proliferation assays. Wild-type expression of *SAMD9L* and *SAMD9* protein was evaluated by western blot in fibroblasts, CD34+ cells, and NK cells at steady state and after stimulation with IFN- α and IFN- γ .

A vast amount of HTS data has been deposited online for other researchers to use. In **paper I** genetic data from 1000G (The 1000 Genomes Project Consortium, 2015) was used to compare genetic variation in HLH genes between healthy individuals and patients with secondary HLH. In **paper IV**, cap analysis gene expression (CAGE) data from the FANTOM project (The FANTOM Consortium and the RIKEN PMI and Clst (dgt), 2014) was used to define the transcriptional start site of *RAB27A* in immune cells and melanocytes.

3.5 IMMUNOLOGICAL ASSAYS FOR DIAGNOSIS OF PRIMARY HLH

The assays outlined below have been used in **Papers I-IV**. These assays were performed immediately after reception of blood samples, since they require viable cells to work properly. Therefore, in most cases, results from immunological assays were already available during analysis of genetic data, facilitating interpretation of genetic findings. In some cases, immunological assays were repeated after achievement of a molecular diagnosis, with analysis of peripheral blood mononuclear cell (PBMC) at resting state and after incubation with IL-2.

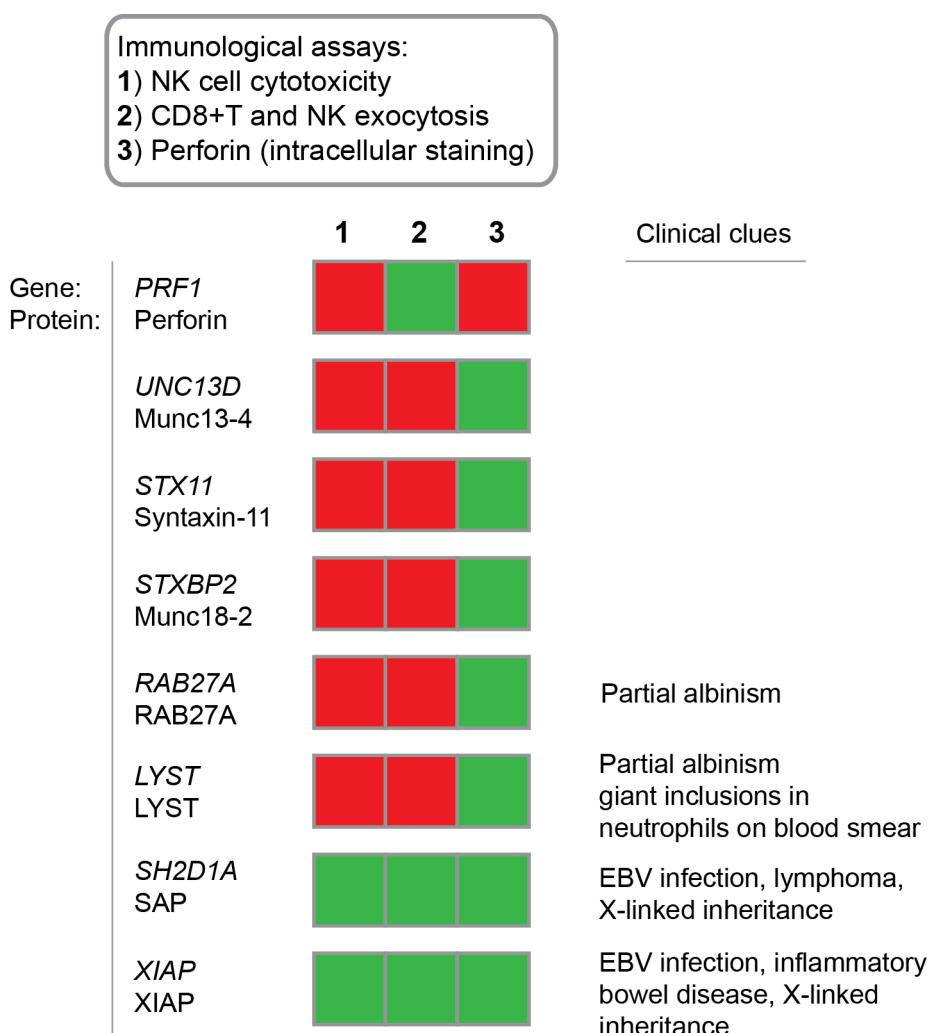


Figure 6. Schema of expected results from immunological assays in patients with primary HLH based on the underlying genetic defect. Red and green are used to illustrate, respectively, an abnormal (defective) result, and a normal result.

NK cell cytotoxicity was evaluated with the well-established chromium release assay. K562 is a human erythroleukemia cell line that lacks expression of MHC class I molecules, and therefore elicits prompt killing by NK cells. K562 cells were first labeled with ^{51}Cr . K562 cells that are targeted and killed by effector NK cells release their internally stored ^{51}Cr . The

higher the level of ^{51}Cr detected, the more K562 cells were killed. The ^{51}Cr -release in the supernatant was measured with a gamma-counter. The data was displayed as % specific lysis for different effector to target ratios, but also summarized as lytic units (LU) at 25% specific lysis. A value of less than 10 LU (at 25% specific lysis) was considered pathologic.

Even though NK cell cytotoxicity is included among the HLH-2004 diagnostic criteria (Henter et al., 2007), degranulation assays have proved to be a more reliable assay to detect primary HLH due to defect in exocytosis (Bryceson et al., 2012; Chiang et al., 2013; Rubin et al., 2017) (Figure 6). In **paper I**, **III**, and **IV**, $\text{CD8}^+\text{CD57}^+$ T and NK cells exocytosis was assessed by flow cytometry. Upon exocytosis of cytotoxic granules, cytotoxic lymphocytes expose CD107a, also known as Lamp-1, on the cell surface. Instead, at resting state, CD107a is associated with cytotoxic granules. Assays to measure exocytic activity of $\text{CD8}^+\text{CD57}^+$ T and NK cells use surface expression of CD107a after incubation with target cells as a proxy to evaluate degranulation activity. Stimulations used include K562, or P815 with anti-CD16 antibody for NK cells and P815 with anti-CD3 antibody for $\text{CD8}^+\text{CD57}^+$ T cells. Effector and target cells were co-incubated at a 1:1 ratio for 3 hours at 37°C. Cells were then stained with anti-CD107a antibody as well as immunophenotypic markers before acquisition by flow cytometry (LSR Fortessa, BD Biosciences). Data was analyzed on FlowJo (v7 or v9, TreeStar). The percentage of degranulation was calculated as $\Delta\text{CD107a}\%$, that is the % difference in CD107a surface expression between stimulated and unstimulated effector cells. Intracellular staining was used to determine expression levels of perforin, granzymes, CD107a, and SAP in **papers I** and **II**, as described by Chiang et al. (Chiang and Bryceson, 2016). Briefly, PBMCs were surfaced-stained with CD3 and CD56 fluorochrome-conjugates antibodies and with a dye to discriminate between alive and dead cells. Cells were thereafter fixed and stained for the intracellular proteins above. Cells were analyzed by flow cytometry (LSR Fortessa, BD Biosciences) and the data produced was analyzed on FlowJo (v7 or v9, TreeStar).

4 RESULTS AND DISCUSSION

4.1 GENETIC ARCHITECTURE OF HLH

4.1.1 Variant spectrum of primary HLH

HLH is a genetically heterogeneous disease. Besides classical FHL forms and syndromes with albinism, other PIDs causing susceptibility to EBV infection, such as XLP1 and XLP2, ITK deficiency, and others, are associated with HLH (Parvaneh et al., 2013). Quickly, the list of genes to be evaluated can grow to a length that is incompatible with a Sanger-based laboratory approach. Moreover, primary and secondary forms of HLH are hard to distinguish; yet such distinction is clinically important and valuable for treatment decisions (Henter et al., 2007). To improve diagnostics of HLH, in **paper I** we designed and used a HTS-based gene panel to simultaneously investigate 12 genes, associated with primary HLH or with susceptibility to EBV infections, in 58 prospectively recruited patients with HLH. Through bioinformatics analysis and filtering, 71 rare variants were subjected to manual curation, and 19 of them were classified as disease-causing. In addition, one indel in *RAB27A* was detected through visual inspection of reads in IGV, and two exonic deletions, in *XIAP* and in *STX11*, through analysis of coverage. Overall, 22 distinct disease-causing variants provided a molecular diagnosis for 22 out of the 58 patients analyzed, with a diagnostic rate of 38%. For eight disease-causing variants not previously reported in FHL patients, pathogenicity was confirmed through segregation studies and concordance with results from functional assays when available. Seven patients were diagnosed with FHL2 (*PRFI*), six with FHL3 (*UNC13D*), and four with FHL4 (*STX11*). Among patients with albinism, two were diagnosed with GS2 (*RAB27A*), and two with CHS (*LYST*). An additional patient was diagnosed with XLP2, due to a hemizygous deletion in *XIAP* (Figure 7A,B). One patient with albinism and HLH (P1) was instead diagnosed with FHL3, due to a homozygous splice-site variant in *UNC13D*. This patient might have two distinct Mendelian disorders that cause a compound phenotype, something not uncommon especially in consanguineous families. Indeed, in another patient with albinism, HLH, and low CD8⁺T cells, we established a double diagnosis of SCID, due to a homozygous pathogenic variant in *JAK3*, and AR albinism oculocutaneous, due to a homozygous pathogenic variant in *TYR* (**paper VII**). No disease-causing variants were found in the other sequenced genes. A separate analysis of evolutionarily conserved intronic regions of FHL genes did not reveal any additional pathogenic variant.

The cohort in **paper I** was heterogeneous for ethnic groups, since recruitment happened through a network of collaborating physicians worldwide. Therefore, this cohort, also small in size, is not suitable to draw conclusion on the contribution of the respective genes. However, the majority of patients were diagnosed with FHL2 or FHL3, as observed in other cohorts (Sieni et al., 2014; Cetica et al., 2016). Among patients from Northern Europe (n=3), two were diagnosed with FHL3 and one with FHL2. One of the FHL3 patients (P11) was

compound heterozygous for a splice-site variant and for the 253kb *UNC13D* inversion, which is the most common FHL aberration in Northern Sweden (Meeths et al., 2011). The 253kb *UNC13D* inversion was detected through a separate multiplex-PCR assay, as previously described (Meeths et al., 2011). The cohort included several patients of Turkish origin, as reflected in the higher frequency of disease-causing variants in *STX11* (FHL4) compared to what was observed in Swedish and Italian cohorts (Meeths et al., 2011; Cetica et al., 2016).

In **paper I** we focused mainly on pediatric HLH, and only eight adult patients with HLH were included. In the pediatric sub-cohort (n=50), the overall diagnostic rate was 44%. The diagnostic rate was highest (65%, 13 out of 20) among children diagnosed with HLH during the first year of age (Figure 7A). A similar diagnostic rate was reported in another study (Cetica et al., 2016) and it is concordant with epidemiological data showing a median age at diagnosis of HLH of 5 to 8 months (Trottestam et al., 2011; Meeths et al., 2015). No case of primary HLH was identified among the few adult patients studied.

In 36 patients no disease-causing variants were found. Among the patients evaluated with functional assays, some showed defects in one of the assays, but none showed a combined defect highly indicative of FHL, GS2, or CHS, suggesting no pathogenic variants in those genes were missed. To ensure this, visual inspection of sequencing reads and Sanger-based analysis of poorly covered exons was performed. It is possible that some of these patients harbor other genetic disorders. Parental consanguinity was indeed frequent even among undiagnosed cases. Others instead will represent truly secondary HLH cases, something that is also suggested by a higher frequency of known triggers of HLH in the undiagnosed group.

In **paper II**, through a retrospective study, we report FHL2 patients with atypical presentations and intra-familial variability. All ten patients were reported to have biallelic missense pathogenic variants in *PRF1*, for a total of seven different missense variants, of which six had previously been reported in atypical or late-onset cases of FHL2. Two of the variants were also identified in **paper I**: p.Thr450Met in P17, who developed HLH at 5 years of age, and p.Arg225Trp in P35, who developed HLH at the age of 8.5 years.

4.1.2 Non-coding pathogenic variants

Most disease-causing variants underlying Mendelian diseases are located in exons and splice junctions. As a result, diagnostic efforts in patients with Mendelian diseases have for a long time focused mainly on coding sequences, previously with Sanger sequencing and now with ES. Nonetheless, very few Mendelian disorders can be molecularly diagnosed with a 100% diagnostic rate, and the average diagnostic rate across many disorders is ~50% (Chong et al., 2015). At this regard, today's diagnostic success for primary HLH can be considered to be above average. In fact, in **paper I**, despite an overall 44% diagnostic rate for pediatric cases, all cases with a functional defect suggestive of an inborn defect of lymphocyte cytotoxicity received a molecular diagnosis. No disease-causing variants in FHL genes were found in

patients without a typical functional defect. The positive and negative predictive value of combined functional assays is high (Chiang, 2015; Rubin et al., 2017), and therefore non-coding pathogenic variants should be suspected in molecularly undiagnosed HLH patients with functional defects suggesting primary HLH. With this approach and focusing on evolutionarily conserved intronic regions, previous studies from our research group have identified non-coding disease-causing variants in *UNC13D* (Meeths et al., 2011; Entesarian et al., 2013).

In **paper IV**, we studied five patients from five unrelated families from Lithuania (n=3), Sweden (n=1), and Russia (n=1) with an atypical form of HLH. In three of these patients, clinical sequencing revealed monoallelic variants in *RAB27A*, including two previously reported as disease-causing, c.239G>C (p.Arg80Thr) in P2, and c.550C>T (p.Arg184X) in P5 (Meeths et al., 2010a). The monoallelic *RAB27A* variant c.559C>T (p.Arg187Trp) identified in P1 was instead classified as a variant of unknown significance. Clinical sequencing had not revealed any pathogenic variants in P3 and P4, yet functional assays showed defective exocytosis of CD8⁺CD57⁺ T cells and NK cells, suggestive of FHL3-5, GS2, or CHS.

GS2 is a PID characterized by partial albinism, silvery-gray hair, and susceptibility to HLH (Ménasché et al., 2000). GS2 patients with truncating variants in *RAB27A* develop HLH in early infancy, and more than half the patients have neurological involvement (Meeths et al., 2010a). None of the patients reported in **paper IV** displayed signs of oculocutaneous albinism. Recently, some patients with GS2 yet without oculocutaneous albinism have been reported (Cetica et al., 2015; Netter et al., 2016). We therefore suspected that a pathogenic variant on the other allele of *RAB27A*, potentially non-coding, had gone undetected in clinical sequencing. To test our hypothesis, we performed WGS on genomic DNA from P1 and P2.

We used IGV to visually analyze WGS data at the *RAB27A* locus. In both patients, we identified several aberrant read pairs at the 5'UTR of *RAB27A*, also involving the first 5 exons of the adjacent gene *PIGB*, which is required for glycosylphosphatidylinositol anchor biosynthesis. Read pairs were aberrant both for insert size and orientation, suggesting events of deletion/duplication as well as inversion, as confirmed by the SV calling software LUMPY. Interestingly, two SV events were shared between P1 and P2. The detected SVs were further analyzed by split reads mapping and breakpoint-PCR. A custom MLPA-assay was used to define gains and losses in the region (Figure 7C).

P1 and P2 shared a heterozygous complex SV characterized by two non-adjacent duplicated regions, with identical breakpoints. Duplication-normal-duplication pattern can associate with cryptic and complex rearrangements (Brand et al., 2015). Indeed, one breakpoint suggested that one of the duplicated regions was also inverted (Figure 7D). P1 also carried a heterozygous 65kb deletion of *RAB27A*, in *cis* with the c.559C>T missense variant. Due to clinical resemblance and geographic proximity, the other three patients (P3-P5) were

analyzed by MLPA and breakpoint-PCR. P3 and P4 were homozygous for the complex SV identified in P1 and P2, while P5 was heterozygous. With biallelic aberrations in *RAB27A*, a molecular diagnosis of GS2 was established in all five patients.

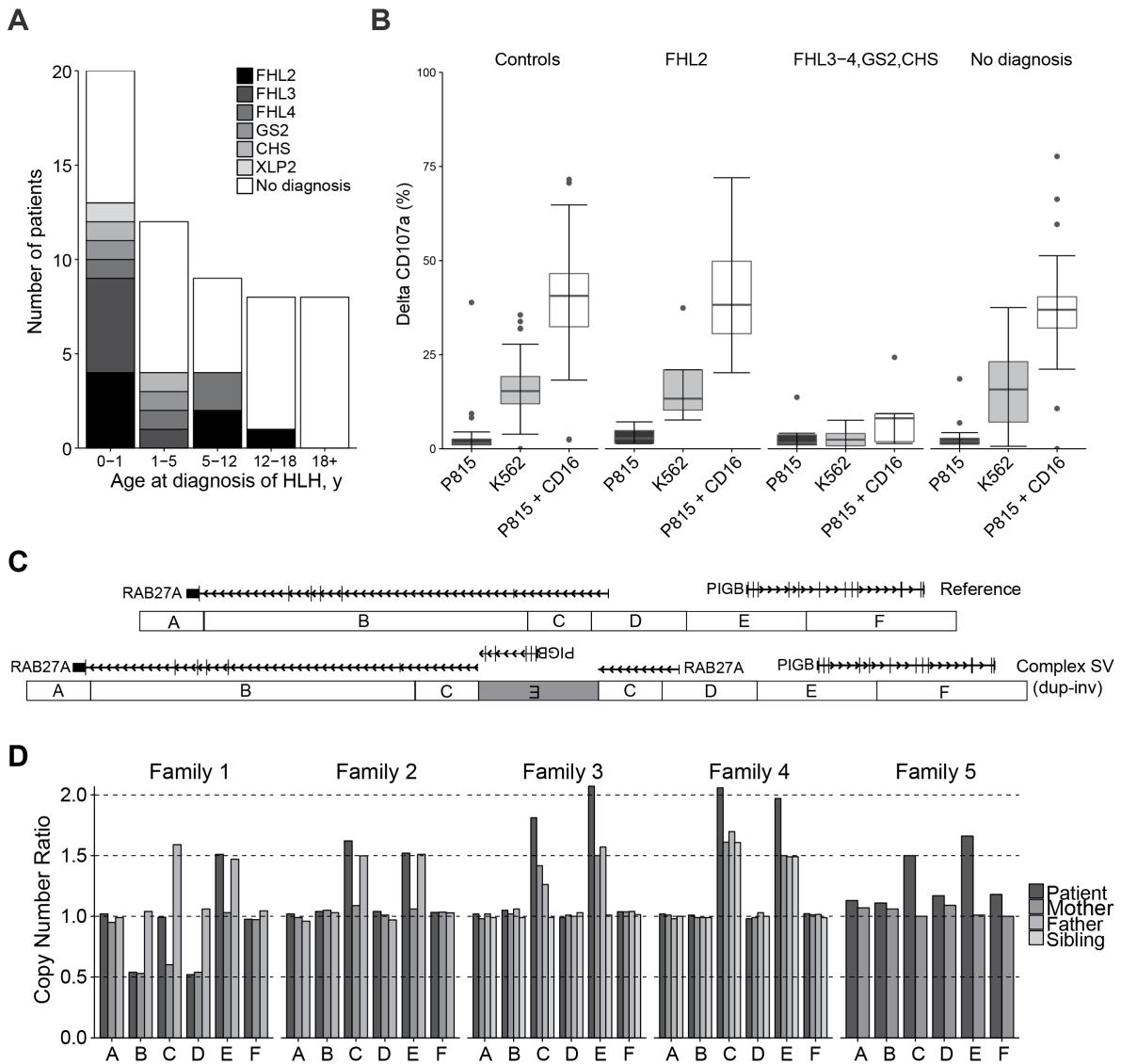


Figure 7. A) Molecular diagnoses identified in **paper I**. B) NK cells exocytic activity of patients studied in **paper I**. Patients are grouped by their molecular diagnosis (FHL2, n=4; FHL3-4, GS2, CHS, n=9; no diagnosis, n=20). C) Gain and losses shown by a custom MLPA assay over the *RAB27A/PIGB* loci in GS2 patients from **paper IV**. Location of MLPA probes is depicted in (D), together with a model for the complex SV as supported by split reads and breakpoint-PCR.

Our reconstruction of the shared complex SV suggests a model (Figure 7C) where only the transcriptional start site (TSS) of one transcript of *RAB27A* (NM_183235.2) is disrupted. Western blot analysis of Rab27a expression in PBMCs from the patients showed markedly decreased expression. Similarly, qPCR showed diminished expression of NM_183235.2 transcript, confirming that the complex SV disrupts gene expression by removing the TSS of

transcript NM_183235.2 and that this transcript is expressed in lymphocytes. According to this model, a copy of *PIGB* remains intact, and qPCR showed preserved expression.

We therefore hypothesized that the lack of albinism in these patients might be due to distinct TSS usage between lymphocytes and melanocytes. Analysis of cap analysis gene expression (CAGE) data from the FANTOM5 project (The FANTOM Consortium and the RIKEN PMI and Clst (dgt), 2014) revealed that in nearly all cell types, including cytotoxic lymphocytes, NM_183235.2 is the predominant transcript (Figure 8). However, primary as well as embryonic stem cell-derived melanocytes also use alternative, downstream TSSs, which are not disrupted by the complex SV. Interestingly, acquisition of downstream TSS seems to occur during differentiation from embryonic stem cells. qPCR data confirmed a smaller contribution of the NM_183235.2 transcript to the total *RAB27A* transcription in melanocytes.

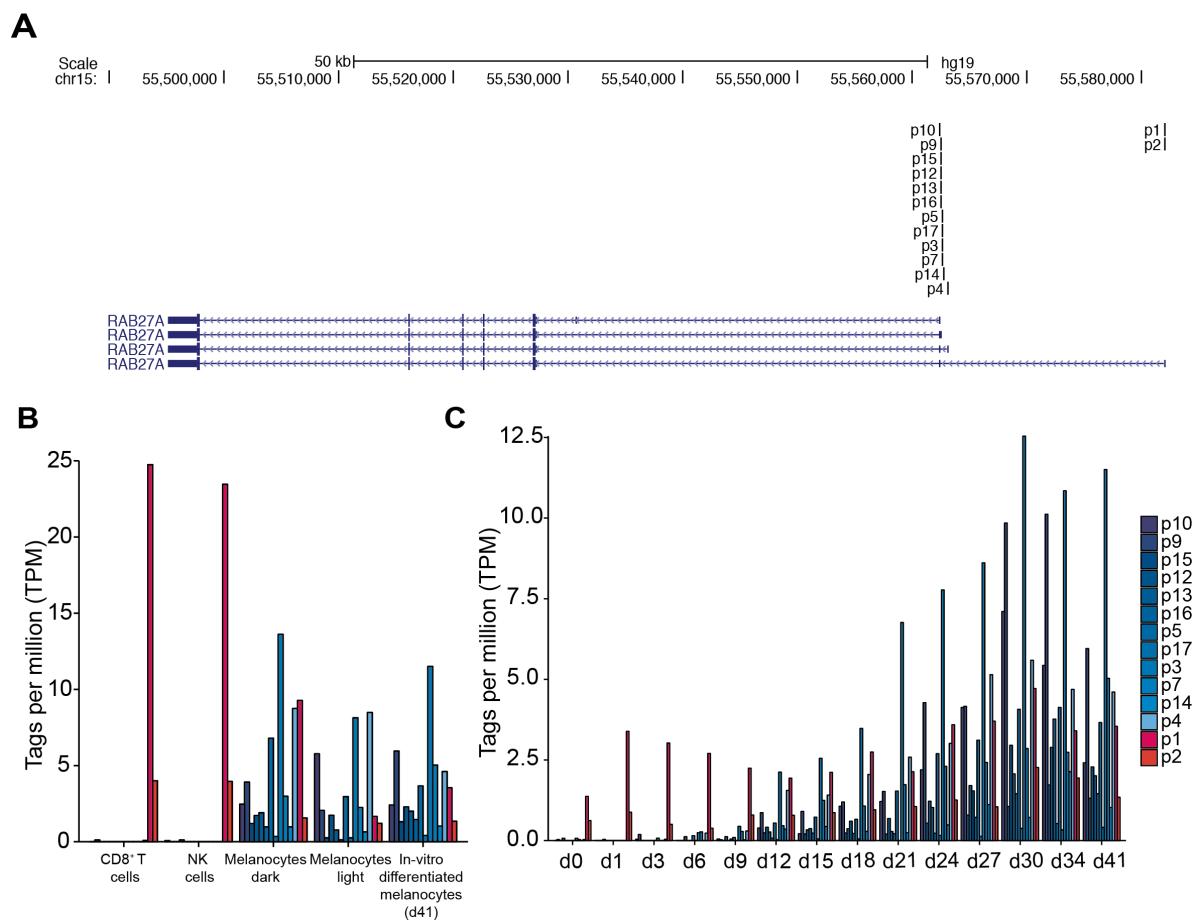


Figure 8. (A) UCSC Genome browser screenshot showing localization of the *RAB27A* TSS according to FANTOM5 CAGE data from human cells. p1 is the TSS for the transcript NM_183235.2. (B) Tags per millions at the different *RAB27A* TSS from the FANTOM5 CAGE data. (C) Tags per millions at the different *RAB27A* TSS during differentiation of human embryonic stem cells into pigmented melanocytes.

In conclusion, we identified SVs affecting the 5'UTR of *RAB27A* in patients with atypical HLH. The lack of albinism was attributed to differential usage of *RAB27A* TSS between

lymphocytes and melanocytes. Moreover, the detection of a recurrent complex SV in geographically related patients suggests a founder effect in the Baltic population. This finding ensures access to a molecular diagnosis for future patients from this area.

4.1.3 HLH beyond defects of cytotoxic lymphocytes: HLH in other PIDs

HLH is the phenotypic manifestation of a dysregulated immune response. In the introduction, we discussed how immune dysregulation and inflammation are present across several types of PIDs. Therefore, the report of HLH in patients with PIDs other than inborn defects of lymphocyte cytotoxicity might not seem surprising. On the other hand, although these patients do not represent *bona fide* primary HLH cases, they still have an underlying genetic defect, and therefore should be distinguished from truly secondary HLH cases. Establishing the right molecular diagnosis in these patients might provide access to tailored treatment and appropriate genetic counseling. Lastly, studying other defects of immunity that can lead to HLH might increase our knowledge about pathogenic mechanisms of hyperinflammatory syndromes.

In **paper III** we report a 2-month-old infant and a 4-year-old child diagnosed with HLH due to disseminated *Mycobacterium bovis* following Bacillus Calmette–Guérin vaccination and *Mycobacterium tuberculosis* infection, respectively. We performed ES on patient A, and, due to parental consanguinity, focused on homozygous variants. Analysis of CNVs revealed a homozygous exonic deletion in *IFNGR2*. RNA studies showed that the deletion of exon 2 led to an early frameshift. Cellular analyses confirmed abrogated IFN- γ -mediated STAT1 phosphorylation, whereas cytotoxic lymphocyte degranulation and perforin expression were normal. In patient B, Sanger sequencing of *IFNGR1* and *IFNGR2* revealed a homozygous missense variant in *IFNGR1*, c.655G>A p.Gly219Arg. The variant was deemed pathogenic based on segregation studies and low frequency in the population. The variant was absent from 1000G and, at the time of writing, is present in 2 out of 8624 individuals from South-East Asia in gnomAD (Lek et al., 2016). Moreover, an in-frame deletion affecting codon 218 was previously reported in a patient with *IFNGR1* deficiency (Jouanguy et al., 2000). *IFNGR1* and *IFNGR2* code for the subunits of the IFN- γ receptor, and a molecular defect in either one of the IFN- γ receptor chains is sufficient to impair IFN- γ signaling.

Pathogenesis of primary HLH is thought to be driven by an excessive production of IFN- γ by activated CD8 $^+$ T cells. In turn, IFN- γ activates macrophages. Development of HLH in the context of abolished IFN- γ signaling is puzzling. Four additional patients with HLH in the context of Mendelian susceptibility to mycobacterial disease due to pathogenic variants in *STAT1* or *IFNGR1* have recently been reported (Staines-Boone et al., 2017). Interestingly, HLH/macrophage-activation-like disease can be induced in wild-type mice through repeated Toll-like receptor 9 stimulation. In this model, IFN- γ -knockout mice developed disease comparably to wild-type mice, suggesting IFN- γ may be dispensable for the development of

immune pathology (Canna et al., 2013). This observation may be of significance in the context of current attempts to treat HLH with anti-IFN- γ therapy.

In **paper VIII**, a multicenter, systematic survey of published and unpublished cases of HLH in other PIDs was conducted. We contributed with the report of a British boy (patient 24) diagnosed with EBV-driven HLH at the age of 4 years. Through ES we revealed a known disease-causing variant (c.838A>C, p.Ser280Arg) in *DKC1* and established a diagnosis of X-linked dyskeratosis congenita. Indeed, the patient had a history of nail dystrophy and other features of dyskeratosis congenita, including short telomeres (unpublished data). In total, 63 cases of HLH in PIDs were identified through literature search and survey among PID/HLH centers worldwide. Importantly, in about half of the cases HLH represented the first manifestation of an underlying PID, as in the patients from **paper III**. Moreover, in several patients, HLH developed despite low numbers of T cells and NK cells, suggesting a lymphocyte-independent pathogenesis for HLH in T cell deficiencies.

Besides PIDs, HLH has been reported also in metabolic disorders, such as Wolman disease. Wolman disease is caused by biallelic disease-causing variants in the gene *LIPA*, encoding for lysosomal acid lipase (Taurisano et al., 2014). Indeed, in two patients left without a molecular diagnosis after the screening described in **paper I** we identified by ES homozygous pathogenic variants in *LIPA*, thus establishing a diagnosis of Wolman disease (unpublished data). Both patients developed HLH within the first months of life. With two more patients diagnosed by ES, the diagnostic rate for the infants with HLH studied in **paper I** reached 75%.

4.1.4 Genetic variation in HLH genes in the population

The detection of rare monoallelic variants in FHL genes in patients with secondary HLH is not uncommon (Zhang et al., 2011). Experimental validation of an *in vitro* dominant negative effect has been provided for a few variants affecting FHL genes, yet the genetic findings in these studies were based on sequencing focused on coding regions (Spessott et al., 2015; Zhang et al., 2016). In an Italian cohort of 240 patients with secondary HLH, 43 (18%) were reported to have one monoallelic pathogenic variant (Cetica et al., 2016). The definition of pathogenic variants was somewhat relaxed, since some of the variants reported as monoallelic pathogenic variants are quite frequent in the population, as the *PRF1* variant c.272C>T p.Ala91Val, which occurs in about 5% of Caucasian individuals (Lek et al., 2016). This variant is associated with HLH in homozygous state or *in-trans* to a more damaging variant, as in three patients studied in **papers I-II**. The penetrance of this variant in homozygous state is incomplete and its effect in monoallelic state is unknown. The interpretation of monoallelic variants is also complicated by the fact that such variants are almost always inherited from a healthy parent. Moreover, secondary HLH cases are usually only tested by targeted sequencing of FHL genes, which might overestimate the contribution of heterozygous variants to HLH susceptibility.

Large-scale sequencing projects have now established that rare genetic variation is common, suggesting that most of it is harmless (Lek et al., 2016). Using publicly available data the amount of genetic variation between specific patients group and the general population can be compared. In **paper I**, to quantitatively examine the contribution of monoallelic variants to secondary HLH, we evaluated the frequency of variants in HLH genes among 2504 unrelated individuals from the 1000 Genomes project (The 1000 Genomes Project Consortium, 2015). The data was analyzed with the same filtering strategy used for HLH patients: protein-altering or splice-site variants with a minor allele frequency lower than 0.05, and, for missense-variants, at least one damaging prediction *in-silico*, were considered as possibly damaging. 636 individuals in 1000G (25%) carried at least one variant matching these criteria, which are relaxed but resemble the criteria used in the literature for secondary HLH. Against a frequency of 25% in presumably healthy individuals, monoallelic variants in FHL genes were not enriched in patients with secondary HLH. In **paper I** nine patients (25%) without a definitive molecular diagnosis carried a monoallelic variant that matched the above filtering criteria. Thus, the identification of a heterozygous variant in FHL genes in patients with HLH should be interpreted cautiously in respect to causality. As illustrated in **paper VI** and previous literature (Meeths et al., 2011; Entesarian et al., 2013), more advanced genetic studies might be required to uncover variants that are hard to detect with conventional methods.

4.1.5 Integrated genetic and functional diagnostics for HLH

In **paper I**, the screening of prospective HLH patients was preceded by a validation phase on 13 previously diagnosed patients. In total, 18 disease-causing variants and 56 polymorphisms were used for validation. All variants could be detected by visual inspection, but only 72 out of 74 were called by the software, corresponding to a sensitivity of 97.3% (95% confidence interval 90.7-99.2). The disease-causing variant not called by the software was located in a homopolymeric stretch, an Achilles heel of Ion Torrent technology. Coverage for the regions of interest was 96.6%, but sequencing depth was on average above 200X. Even ES, which too relies on a capture step, suffers from incomplete coverage (Yang et al., 2013). While it is possible to keep track of gaps in a small panel, and to potentially fill them by an alternative approach, this becomes harder on a large-scale. WGS compromises on sequencing depth (usually 30X), but provides uniform coverage, reliable CNV calling, and information on non-coding regions. WGS is therefore replacing gene panels and ES in clinical and research settings. Yet further optimization of strategies for sequencing, analyses, and validation are required to increase the diagnostic success of WGS (Taylor et al., 2015).

In **papers I, III and IV**, different applications of HTS and functional assays for diagnostics of primary HLH were used to achieve a molecular diagnosis. In all cases, functional assays provided important information. In **paper I**, they aided interpretation of genetic findings in patients with pathogenic variants associated with FHL, GS2 and CHS, while in **paper IV**, defective CD8⁺CD57⁺T cell and NK cell exocytosis indicated the presence of non-coding

pathogenic variants. Importantly, in patients with inborn defects of lymphocyte cytotoxicity, functional assays are constitutively abnormal. Thus, abnormal functional assays might provide a strong diagnostic hint for patients with atypical manifestations of primary HLH. Among the functional assays available today, CD8⁺CD57⁺ T cell and NK cell degranulation assays and intracellular staining of perforin outperform NK cell cytotoxicity assay (Rubin et al., 2017). In fact, both patient A from **paper III** and patient 24 from **paper VIII** showed defective NK cell cytotoxicity due to low NK cell numbers. The defect in NK cell cytotoxicity in these patients was comparable with what is seen in FHL and therefore enough to fulfill one of the HLH-2004 diagnostic criteria. However, no defect was seen in NK cell degranulation and perforin expression, prompting us to look for genetic defects other than FHL.

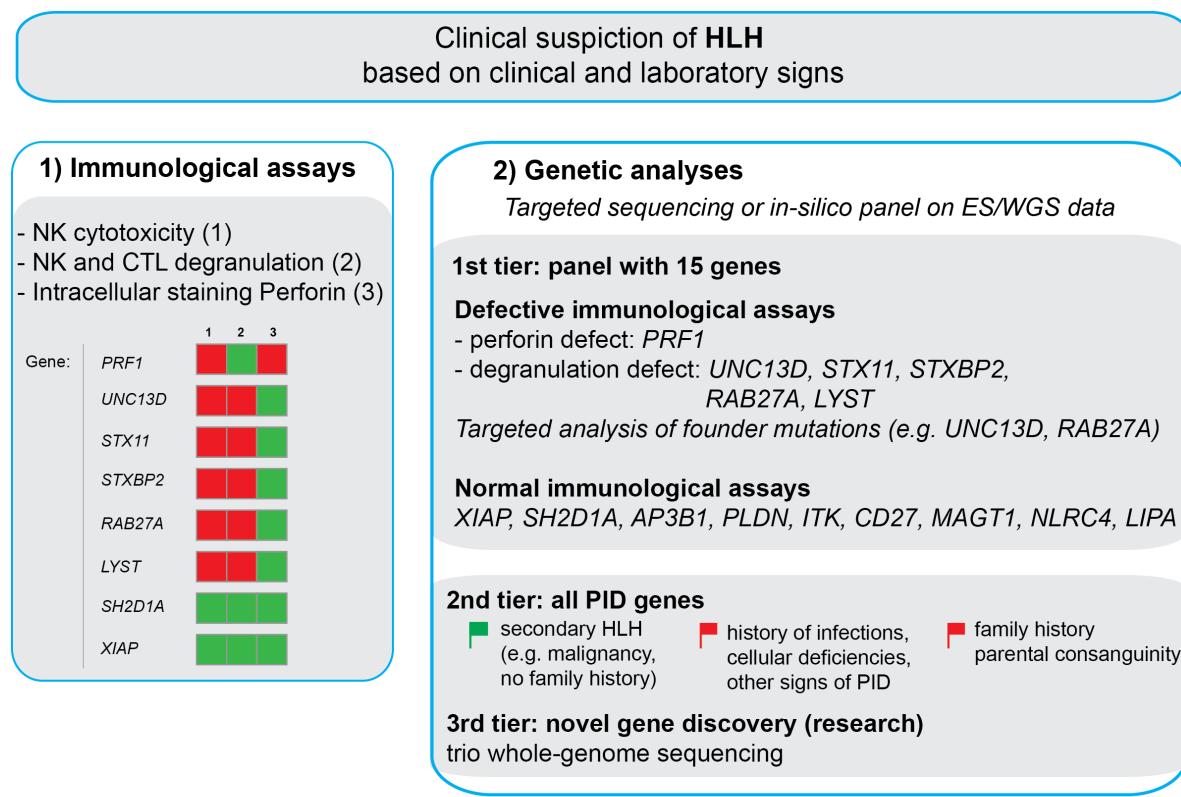


Figure 9. Proposal for a diagnostic approach of patients with HLH based on functional assays and genetic testing. A complete diagnostic algorithm also includes physical investigations and laboratory tests which are not presented here.

Genetic and functional investigations are therefore interdependent for diagnostics of primary HLH. In Figure 9 we propose an updated diagnostic approach to patients with HLH. Functional immunological assays for primary HLH are recommended in all patients as first-line screening. A defect consistent with primary HLH can be sufficient to inform clinical management until a definitive molecular diagnosis is established. To achieve a molecular diagnosis, we propose a first-tier genetic screening based on the analysis of 15 genes through a gene panel or through an *in-silico* panel based on ES/WGS data. Compared to the list used

in **paper I**, the genes *CD27*, *NLRC4*, and *LIPA* have been added. Pathogenic variants in *CD27* cause a PID characterized by severe EBV infection and an increased risk for EBV-driven HLH (Alkhairy et al., 2015). GOF heterozygous variants in *NLRC4* cause an autoinflammatory syndrome that resembles HLH (Canna et al., 2014; Romberg et al., 2014). If sequencing data is based on WGS, analysis of non-coding variants and CNV can be performed directly. Alternatively, patients with defective lymphocyte exocytosis and no findings or only monoallelic variants in *UNC13D* or *RAB27A*, should be evaluated for the respective founder mutations with a targeted method (Meeths et al., 2011) (**paper IV**). Patients without disease-causing variants in the genes above but with red flags for PIDs, such as history of infections, autoimmunity or cellular deficiencies (e.g. low T cells numbers in SCID), and with general red flags for genetic diseases, like a positive family history or parental consanguinity, should be evaluated for pathogenic variants in all known PID genes. In research settings, an unbiased analysis of all variants aimed at discovery of novel genes can be performed. For that, a family study design (e.g. trio) is recommended.

4.2 PHENOTYPIC SPECTRUM IN DISORDERS OF IMMUNE DYSREGULATION

4.2.1 Atypical manifestations of defects in lymphocyte cytotoxicity

4.2.1.1 Late-onset HLH

It is now known that primary HLH can develop beyond childhood (Janka and Lehmburg, 2014). However, late-onset of disease can sometimes have atypical manifestations, making the establishment of a diagnosis difficult. Some of the patients studied in **papers I-II** and **IV** were diagnosed with primary HLH during adolescence and adulthood. In **paper I** the oldest patient diagnosed with primary HLH was 16 years old at diagnosis (P48). P48 was compound heterozygous for two variants in *PRF1*, c.272C>T (p.Ala91Val) and c.1288G>T (p.Asp430Tyr). In **paper II** we report a patient (E) diagnosed with HLH at the age of 38 years with a homozygous *PRF1* Ala91Val variant, and two patients (A:1, B:1) diagnosed, respectively, at 13 and 19 years of age. In **paper IV**, P2 developed HLH at 14.5 years of age and was the oldest case included in the Swedish incidence study by Meeths et al.. P1 and P3 from **paper IV** developed HLH at 14 and at 9 years of age, respectively, while P4 and P5 have so far not developed clinical signs of HLH. In larger cohorts, primary HLH among adults explain a small percentage of all HLH cases (Zhang et al., 2011; Wang et al., 2014b; Zhang et al., 2011), yet a diagnosis of primary HLH has clinical consequences as important in adulthood as in childhood.

4.2.1.2 Cancer

Inflammation and immune cell infiltration are observed in most cancers. Tumor-associated inflammatory response results from the immune system attempt to kill cancer cells. At the same time, the presence of an inflammatory state promotes cancer growth and acquisition by cancer cells of their functional hallmarks (Hanahan and Weinberg, 2011). Occurrence of

cancer has been observed in patients with primary HLH or other PIDs (Leechawengwongs and Shearer, 2012; Meeths et al., 2014). Moreover, in a prospective Japanese study, individuals with decreased NK cell cytotoxicity displayed an increased incidence of cancer (Imai et al., 2000). In **papers II** and **IV** we studied patients who developed lymphoma due to biallelic pathogenic variants in *PRF1* or *RAB27A*. Importantly, development of lymphoma was in some cases the only manifestation of an inborn defect of lymphocyte cytotoxicity.

In **paper I** we report two families (A, B) each with two siblings with biallelic pathogenic variants in *PRF1*, of which one sibling developed HLH in adolescence and the other one Hodgkin lymphoma in adolescence, but no HLH. Instead, patient G developed both HLH and a non-Hodgkin T cell lymphoma at 22 months of age. In **paper IV** P5, diagnosed with biallelic disease-causing variants in *RAB27A*, developed Hodgkin lymphoma at 13 years of age, after recurrent fever episodes, but did not fulfill the HLH-2004 diagnostic criteria. Of note, the sister of P2 developed brain lymphoma, but, due to lack of material, could not be genetically tested. Previous case reports of patients with biallelic hypomorphic variants in *PRF1* who develop both HLH and lymphoma or leukemias exist (Clementi et al., 2005; Chia et al., 2009). This co-occurrence has also been observed for other genes responsible for primary HLH, such as *STXBP2* (Machaczka et al., 2012). The occurrence of cancer as the only manifestation of an inborn defect of lymphocyte cytotoxicity in patients from **papers II** and **IV** strengthens further the link between perforin-dependent lymphocyte cytotoxicity and susceptibility to lymphoma. Besides a role for a diminished lymphocyte-mediated killing of malignant cells, it can be speculated that a persistent low-grade immune dysfunction and inflammation might promote a microenvironment suitable for tumor development (Carbone et al., 2014). In animal models, perforin-knockout mice develop B-cell lymphoma with age, while infection with LCMV is required to develop HLH (Smyth et al., 2000). In conclusion, biallelic hypomorphic variants in genes associated with primary HLH might represent a cause of familial hematological malignancies.

4.2.1.3 Predominant neurological manifestations

Neurological manifestations are commonly reported in patients with HLH, and they are often present already at diagnosis (Horne et al., 2008; Deiva et al., 2012). In some cases neurological manifestations predominate and the signs of systemic inflammation typical for HLH may be scarce (Meeths et al., 2014). In these cases, achieving the correct diagnosis is hard, also because specialists less familiar with HLH might see these patients. At the same time, achieving a diagnosis is very important also in patients with *neuro-HLH* since treatment can reduce development of additional brain damage and a timely HSCT can even resolve neurological symptoms (Hussein et al., 2014).

In this thesis, five patients with an inborn defect of lymphocyte cytotoxicity and predominant neurological manifestations were studied, two with biallelic missense variants in *PRF1* and three with biallelic variants in *RAB27A*. Three of them, patient B:1 from **paper II**, and

patients P1 and P3 from **paper IV**, developed neurological manifestations several years before the onset of systemic HLH. Indeed, it was only the onset of HLH that triggered the initiation of a diagnostic work-up for HLH. Patient B:1 from **paper II** suffered from recurrent headaches, with vomiting and fever for five years before onset of systemic HLH, while P1 and P3 from **paper IV** suffered from recurring neuroinflammation with acute disseminated encephalomyelitis (ADEM)-like features for several years before the onset of systemic HLH. Because of a relapsing-remitting course of disease, P1 had received a diagnosis of multiple sclerosis. In addition to neuroinflammation, P3 also suffered from granulomatous dermatitis and lung infiltrates. Instead, patient F from **paper II** and P4 from **paper IV** only presented with *neuro-HLH*, without systemic signs of HLH. P4 developed ataxia after varicella infection at the age of 1 year and 9 months. Additional episodes followed, and a biopsy of a brain lesion showed lymphocyte infiltration. A similar presentation of ataxia triggered by infections was reported in two siblings in whom a diagnosis of FHL2 was established only post-mortem through ES (Dias et al., 2013). In clinical practice, an inborn defect of lymphocyte cytotoxicity should be considered in the differential diagnosis of patients with neuroinflammation and functional assays should be performed. The inclusion of genes associated with primary HLH into gene lists for the genetic evaluation of patients with ataxia and seizures might reveal that inborn defects of lymphocyte cytotoxicity is an under-recognized cause of neurologic disease.

4.2.1.4 Genotype-phenotype correlations in atypical primary HLH

Severe perforin deficiency due to truncating variants is associated with very early onset HLH (Trizzino et al., 2008). As observed in **paper II**, biallelic missense variants associate with older age at presentation and atypical manifestations. Interestingly, in **paper II** we found sibling pairs discordant for clinical presentations, suggesting that the phenotypic effect of hypomorphic variants is harder to predict. The majority of the *PRF1* variants reported in **paper II** was previously reported in patients with late-onset HLH, demonstrating a consistent genotype-phenotype correlation. In *in-vitro* evaluations, compared to null variants, these variants were temperature-sensitive, and their function could be restored at 30°C (Chia et al., 2009). In **paper II** we show that compared to FHL2 patients with early onset of disease, NK cells from atypical FHL2 patients recover cytotoxic function after stimulation with IL-2. In **paper IV**, recovery of cytotoxicity and exocytic capacity after IL-2 stimulation is also shown for CD8⁺CD57⁺ T cells and NK cells from P3 and P4. Although the mechanism of IL-2-mediated recovery is not known, this phenomenon has previously been observed in other patients with late-onset or atypical manifestations (Meeths et al., 2010b; Rohr et al., 2010).

All the GS2 patients studied in **paper IV** carry the TSS-disrupting SV in *RAB27A* at least in heterozygous state. Clinically, recurring episodes of neuroinflammation, skin granulomas, late-onset HLH, susceptibility to lymphoma, and lack of skin and hair hypopigmentation define this atypical form of GS2.

4.2.2 Severe autoimmunity in LRBA deficiency

In **paper V** we studied a patient, born to consanguineous parents of Turkish origin, who suffered from several autoimmune manifestations and failure to thrive since the age of seven years. At age seven he developed autoimmune hemolytic anemia, hepatosplenomegaly, autoimmune thyroiditis, and hypogammaglobulinemia requiring intravenous immunoglobulins (IVIG). He then developed immune thrombocytopenia, repeated infections, and a severe enteropathy with colitis, gastritis, and subtotal villus atrophy in the duodenum resembling celiac disease. At the age of 14 the patient developed pancreatitis and neurological symptoms with brain lesions resembling ADEM. A clinical diagnosis of CVID was made. Due to the severe course of disease, the patient received HSCT at the age of 15 years from the HLA-matched mother. With HSCT, a clinical, radiological, and histological resolution of most manifestations was achieved. At the time of writing, the patient is still doing well more than 4 years after HSCT.

To establish a molecular diagnosis ES was performed post-HSCT. We identified a homozygous frameshift variant (c.7162delA, p.T2388Pfs*7) in *LRBA*. Both parents were heterozygous carriers. At the time of publication, this pathogenic variant had recently been reported in two siblings with LRBA deficiency (Seidel et al., 2015). One of those siblings was the first patient with LRBA deficiency for whom a successful outcome after HSCT was reported.

Disease-causing variants in *LRBA* were initially identified in patients with early onset CVID (Lopez-Herrera et al., 2012). Since then, they have also been identified in patients with IPEX-like phenotype (Charbonnier et al., 2015) and multiple autoimmune endocrinopathy, such as neonatal diabetes (Johnson et al., 2017). Review of large cohorts of patients with LRBA deficiency reveals that most patients suffer from immune dysregulation, in the form of enteropathy, autoimmune hemolytic anemia, thrombocytopenia, and type I diabetes (Alkhairy et al., 2016; Gámez-Díaz et al., 2016). Recurrent infections are also common. Some patients develop neurological manifestations. Moreover, ADEM has been reported in patients with CTLA4 deficiency (Schubert et al., 2014).

LRBA deficiency leads to a secondary deficiency of CTLA4 (Lo et al., 2015). This observation has prompted the use of abatacept, a CTLA4-immunoglobulin fusion drug, for treatment of patients with LRBA deficiency, with good responses (Lo et al., 2015). A recent survey of published and unpublished patients with LRBA deficiency revealed that 12 out of 72 cases received a HSCT due to severe manifestations (Seidel et al., 2017). The patient studied in **paper IV** was also included in this survey. Overall, survival post-HSCT was 67% (8/12). Four patients died within 3 months post-transplant due to transplant-related complications. Instead, surviving patients achieved a good level of remission including some complete remissions. This data suggests a curative potential of HSCT for patients with LRBA

deficiency. Seidel *et al.* suggest proceeding to early HSCT in LRBA deficient patients with severe manifestations.

4.3 GAIN-OF-FUNCTION VARIANTS IN *SAMD9L* PREDISPOSE TO MDS/AML WITH MONOSOMY 7

4.3.1 *SAMD9L*-related phenotypes

In **paper VI** we studied two families with variable expressivity of cytopenia, MDS with cytogenetic aberrations of chromosome 7, immunodeficiency, and progressive cerebellar dysfunction. Two individuals from family 1, F1:III-1 and his grandfather F1:I-3, and one individual from family 2, F2:II-4, developed MDS. F1:III-1 and F2:II-4 developed MDS in childhood and F1:I-3 at the age of 56. In family 1, due to familial MDS and cellular immunodeficiency affecting B, NK cells, and monocytes, *GATA2* deficiency was suspected (Spinner *et al.*, 2014). However, no pathogenic variants in *GATA2* were identified. Three individuals suffered from transient cytopenia in childhood, which resolved spontaneously. Several family members, including F1:III-1, F1:I-3, and F2:II-4, suffered from very mild to moderate balance impairment. A summary of the clinical manifestations of affected individuals is provided in Table 5.

ES was performed in family 1, and affected individuals from three generations were analyzed. Rare variants consistent with an autosomal dominant or X-linked recessive inheritance model were reviewed. We identified a heterozygous missense variant, c.2956C>T (p.Arg986Cys) in the gene sterile alpha motif protein 9-like (*SAMD9L*, NM_152703), located on 7q21.2, a region commonly deleted in myeloid malignancies (Asou *et al.*, 2009). By Sanger sequencing, another rare heterozygous missense variant in *SAMD9L* was identified in family 2, c.2672T>C (p.Ile891Thr). Both variants were not present in ExAC, affected evolutionarily conserved amino acid residues, and had damaging *in-silico* predictions. At the time a *SAMD9L* variant was discovered in family 1, no human disease was associated with genetic variants in *SAMD9L*. Instead, haploinsufficiency for the same gene in mouse was associated with development of MDS with age (Nagamachi *et al.*, 2013), making *SAMD9L* an interesting candidate in these families.

In 2016, Chen *et al.* found heterozygous missense pathogenic variants in *SAMD9L* in two families with ataxia-pancytopenia (ATXPC) syndrome (MIM 159550), including the first family sharply described by Li *et al.* in 1978 (Li *et al.*, 1978, 1981; Chen *et al.*, 2016a). All the six affected individuals reported by Li *et al.* displayed cerebellar ataxia, and cerebellar atrophy was reported in five. Two developed AML and other had cytopenias. Similarly, the additional family studied by Chen *et al.* included nine individuals with cerebellar ataxia and variable hematologic abnormalities. Family 2 from **paper VI** had received a clinical diagnosis of ATXPC before a genetic defect was identified. Instead, several individuals from family 1 had only mild neurological manifestations and were mainly followed for their hematologic abnormalities.

Individual	Current age	Cytopenia	Myelodysplasia	Immunodeficiency	Neurological manifestations
F1:I-2	64	Neutropenia and thrombocytopenia at 61 years of age	-	No frequent or unusual infections	Nystagmus, mild balance impairment, lower limb hyperreflexia
F1:I-3	58†	B, NK, and monocytopenia at 54 years of age	MDS with der(1;7)(q10;p10) at 56 years of age. Bone marrow with prominent erythropoiesis and dysplastic features in megakaryocytopoiesis followed by MDS without elevated blast-levels but with increased T cell numbers.	Recurrent bacterial infections requiring hospitalization	Very mild impairment of balance
F1:II-1	38	-	-	-	-
F1:II-2	33	-	-	-	Nystagmus, very mild balance impairment
F1:II-4	28	Tri-lineage cytopenia from 18 months to 3 years of age	-	-	-
F1:III-1	6	Thrombocytopenia followed by tri-lineage cytopenia at 5 months of age	MDS with -7 and -22 at 4 years of age. Severely hypoplastic bone marrow diagnosed as pediatric refractory cytopenia	Congenital CMV infection, hypogamma-globulinemia	Nystagmus, unsteady gait, lower limb stiffness and weakness (after HSCT)
F1:III-2	4	Transient episode of thrombocytopenia at 21 months of age	-	-	-
F2:I-2	34	-	-	Recurrent otitis in childhood, recurrent sinusitis	Nystagmus, memory problems, muscle weakness, cerebellar atrophy
F2:II-1	13	Transient episode of severe aplastic anemia at 14 months of age, after use of chloramphenicol (ear drops)	Bone marrow biopsy showed markedly hypoplastic bone marrow, cellularity was assessed to be 5%	Recurrent otitis from 4 months of age	Nystagmus, ADHD, intermittent vertigo
F2:II-4	9	Thrombocytopenia at 2 months followed by tri-lineage cytopenia at 14 months of age	MDS -7 at 18 months of age. Normal bone marrow cellularity and dysplastic megakaryocytes and clear myelodysplasia in the megakaryopoiesis	Bronchiolitis at 6 weeks and varicella infection at 4 months of age, both requiring hospitalization; recurrent otitis and laryngitis before the HSCT	Nystagmus, balance problems, dysmetria, lower limb stiffness, ADHD

Table 5. Clinical characteristics of the patients studied in **paper VI** (adapted from there).

4.3.2 Effect of *SAMD9L* pathogenic variants

SAMD9L is located on 7q21.2 adjacently to *SAMD9*, a highly conserved paralogous gene. Biallelic pathogenic variants in *SAMD9* cause normophosphatemic familial tumoral calcinosis, an autosomal recessive disease characterized by inflammatory skin lesions with calcium deposits (Topaz et al., 2006; Chefetz et al., 2008). In 2016, heterozygous *de novo* variants in *SAMD9* were identified in individuals with Myelodysplasia, Infection, Restriction of growth, Adrenal hypoplasia, Genital phenotypes (anomalies of external and internal genitalia), and Enteropathy, defining a new syndrome called MIRAGE (MIM 610453) (Narumi et al., 2016). Evidence for a tumor suppressor function exists for both *SAMD9* and *SAMD9L* (Nagamachi et al., 2013; Ma et al., 2014; Wang et al., 2014a). Moreover, *SAMD9* and *SAMD9L* are IFN- γ regulated genes and antiviral activity is documented for *SAMD9* (Pappas et al., 2009; Hershkovitz et al., 2011; Liu and McFadden, 2015). In mouse, Samd9l is localized to endosomes and inhibits growth factor receptor signaling (Nagamachi et al., 2013). Therefore, both genes might have antitumor and antiviral function.

To understand the effect on cell proliferation of the *SAMD9L* mutants Arg986Cys and Ile891Thr, identified in **paper VI**, and of His880Gln and Cys1196Ser, identified by Chen *et al.*, we performed dye dilution assays in 293FT cells transiently transfected with wild-type or mutant N-terminal teal fluorescent protein (TFP)-tagged *SAMD9L*. Compared to cells transfected with vector only, 293FT cells transfected with wild-type TFP-*SAMD9L* showed reduced cell proliferation, consistently with a growth suppressing activity of *SAMD9L*. The inhibition on proliferation was much stronger when 293FT cells were transfected with the TFP-*SAMD9L* mutants, which therefore represent GOF variants (Figure 10A). A similar growth-restricting effect was shown for *SAMD9* variants associated with the MIRAGE syndrome (Narumi et al., 2016). At first, the observation of GOF variants in a tumor suppressor gene appeared paradoxical in the context of a cancer predisposing syndrome.

4.3.3 Predisposition to MDS/AML with monosomy 7

As reviewed in the introduction, Knudson's "2-hit" model for cancer predisposition presupposes the occurrence of two inactivating variants, one germline and one acquired somatic, in a tumor suppressor gene (Knudson, 1971). In **paper VI** the three individuals who developed MDS displayed complete or partial deletion of chromosome 7, on which *SAMD9L* and *SAMD9* are located. Monosomy 7 was present also in all the other reported individuals with *SAMD9L* or *SAMD9* variants who developed MDS/AML (Chen et al., 2016a; Narumi et al., 2016; Buonocore et al., 2017).

Under a classic Knudson's "2-hit" model, a germline predisposing variant in a tumor suppressor gene might display loss-of-heterozygosity (LOH) in tumor cells. To quantify LOH in MDS cells, variant-specific dPCR assays were used. MDS-derived DNA was available from F1:I-3 and F2:II-4. Unexpectedly, the amount of mutated *SAMD9L* alleles was decreased in MDS cells compared to fibroblasts, demonstrating that the germline *SAMD9L*

variants were located on the deleted chromosome 7 (Figure 10B,C). Similarly, germline *SAMD9* variants were depleted in MDS cells of patients with MIRAGE syndrome (Narumi et al., 2016; Buonocore et al., 2017). Therefore, in individuals with germline growth-inhibiting variants in *SAMD9L* or *SAMD9*, those HSC that loose the growth-inhibiting allele through somatic aneuploidy gain a competitive advantage and can expand clonally. Narumi *et al.* described this phenomenon as *adaptation-by-aneuploidy*. Most likely, haploinsufficiency of other genes on 7q and other somatically acquired mutations contribute to the development of a mature MDS clone.

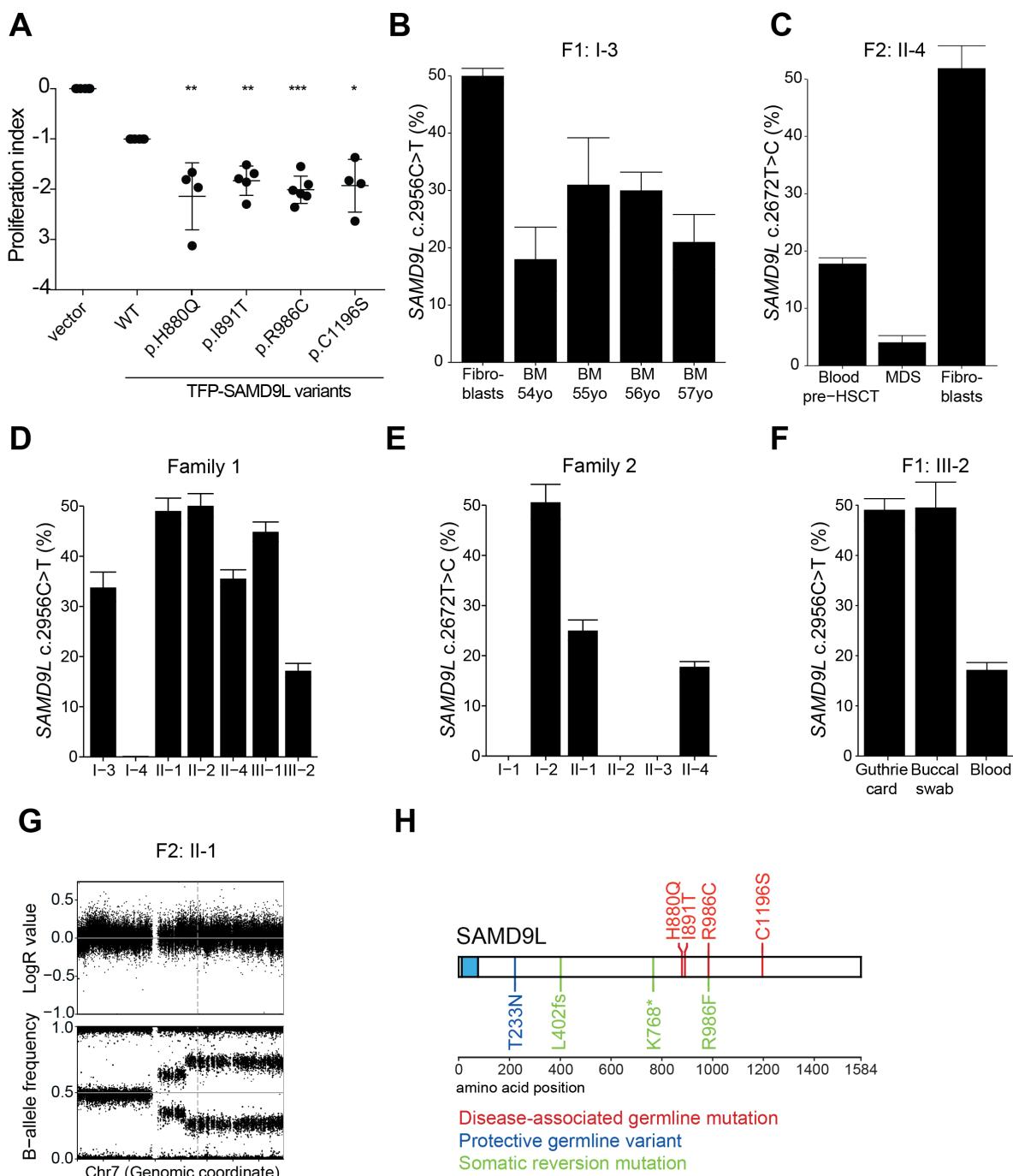


Figure 10. (Previous page) (A) Compilation of results from dye dilution assays performed on 293FT cells transiently transfected (Tx) with TFP-SAMD9L wild-type (WT) or disease-associated variants reported in **paper VI** and by Chen *et al.*. (B) Quantification of *SAMD9L* c.2956C>T variant by variant-specific dPCR in DNA derived from fibroblasts and serial bone marrow samples from F1:I-3. (C) Quantification of *SAMD9L* c.2672T>C variant in DNA derived from MDS, peripheral blood, and fibroblasts from F2:II-4. (D) Quantification of the frequency of the *SAMD9L* c.2956C>T in family 1 and (E) *SAMD9L* c.2672T>C in family 2 in peripheral blood-derived DNA. (F) Quantification of the *SAMD9L* c.2956C>T in DNA from a neonatal dried blood spot and a buccal swab from F1:III-2. For each sample, values represent the mean of at least two chips. Error bars denote 95% confidence levels. (G) SNP array data for chromosome 7 of F2:II-1. The data is displayed as Log R ratio of intensity signal and B-allele frequency. The dashed vertical line indicates the position of the *SAMD9L* c.2672T>C pathogenic variant. (H) Overview of *SAMD9L* structure. The SAM domain is depicted in blue. The position of disease-associated germline *SAMD9L* GoF variants (red), germline LoF variants (blue) and somatic reversion mutations (green) are indicated.

4.3.4 Frequent hematopoietic revertant mosaicism in carriers of *SAMD9L* GOF variants

During Sanger-based segregation studies we observed that in some individuals from family 1 and 2 the mutant allele was underrepresented in peripheral blood-derived germline DNA, suggesting mosaicism. This observation was confirmed by variant-specific dPCR assays. Out of the 10 carriers of either *SAMD9L* variant, the mutant allele was underrepresented in the three individuals with MDS and in four other carriers (F1:I-2, F1:II-4, F1:III-3, and F2:II-1) (Figure 10D,E). The finding in F1:III-3 was concordant with the observation from ES data that the *SAMD9L* variant c.2956C>T was present only in 12% of the sequencing reads. Using ES data from F1:III-3, the allelic frequency of all heterozygous sites on chromosome 7 was evaluated. This analysis revealed LOH of 7q, defined by SNP-arrays as copy neutral (CN)-LOH due to mosaic segmental uniparental disomy (UPD) of 7q in ~80% of cells (Figure 10G). Mosaic UPD(7q) was confirmed also in F1:I-2, F1:II-4, and F2:II-1. In F1:I-2 and F2:II-1 two clones with UPD segments of different lengths, both including *SAMD9L*, were clearly distinguishable (Figure 10G). Variant-specific dPCR assay revealed that the UPD(7q) of F2:II-1 was absent from DNA isolated from fibroblasts and from a dried blood spot collected at birth. This implies that the UPD(7q) was acquired post-natally and was confined to blood, *i.e.* an event of revertant mosaicism in blood cells (Figure 10F). Sorting of bone marrow cells from F1:II-4 showed that the reversion happened during differentiation of HSCs.

Revertant mosaicism is the result of a somatic genetic event that corrects an inborn genetic defect in a cell, followed by clonal expansion of the same cell. Its prevalence is unknown, but it has been reported in bone marrow failure syndromes, PIDs, and in inherited skin diseases. Revertant mosaicism is usually associated with a phenotypic improvement (Hirschhorn, 2003; Forsberg *et al.*, 2017). Different kinds of somatic events can lead to revertant mosaicism. Revertant mosaicism with segmental UPD is most likely the result of mitotic

recombination events, but second-site mutations are also a common mechanism. Indeed, review of available ES data or analysis by Sanger sequencing of the coding region of *SAMD9L* in other carriers revealed second-site mutations in F1:II-4 (c.2957G>T and c.1204_1208delCTCAT), who also displayed UPD(7q), and in F2:I-2, c.2302A>T (p.Lys768Ter) (Figure 10H). The latter was confirmed to be located in *cis* with the disease-causing *SAMD9L* variant c.2972T>C. In total, 5 out of the 7 carriers who have not developed MDS displayed revertant mosaicism. The occurrence of multiple reverted clones in at least three carriers illustrates the strong selective advantage of cells that manage to eliminate the growth-inhibiting effect of germline GOF *SAMD9L* variants. Phenotypically, hematopoietic revertant mosaicism was present in carriers with a history of spontaneously resolved cytopenia in childhood and/or carriers with predominant neurological manifestations, suggesting a milder course of disease. Even individuals with MIRAGE syndrome and *SAMD9* GOF variants display hematopoietic revertant mosaicism leading to increased survival (Buonocore et al., 2017). In a broader context, cancer predisposition due to *aneuploidy-by-adaptation* can be interpreted as the phenotypic result of revertant mosaicism and growth advantage of reverted cells. Moreover, hematopoietic revertant mosaicism represents by definition a form of clonal hematopoiesis, which confers an increased risk for hematological malignancies (Jaiswal et al., 2014).

Two carriers from family 1 (F1:II-1 and F1:II-2), with minimal or no signs of disease, did not display hematopoietic revertant mosaicism. Notably, both carried a germline *in-trans* rare missense variant in *SAMD9L* c.698C>A (p.Thr233Asn). Dye dilution assays performed on 293FT cells after transfection with the TFP-tagged *SAMD9L* Thr233Asn mutant showed that the Thr233Asn *SAMD9L* variant has a LOF effect and promotes cell proliferation. Germline compound heterozygosity for a GOF and a LOF variant in *SAMD9L* might protect from disease and diminish the selective advantage of reverted clones.

Phenomena like cancer predisposition through *adaptation-by-aneuploidy* and a high frequency of revertant mosaicism make the genes *SAMD9L* and *SAMD9* very interesting, especially considering that their function is largely unknown. In **paper VI** we confirm IFN-regulation for both *SAMD9L* and *SAMD9*, even in CD34⁺ cells. This finding, coupled with the observation that onset of cytopenia in carriers was often preceded by infections, suggests a role for these genes in the regulation of HSC proliferation during demand-adapted hematopoiesis. Efforts to prove this hypothesis and elucidate the function of *SAMD9* and *SAMD9L* are underway. Based on the strong inhibition of proliferation displayed by gain-of-function mutations in these genes, the identification of compounds able to perturb expression and/or function of *SAMD9L* and *SAMD9* may also prove useful in cancer therapy.

In clinical settings, it will be important to understand how large proportion of pediatric and adult cases MDS/AML can be attributed to germline mutations in *SAMD9L* or *SAMD9*. Long-term follow up of carriers of *SAMD9L* mutation with hematopoietic reversion will be

required to understand progression of disease and define appropriate strategies for clinical monitoring.

5 CONCLUSIONS

The main conclusions of this thesis are summarized below.

- About 44% of pediatric patients fulfilling HLH-2004 diagnostic criteria have biallelic pathogenic variants in genes associated with hemophagocytic lymphohistiocytosis (HLH). Targeted sequencing through a high-throughput sequencing-based gene panel is a suitable method for genetic diagnostics of primary HLH.
- Combined functional assays for diagnostics of primary HLH, especially CD8⁺CD57⁺ T-cells and NK cells degranulation assay and intracellular staining of perforin, have high positive and negative predictive values. A molecular diagnosis is found for virtually all patients with a functional defect highly typical for primary HLH. Non-coding pathogenic variants should be suspected in molecularly undiagnosed patients with a functional defect highly typical for primary HLH.
- Rare monoallelic variants with *in-silico* possibly damaging predictions in genes associated with primary HLH are present in 25% of individuals in the 1000G cohort. Thus, caution is required when assessing the causality of monoallelic variants in HLH-associated genes in patients with secondary HLH.
- Biallelic *PRF1* missense pathogenic variants are associated with late-onset HLH, susceptibility to lymphoma, predominant neurological manifestations, and gastrointestinal inflammation. Intra- and inter-familial phenotypic variability is a feature of perforin-deficiency due to biallelic *PRF1* missense mutations.
- A non-coding, complex structural variant (SV) disrupting the transcriptional start site (TSS) of *RAB27A* explains atypical cases of Griscelli syndrome type 2 (GS2) in the Nordic and Baltic population. The identification of a founder mutation ensures access to a molecular diagnosis for future patients from this area.
- This atypical form of GS2, with a shared genetic defect, is characterized by recurrent episodes of neuroinflammation, skin granulomas, late-onset HLH, susceptibility to lymphoma, and lack of skin and hair hypopigmentation.
- Differential TSS usage between cytotoxic lymphocytes and melanocytes explains the lack of hypopigmentation in patients with a non-coding, complex SV disrupting the TSS of the *RAB27A* transcript NM_183235.2.
- Lymphoma or neurological manifestations may be the only phenotypic manifestation of an inborn defect of lymphocyte cytotoxicity. Even in patients with atypical manifestations of inborn defects of lymphocyte cytotoxicity, functional assays are constitutively abnormal and can aid diagnostics. Recovery of cytotoxic function after IL-2 stimulation is commonly observed in atypical cases. Similarly, genes associated

with primary HLH should be included in the genetic investigation of familial hematological malignancies, as well as of patients with unexplained ataxia and seizures associated with inflammation.

- Fatal HLH can develop in patients with IFN- γ receptor-deficiency due to pathogenic variants in the genes *IFNGR2* and *IFNGR1* upon severe mycobacterial infections. This finding suggests IFN- γ -independent mechanisms for the development of HLH, of importance in the context of current attempts to treat HLH with anti-IFN- γ therapy.
- HLH can represent the first clinical manifestation of an underlying primary immunodeficiency (PID) other than inborn defects of lymphocyte cytotoxicity. Exome or whole-genome sequencing might be required to exclude an underlying PID.
- HSCT should be considered in patients with severe LRBA deficiency. LRBA deficiency should be suspected in patients with several autoimmune manifestations, including early-onset inflammatory bowel disease.
- Autosomal-dominant *SAMD9L* gain-of-function variants cause a syndrome characterized by cytopenia, myeloid malignancies involving chromosome 7 aberrations, immunodeficiency, and cerebellar dysfunction, with marked inter- and intrafamilial clinical variability.
- Gain-of-function (GOF) *SAMD9L* variants strongly inhibit cell proliferation, confirming that *SAMD9L* acts as tumor suppressor. Somatic restoration of a wild-type *SAMD9L* function induces a strong selective advantage in a hypoplastic bone marrow.
- Predisposition to myeloid malignancies in patients with *SAMD9L* GOF variants does not follow the classic “2-hit” model for cancer predisposition. Myeloid malignancies with monosomy 7 or partial deletion of 7q result from *adaptation-by-aneuploidy* events: HSCs loose the chromosome 7 carrying the growth-inhibiting *SAMD9L* variant to restore a normal level of cell proliferation.
- Besides aneuploidy, hematopoietic revertant mosaicism through mitotic recombination and second-site mutations is a frequent post-natal event in carriers of *SAMD9L* GOF variants and is associated with milder manifestations of disease.

5.1 SNAPSHOT OF THE FIELD IN 2017

Four years after the initiation of this PhD project, in virtually all patients with functional defects characteristic for an inborn defect of lymphocyte cytotoxicity coding or non-coding pathogenic variants in genes associated with primary HLH are now identified (**paper I, IV**, unpublished data, Cetica et al., 2016). This suggests it is unlikely that a *bona fide* new genomic locus for primary HLH will be discovered. Nonetheless, it has become clear that the group of so-called secondary HLH include several patients with genetic defects other than inborn errors of lymphocyte cytotoxicity, such as known and novel PIDs, and metabolic disorders (**paper III**, unpublished data, Canna et al., 2014; Bode et al., 2015). At the same time, the phenotypic spectrum of primary HLH is getting broader, as shown in **paper II** and **IV** and by others, requiring attention by different medical specialties. Within basic science, studies of murine models of HLH have revealed that CD8⁺ T cells and NK cells have non-redundant roles in the pathogenesis of HLH (Sepulveda et al., 2015), and proposed novel strategies for targeted treatment (Das et al., 2016; Maschalidi et al., 2016). However, the genetic basis of secondary HLH is still largely unknown. Variants in genes associated with primary HLH were detected in patients with secondary HLH (Kaufman et al., 2014; Schulert et al., 2016). Murine models seem to suggest an additive effect of mutations in different genes associated with primary HLH is possible (Sepulveda et al., 2016). Interestingly, an increased risk of cancer has been shown for likely heterozygous carriers of HLH-causing mutations (Löfstedt et al., 2015). Further quantitative approaches, such as the one used in **paper I**, are required to understand genetic predisposition to secondary HLH in a genome-wide dimension.

As expected, the use of HTS in the field of PIDs has been tremendously useful to identify new inborn defects of immunity. In four years, about 50 new PIDs have been described. Moreover, access to ES/WGS has revealed a large degree of phenotypic overlap between different disorders (Meyts et al., 2016). More than 70 patients with LRBA deficiency have been reported and, due to an increased understanding of the pathogenesis of this disorder, a targeted treatment with a CTLA4-immunoglobulin fusion drug is now available (Lo et al., 2015). Case reports, as **paper V**, and international surveys are providing evidence for a beneficial effect of HSCT in severely affected patients with LRBA deficiency (Seidel et al., 2017).

In the field of hereditary myeloid malignancies and PIDs, several patients with GATA2 deficiency have been reported, delineating a broad spectrum of phenotypes associated with germline mutations in *GATA2*, with a clear overlap between immunodeficiency and risk for MDS/AML (Spinner et al., 2014). Moreover, germline mutations were found in up to 37% of pediatric MDS cases with monosomy 7 (Włodarski et al., 2016), revealing that germline sequencing is indicated for children with MDS with monosomy 7. The increased attention and clinical relevance of genetic predisposition to myeloid malignancies is also demonstrated by the inclusion of inherited myeloid malignancies as a separate group in the WHO 2016

classification of tumors of the hematopoietic and lymphoid tissues (Arber et al., 2016). We and others have revealed new genetic predisposition to MDS/AML with monosomy 7 due to germline mutations in *SAMD9L* characterized by a high frequency of hematopoietic revertant mosaicism (**paper VI**, Chen et al., 2016a) and *SAMD9* (Narumi et al., 2016; Buonocore et al., 2017). This finding may explain additional familial and sporadic cases of MDS with monosomy 7.

The purpose of the above section was to give a quick overview of the development seen in the field of HLH, PID and inherited myeloid malignancies during the last four years. One page is of course not enough to capture four years of science! A selection has been made with the intention to present the most relevant literature in relation to the studies included in this thesis.

6 FUTURE PERSPECTIVES

“JACK: Yes, but you said yourself that a severe chill was not hereditary.

ALGERNON: It usen’t to be, I know - but I daresay it is now. Science is always making wonderful improvements in things.”

Oscar Wilde, The Importance of Being Earnest, 1895

In lack of a molecular diagnosis consistent with primary HLH, patients usually receive the umbrella diagnosis of secondary HLH, which includes disorders heterogeneous for underlying etiologies, pathogenic mechanisms, and therapy requirements. A small proportion of secondary HLH cases can be *solved* through ES/WGS and reclassified as other PIDs or metabolic diseases. However, the genetic basis of *bona fide* secondary HLH is most likely polygenic and multi-factorial (Meeths and Bryceson, 2016). Moreover, since secondary HLH can be triggered by different underlying conditions and triggers are not always easily identified, the design of well-controlled and well-powered genome-wide studies is harder to achieve compared to other diseases. Yet, it would be interesting to see an international collaboration form to tackle this question. Genetic studies combined with fine-tuned functional characterization of patients with secondary HLH, through transcriptome analysis or deep immunophenotyping, might also help to subgroup patients and identify gene pathways important for disease pathogenesis. Here, an additional problem might be how to discern real HLH-related changes from the effect of the underlying disease and of the ongoing treatment. Efforts to establish predictive biomarkers to identify patients at risk to develop secondary HLH are also needed.

Paradoxically, useful answers might come from studies of healthy or unselected individuals. Strategies based on combined deep immunophenotyping and GWAS have already proved powerful to dissect heritability of immune traits (Brodin and Davis, 2017). Extended to functional phenotyping of CD8⁺ T cells and NK cells, this strategy might reveal other genetic factors, likely with smaller effect size, important for the control of lymphocyte cytotoxicity. Phenome-wide association studies (PheWAS), based on biorepositories with access to genetic data and linked electronic health records (Bush et al., 2016), can provide a powerful resource to understand the impact of genetic variation in HLH-related genes on susceptibility to cancer and other inflammatory diseases besides HLH. If coupled with quantification of lymphocyte cytotoxicity, PheWAS can also inform about penetrance and phenotypic spectrum of decreased lymphocyte cytotoxicity.

In the recent years one of the focus of human genetics has been to discover the genetic basis of Mendelian diseases. However, for all Mendelian genes to be discovered, it will be necessary to overcome some of the existing bottlenecks in the studies of rare diseases (Boycott et al., 2017). One bottleneck is represented by the lack of infrastructure for sharing

of genetic data. This makes the collection of several patients with the same genetic defect slow. Current solutions are represented by social medias, which are more and more popular among patients and families as a way to connect individuals with the same genetic defect, and by services like Matchmaker Exchange (Philippakis et al., 2015). However, long-term solutions are needed. At the same time, existing knowledge gaps in our ability to interpret genetic variants hamper gene discovery. In diseases like HLH and other PIDs, the existence of robust functional or immunophenotypic read-outs to assess pathogenicity is extremely valuable. Yet, such diagnostics tests do not exist for the majority of rare diseases. This leads to the generation of thousands of coding variants of uncertain significance and to inconclusive genetic tests whose result cannot be used for patient care. With many patients undergoing WGS, a better understanding of the effect of non-coding variations is also needed to enable analysis of these regions in clinical settings.

In conclusions, this is a very exciting time for the field of human genetics! Sequencing technologies are now very mature and the initial genetic screening for diagnostic purposes is increasingly performed in clinical settings. This creates an opportunity for researchers to move from the *hypothesis-free* phase of the recent years, focused on gene discovery, to more *hypothesis-driven* research aimed to understand the biology of these new and old Mendelian disorders with the hope to soon be able to deliver to our patients not just a diagnosis, but a deep understanding of their diseases and targeted treatment opportunities.

7 ACKNOWLEDGEMENTS

My research would not have been possible without the cooperation of **patients, their families and their physicians**.

I was introduced to the stark reality of HLH by **Jan-Inge Henter**, my main supervisor. He created a ring of knowledge and opportunities around me and gave precious freedom to tackle the disease according to my own ideas and inclinations. The results I have achieved are in direct proportion to his care and encouragement.

Co-supervisor **Yenan Bryceson** always raised the bar a little higher. His trust and confidence in my abilities were often greater than my own. My PhD journey has been a great learning experience thanks to his constructive feedback.

Magnus Nordenskjöld never lost an opportunity to infuse humor into our conversations. His open-door attitude and prompt action contributed to a worry-free work environment. He guided me instrumentally in my starting steps as a researcher and doctor and I cannot thank him enough for that.

Marie Meeths is a good listener. Her ability to reframe my problems positively was close to magic. I have enjoyed our collaboration as scientists in CMM and colleagues at the Clinical Genetic unit.

I too have seen the joys and sorrows of producing papers. I express sincere gratitude to these co-authors for having made the experience lighter: the devoted flow-cytometer master **Samuel Chiang**; molecular biology and SAMD9L tour-guide **Matthias Voss**; **Ulf Tedgård**, **Elisa Rahikkala**, **Joao Farela Neves** and **Elena Sieni** who scouted interesting families for our lab; **Jelena Rascon**, my fellow-detective in the Puzzle of the Lithuanian Patients; **Kristina Lagerstedt** for all things Ion Torrent.

I am proud of my students **Viktoria Lundbring** and **Eya Ben Bdira**. Supervising their projects was a humbling experience and a time-machine trip back to my own young student days.

Colleagues at CMM hold a special place in my heart. **Vasilios “Vasse” Zachariadis** was my go-to reviewer of nascent thoughts and R-code. Our incessant conversations about science have honed my critical thinking. Kind **Nina Jäntti** taught me how smiling eases lab-work and laughing fights troubles. My work has benefited a lot from **Fulya Taylan**, a first-rate novel gene hunter and walking encyclopedia. **Miriam Entesarian** gently guided me when I started at CMM. **Raquel Vaz** spread cheer in English, Swedish, and Nespresso-ish. **Alice Costantini**, **Elisa Martini**, and **Miriam Armenio** gave much needed Italian companionship. Not to forget the rest of the group - **Wolfgang**, **Jesper**, **Ellika**, **Benedicte**, **Maria**, **Anders**, **Anna H**, **Josephine**, **Johanna W**, **Ingegerd**, **Ameli**, **Tobias**, **Stanley**, **Sofia**, **Malin**, **Anh**

Nhi, Britt-Marie, Daniel, Johanna L, Agne, Måns, Jessada, Christina, Samina, Aron, Noor - who shaped a pleasant, helpful workplace.

Ann Nordgren and **Anna Lindstrand** welcomed me to all activities of the Rare Diseases group. The summer outing in Fiskebäckskil is among my fondest memories. It was the ideal combination of science and sea.

“Sugar-kick for the tongue, protein-kick for the mind” was how I began Monday mornings with colleagues from Yenan’s group. **Jelve Zendegani** was a wizard with cakes and samples. **Jakob Theorell** always reminded me that data analysis is complicated. **Tim Holmes** cracked jokes and sorted cells in equal measure. It was a pleasure working with **Heinrich Schlums** on the Immunity paper. **Giovanna Perinetti Casoni** moved with the energy and spirit of a racecar. I have enjoyed working and hanging out with the entire crew including **Frank, Hongya, Lamberto, Martha-Lena, Misty, Stephanie, Donatella, Irene, Beatrice, Tak, Ram, Tamara, Sigrun, Angel, Saeed, Sam, and Matthias**.

Each academic year kicked off with refreshing retreats with colleagues from Jan-Inge’s group. **Tatiana Greenwood, Alexandra Löfstedt, AnnaCarin Horne, and Ida Hed Myrberg** enthusiastically shared their progress in understanding secondary HLH while **Egle, Nikolas, Magda, Selma, and Nikolaos** did the same for LCH. **Yini Wang** supported with crucial lab-work. **Désirée Gavhed** and **Elisabet Bergsten** provided prompt administrative help. **Egle Kvedaraite** and I cooled off after work in bars and parks.

I met a host of helpful colleagues when I started working in the Clinical Genetics Unit. **Peter and Maritta** gave solid introduction to the role of the DNA-doctor. From day one, I felt confident flying under the wing of my clinical supervisor, **Emma Tham, Nadja Pekkola Pacheco** and **Karin Wallander** helped navigate new waters. I look forward to shaping the future of clinical genetics with them and the other junior doctors **Josephine, Cecilia, Sofia, Emma E, and Johanna W!**

We live in an age where knowledge is shared freely and the pain taken to create it is easily forgotten. I am deeply indebted to bioinformatics and open-source software communities for their selfless contributions and troubleshooting guidance.

Thanks to generous support from Jan-Inge, Barncancerfonden, the Karolinska Institutet, the Jeffrey Modell Foundation, the ESID society, Okinawa University, and the Wellcome Genome Center, I was able to attend great courses and conferences all over the world. The awakening and growth I have experienced from these visits are most tangible.

All roads lead not to Rome, but to Florence in my case. **Sabrina Giglio** at the Meyer Children Hospital nurtured my long-lasting interest in clinical genetics. **Aldesia Provenzano** trained me in the art of laboratory techniques. I often think about these early days to remind myself about the extent of my journey.

Friends **Linda, Francesca, Fabio, Laura, Lucia, Chiara, Ilaria, Edoardo**, and **Chiara S** have shown precious care and love despite the distance to Sweden. *Grazie, amici mei! Vi voglio bene!* Thank you **Siddharth** for lasting friendship.

Indrani, Muddu, Aloka and **Naomi** — all part of my extended family showered me with affection and encouragement whenever we met.

My parents **Alberto** and **Nicoletta** used nothing but love and example to teach me the value of education, integrity, hard work, and discipline. Our close-knit communication during my time in Karolinska has been a source of immense serenity and strength. A big hug goes to my brother **Andrea** who inspired me time and again with his courage and dedication as a researcher.

Words are inadequate to reflect on the golden support given by my life-partner **Ananya**. He held my hand throughout this journey, equipped me with mental tools to cope with the occasional harshness of academic life, and made sure small and big successes were celebrated.

8 REFERENCES

- Abouelhoda, M., Faquih, T., El-Kalioby, M., and Alkuraya, F.S. (2016). Revisiting the morbid genome of Mendelian disorders. *Genome Biology* *17*, 235.
- Acuto, O., and Michel, F. (2003). CD28-mediated co-stimulation: A quantitative support for TCR signalling. *Nat Rev Immunol* *3*, 939–951.
- Adès, L., Itzykson, R., and Fenaux, P. (2014). Myelodysplastic syndromes. *Lancet* *383*, 2239–2252.
- Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. *Nat Meth* *7*, 248–249.
- Agrawal, A., Eastman, Q.M., and Schatz, D.G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* *394*, 744–751.
- Aguirre-Gamboa, R., Joosten, I., Urbano, P.C.M., van der Molen, R.G., van Rijssen, E., van, Cranenbroek, B., Oosting, M., Smeekens, S., Jaeger, M., Zorro, M., et al. (2016). Differential Effects of Environmental and Genetic Factors on T and B Cell Immune Traits. *Cell Reports* *17*, 2474–2487.
- Al-Mousa, H., Abouelhoda, M., Monies, D.M., Al-Tassan, N., Al-Ghonaium, A., Al-Saud, B., Al-Dhekri, H., Arnaout, R., Al-Muhsen, S., Ades, N., et al. (2016). Unbiased targeted next-generation sequencing molecular approach for primary immunodeficiency diseases. *J. Allergy Clin. Immunol.* *137*, 1780–1787.
- Alkhairy, O.K., Perez-Becker, R., Driessen, G.J., Abolhassani, H., van Montfrans, J., Borte, S., Choo, S., Wang, N., Tesselaar, K., Fang, M., et al. (2015). Novel mutations in TNFRSF7/CD27: Clinical, immunologic, and genetic characterization of human CD27 deficiency. *Journal of Allergy and Clinical Immunology* *136*, 703–712.e10.
- Alkhairy, O.K., Abolhassani, H., Rezaei, N., Fang, M., Andersen, K.K., Chavoshzadeh, Z., Mohammadzadeh, I., El-Rajab, M.A., Massaad, M., Chou, J., et al. (2016). Spectrum of Phenotypes Associated with Mutations in LRBA. *J Clin Immunol* *36*, 33–45.
- Ameur, A., Dahlberg, J., Olason, P., Vezzi, F., Karlsson, R., Martin, M., Viklund, J., Kähäri, A.K., Lundin, P., Che, H., et al. (2017). SweGen: A whole-genome data resource of genetic variability in a cross-section of the Swedish population. *Eur J Hum Genet*.
- Arber, D.A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M.J., Beau, M.M.L., Bloomfield, C.D., Cazzola, M., and Vardiman, J.W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* *127*, 2391–2405.
- Artis, D., and Spits, H. (2015). The biology of innate lymphoid cells. *Nature* *517*, 293–301.

Asou, H., Matsui, H., Ozaki, Y., Nagamachi, A., Nakamura, M., Aki, D., and Inaba, T. (2009). Identification of a common microdeletion cluster in 7q21.3 subband among patients with myeloid leukemia and myelodysplastic syndrome. *Biochemical and Biophysical Research Communications* *383*, 245–251.

Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* *12*, 745–755.

Bergsten, E., Horne, A., Aricó, M., Astigarraga, I., Egeler, R.M., Filipovich, A.H., Ishii, E., Janka, G., Ladisch, S., Lehmburg, K., et al. (2017). Confirmed efficacy of etoposide and dexamethasone in HLH treatment: Long term results of the cooperative HLH-2004 study. *Blood* *blood-2017-06-788349*.

Blair, D.R., Lytle, C.S., Mortensen, J.M., Bearden, C.F., Jensen, A.B., Khiabanian, H., Melamed, R., Rabadan, R., Bernstam, E.V., Brunak, S., et al. (2013). A Nondegenerate Code of Deleterious Variants in Mendelian Loci Contributes to Complex Disease Risk. *Cell* *155*, 70–80.

Bode, S.F., Ammann, S., Al-Herz, W., Bataneant, M., Dvorak, C.C., Gehring, S., Gennery, A., Gilmour, K.C., Gonzalez-Granado, L.I., Groß-Wieltsch, U., et al. (2015). The syndrome of hemophagocytic lymphohistiocytosis in primary immunodeficiencies: Implications for differential diagnosis and pathogenesis. *Haematologica* *100*, 978–988.

Boisson, B., Quartier, P., and Casanova, J.-L. (2015). Immunological loss-of-function due to genetic gain-of-function in humans: Autosomal dominance of the third kind. *Curr. Opin. Immunol.* *32*, 90–105.

Booth, C., Gilmour, K.C., Veys, P., Gennery, A.R., Slatter, M.A., Chapel, H., Heath, P.T., Steward, C.G., Smith, O., O'Meara, A., et al. (2010). X-linked lymphoproliferative disease due to SAP/SH2D1A deficiency: A multicenter study on the manifestations, management and outcome of the disease. *Blood* *117*, 53–62.

Bousfiha, A., Jeddane, L., Al-Herz, W., Ailal, F., Casanova, J.-L., Chatila, T., Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Franco, J.L., et al. (2015). The 2015 IUIS Phenotypic Classification for Primary Immunodeficiencies. *J. Clin. Immunol.* *35*, 727–738.

Bousfiha, A.A., Jeddane, L., Ailal, F., Benhsaien, I., Mahlaoui, N., Casanova, J.-L., and Abel, L. (2013). Primary immunodeficiency diseases worldwide: More common than generally thought. *J. Clin. Immunol.* *33*, 1–7.

Boycott, K.M., Rath, A., Chong, J.X., Hartley, T., Alkuraya, F.S., Baynam, G., Brookes, A.J., Brudno, M., Carracedo, A., Dunnen, J.T. den, Dyke SOM, Estivill X, et al. (2017). International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *The American Journal of Human Genetics* *100*, 695–705.

Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnipotent. *Cell* *169*, 1177–1186.

Brand, H., Collins, R.L., Hanscom, C., Rosenfeld, J.A., Pillalamarri, V., Stone, M.R., Kelley, F., Mason, T., Margolin, L., Eggert, S., et al. (2015). Paired-Duplication Signatures Mark Cryptic Inversions and Other Complex Structural Variation. *The American Journal of Human Genetics* *97*, 170–176.

Brenner, null (1998). Refuge of spandrels. *Curr. Biol.* *8*, R669.

Brodin, P., and Davis, M.M. (2017). Human immune system variation. *Nat Rev Immunol* *17*, 21–29.

Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C.J.L., Furman, D., Shen-Orr, S., Dekker, C.L., Swan, G.E., Butte, A.J., et al. (2015). Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. *Cell* *160*, 37–47.

Broz, P., and Monack, D.M. (2013). Newly described pattern recognition receptors team up against intracellular pathogens. *Nat Rev Immunol* *13*, 551–565.

Bryceson, Y.T., March, M.E., Barber, D.F., Ljunggren, H.-G., and Long, E.O. (2005). Cytolytic granule polarization and degranulation controlled by different receptors in resting NK cells. *J Exp Med* *202*, 1001–1012.

Bryceson, Y.T., March, M.E., Ljunggren, H.-G., and Long, E.O. (2006). Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion. *Blood* *107*, 159–166.

Bryceson, Y.T., Chiang, S.C., Darmanin, S., Fauriat, C., Schlums, H., Theorell, J., and Wood, S.M. (2011). Molecular Mechanisms of Natural Killer Cell Activation. *Journal of Innate Immunity* *3*, 216–226.

Bryceson, Y.T., Pende, D., Maul-Pavicic, A., Gilmour, K.C., Ufheil, H., Vraetz, T., Chiang, S.C., Marcenaro, S., Meazza, R., Bondzio, I., et al. (2012). A prospective evaluation of degranulation assays in the rapid diagnosis of familial hemophagocytic syndromes. *Blood* *119*, 2754–2763.

Buelow, B.J., Routes, J.M., and Verbsky, J.W. (2014). Newborn screening for SCID: Where are we now? *Expert Rev Clin Immunol* *10*, 1649–1657.

Buonocore, F., Kühnen, P., Suntharalingham, J.P., Valle, I.D., Digweed, M., Stachelscheid, H., Khajavi, N., Didi, M., Brady, A.F., Blankenstein, O., et al. (2017). Somatic mutations and progressive monosomy modify *SAMD9*-related phenotypes in humans. *J Clin Invest* *127*, 1700–1713.

Bush, W.S., Oetjens, M.T., and Crawford, D.C. (2016). Unravelling the human genome–phenome relationship using genome-wide association studies. *Nat Rev Genet* *17*, 129–145.

Busque, L., Mio, R., Mattioli, J., Brais, E., Blais, N., Lalonde, Y., Maragh, M., and Gilliland, D.G. (1996). Nonrandom X-inactivation patterns in normal females: Lyonization ratios vary with age. *Blood* *88*, 59–65.

Canna, S.W., Wrobel, J., Chu, N., Kreiger, P.A., Paessler, M., and Behrens, E.M. (2013). Interferon- γ Mediates Anemia but Is Dispensable for Fulminant Toll-like Receptor 9-Induced Macrophage Activation Syndrome and Hemophagocytosis in Mice. *Arthritis & Rheumatism* *65*, 1764–1775.

Canna, S.W., de Jesus, A.A., Gouni, S., Brooks, S.R., Marrero, B., Liu, Y., DiMattia, M.A., Zaal, K.J.M., Sanchez, G.A.M., Kim, H., et al. (2014). An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. *Nat Genet* *46*, 1140–1146.

Carbone, A., Tripodo, C., Carlo-Stella, C., Santoro, A., and Gloghini, A. (2014). The Role of Inflammation in Lymphoma. SpringerLink 315–333.

Carr, E.J., Dooley, J., Garcia-Perez, J.E., Lagou, V., Lee, J.C., Wouters, C., Meyts, I., Goris, A., Boeckxstaens, G., Linterman, M.A., et al. (2016). The cellular composition of the human immune system is shaped by age and cohabitation. *Nat Immunol* *17*, 461–468.

Casanova, J.-L., and Abel, L. (2005). Inborn errors of immunity to infection. *Journal of Experimental Medicine* *202*, 197–201.

Casanova, J.-L., Conley, M.E., Seligman, S.J., Abel, L., and Notarangelo, L.D. (2014). Guidelines for genetic studies in single patients: Lessons from primary immunodeficiencies. *J Exp Med* *211*, 2137–2149.

Cetica, V., Hackmann, Y., Grieve, S., Sieni, E., Ciambotti, B., Coniglio, M.L., Pende, D., Gilmour, K., Romagnoli, P., Griffiths, G.M., et al. (2015). Patients with Griscelli syndrome and normal pigmentation identify RAB27A mutations that selectively disrupt MUNC13-4 binding. *Journal of Allergy and Clinical Immunology* *135*, 1310–1318.e1.

Cetica, V., Sieni, E., Pende, D., Danesino, C., De Fusco, C., Locatelli, F., Micalizzi, C., Putti, M.C., Biondi, A., Fagioli, F., et al. (2016). Genetic predisposition to hemophagocytic lymphohistiocytosis: Report on 500 patients from the Italian registry. *J. Allergy Clin. Immunol.* *137*, 188–196.e4.

Charbonnier, L.-M., Janssen, E., Chou, J., Ohsumi, T.K., Keles, S., Hsu, J.T., Massaad, M.J., Garcia-Lloret, M., Hanna-Wakim, R., Dbaibo, G., et al. (2015). Regulatory T-cell deficiency and immune dysregulation, polyendocrinopathy, enteropathy, X-linked-like disorder caused by loss-of-function mutations in LRBA. *Journal of Allergy and Clinical Immunology* *135*, 217–227.e9.

Chefetz, I., Amitai, D.B., Browning, S., Skorecki, K., Adir, N., Thomas, M.G., Kogleck, L., Topaz, O., Indelman, M., Uitto, J., et al. (2008). Normophosphatemic Familial Tumoral

Calcinosis Is Caused by Deleterious Mutations in SAMD9, Encoding a TNF- α Responsive Protein. *Journal of Investigative Dermatology* 128, 1423–1429.

Chen, D.-H., Below, J.E., Shimamura, A., Keel, S.B., Matsushita, M., Wolff, J., Sul, Y., Bonkowski, E., Castella, M., Taniguchi, T., et al. (2016a). Ataxia-Pancytopenia Syndrome Is Caused by Missense Mutations in SAMD9L. *The American Journal of Human Genetics* 98, 1146–1158.

Chen, R., Shi, L., Hakenberg, J., Naughton, B., Sklar, P., Zhang, J., Zhou, H., Tian, L., Prakash, O., Lemire, M., et al. (2016b). Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nat Biotech* 34, 531–538.

Chia, J., Yeo, K.P., Whisstock, J.C., Dunstone, M.A., Trapani, J.A., and Voskoboinik, I. (2009). Temperature sensitivity of human perforin mutants unmasks subtotal loss of cytotoxicity, delayed FHL, and a predisposition to cancer. *Proceedings of the National Academy of Sciences* 106, 9809–9814.

Chiang, S.C.C. 2015. Functional assays for the diagnosis of primary defects in lymphocyte cytotoxicity (Inst för medicin, Huddinge / Dept of Medicine, Huddinge).

Chiang, S.C.C., and Bryceson, Y.T. (2016). Measurement of NK Cell Phenotype and Activity in Humans. 300–309.

Chiang, C., Layer, R.M., Faust, G.G., Lindberg, M.R., Rose, D.B., Garrison, E.P., Marth, G.T., Quinlan, A.R., and Hall, I.M. (2015). SpeedSeq: Ultra-fast personal genome analysis and interpretation. *Nat Meth* 12, 966–968.

Chiang, S.C.C., Theorell, J., Entesarian, M., Meeths, M., Mastafa, M., Al-Herz, W., Frisk, P., Gilmour, K.C., Ifversen, M., Langenskiöld, C., et al. (2013). Comparison of primary human cytotoxic T-cell and natural killer cell responses reveal similar molecular requirements for lytic granule exocytosis but differences in cytokine production. *Blood* 121, 1345–1356.

Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *The American Journal of Human Genetics* 97, 199–215.

Churpek, J.E., Pyrtel, K., Kanchi, K.-L., Shao, J., Koboldt, D., Miller, C.A., Shen, D., Fulton, R., O’Laughlin, M., Fronick, C., et al. (2015). Genomic analysis of germline and somatic variants in familial myelodysplasia/acute myeloid leukemia. *Blood* 125, 6411–641100.

Cichocki, F., Schlums, H., Li, H., Stache, V., Holmes, T., Lenvik, T.R., Chiang, S.C.C., Miller, J.S., Meeths, M., Anderson, S.K., et al. (2014). Transcriptional regulation of Munc13-4 expression in cytotoxic lymphocytes is disrupted by an intronic mutation associated with a primary immunodeficiency. *Journal of Experimental Medicine* 211, 1079–1091.

Clamp, M., Fry, B., Kamal, M., Xie, X., Cuff, J., Lin, M.F., Kellis, M., Lindblad-Toh, K., and Lander, E.S. (2007). Distinguishing protein-coding and noncoding genes in the human genome. *PNAS* *104*, 19428–19433.

Clementi, R., Locatelli, F., Dupré, L., Garaventa, A., Emmi, L., Bregni, M., Cefalo, G., Moretta, A., Danesino, C., Comis, M., et al. (2005). A proportion of patients with lymphoma may harbor mutations of the perforin gene. *Blood* *105*, 4424–4428.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature* *464*, 704–712.

Consortium, T.E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* *489*, 57–74.

Consortium, T.I.H.3. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* *467*, 52–58.

Cooper, G.S., Miller, F.W., and Pandey, J.P. (1999). The role of genetic factors in autoimmune disease: Implications for environmental research. *Environ Health Perspect* *107*, 693–700.

Côte, M., Ménager, M.M., Burgess, A., Mahlaoui, N., Picard, C., Schaffner, C., Al-Manjomi, F., Al-Harbi, M., Alangari, A., Le Deist, F., et al. (2009). Munc18-2 deficiency causes familial hemophagocytic lymphohistiocytosis type 5 and impairs cytotoxic granule exocytosis in patient NK cells. *Journal of Clinical Investigation* *119*, 3765–3773.

Crow, Y.J. (2011). Type I interferonopathies: A novel set of inborn errors of immunity. *Annals of the New York Academy of Sciences* *1238*, 91–98.

Crow, Y.J., and Manel, N. (2015). Aicardi-Goutières syndrome and the type I interferonopathies. *Nat Rev Immunol* *15*, 429–440.

Das, R., Guan, P., Sprague, L., Verbist, K., Tedrick, P., An, Q.A., Cheng, C., Kurachi, M., Levine, R., Wherry, E.J., et al. (2016). Janus kinase inhibition lessens inflammation and ameliorates disease in murine models of hemophagocytic lymphohistiocytosis. *Blood* *127*, 1666–1675.

Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature* *334*, 395–402.

de Saint Basile, G., Ménasché, G., and Fischer, A. (2010). Molecular mechanisms of biogenesis and exocytosis of cytotoxic granules. *Nat Rev Immunol* *10*, 568–579.

Deiva, K., Mahlaoui, N., Beaudonnet, F., Basile, G. de S., Caridade, G., Moshous, D., Mikaeloff, Y., Blanche, S., Fischer, A., and Tardieu, M. (2012). CNS involvement at the onset of primary hemophagocytic lymphohistiocytosis. *Neurology* *78*, 1150–1156.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* *43*, 491–498.

Dias, C., McDonald, A., Sincan, M., Rupps, R., Markello, T., Salvarinova, R., Santos, R.F., Menghrajani, K., Ahaghotu, C., Sutherland, D.P., et al. (2013). Recurrent subacute post-viral onset of ataxia associated with a PRF1 mutation. *Eur J Hum Genet* *21*, 1232–1239.

Dickinson, R.E., Griffin, H., Bigley, V., Reynard, L.N., Hussain, R., Haniffa, M., Lakey, J.H., Rahman, T., Wang, X.-N., McGovern, N., et al. (2011). Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood* *118*, 2656–2658.

DiNardo, C.D., Bannon, S.A., Routbort, M., Franklin, A., Mork, M., Armanios, M., Mace, E.M., Orange, J.S., Jeff-Eke, M., Churpek, J.E., et al. (2016). Evaluation of Patients and Families With Concern for Predispositions to Hematologic Malignancies Within the Hereditary Hematologic Malignancy Clinic (HHMC). *Clinical Lymphoma Myeloma and Leukemia* *16*, 417–428.e2.

Doolittle, W.F. (2013). Is junk DNA bunk? A critique of ENCODE. *PNAS* *110*, 5294–5300.

Eddy, S.R. (2012). The C-value paradox, junk DNA and ENCODE. *Current Biology* *22*, R898–R899.

Egeler, R.M., Shapiro, R., Loehelt, B., and Filipovich, A. (1996). Characteristic immune abnormalities in hemophagocytic lymphohistiocytosis. *J. Pediatr. Hematol. Oncol.* *18*, 340–345.

Entesarian, M., Chiang, S.C.C., Schlums, H., Meeths, M., Chan, M.-Y., Mya, S.-N., Soh, S.-Y., Nordenskjöld, M., Henter, J.-I., and Bryceson, Y.T. (2013). Novel deep intronic and missense UNC13D mutations in familial haemophagocytic lymphohistiocytosis type 3. *Br J Haematol* *162*, 415–418.

Evans, D.M., Frazer, I.H., and Martin, N.G. (1999). Genetic and environmental causes of variation in basal levels of blood cells. *Twin Research* (2000) *2*, 250–257.

Fadeel, B., Orrenius, S., and Henter, J.I. (1999). Induction of apoptosis and caspase activation in cells obtained from familial haemophagocytic lymphohistiocytosis patients. *Br. J. Haematol.* *106*, 406–415.

Fakhro, K.A., Staudt, M.R., Ramstetter, M.D., Robay, A., Malek, J.A., Badii, R., Al-Marri, A.A.-N., Khalil, C.A., Al-Shakaki, A., Chidiac, O., et al. (2016). The Qatar genome: A population-specific tool for precision medicine in the Middle East. *Human Genome Variation* *3*, hgv201616.

Farquhar, J.W., and Claireaux, A.E. (1952). Familial haemophagocytic reticulosis. *Archives of Disease in Childhood* *27*, 519–525.

Feldmann, J., Callebaut, I., Raposo, G., Certain, S., Bacq, D., Dumont, C., Lambert, N., Ouachée-Chardin, M., Chedeville, G., and Tamary, H. (2003). Munc13-4 is essential for cytolytic granules fusion and is mutated in a form of familial hemophagocytic lymphohistiocytosis (FHL3). *Cell* *115*, 461–473.

Fischer, A., and Rausell, A. (2016). Primary immunodeficiencies suggest redundancy within the human immune system. *Science Immunology* *1*, eaah5861.

Fischer, A., Provot, J., Jais, J.-P., Alcais, A., and Mahlaoui, N. (2017). Autoimmune and inflammatory manifestations occur frequently in patients with primary immunodeficiencies. *Journal of Allergy and Clinical Immunology*.

Forsberg, L.A., Gisselsson, D., and Dumanski, J.P. (2017). Mosaicism in health and disease — clones picking up speed. *Nat Rev Genet* *18*, 128–142.

Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., et al. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* *449*, 851–861.

Gallucci, S., Lolkema, M., and Matzinger, P. (1999). Natural adjuvants: Endogenous activators of dendritic cells. *Nat Med* *5*, 1249–1255.

Gasteiger, G., and Rudensky, A.Y. (2014). Interactions between innate and adaptive lymphocytes. *Nat Rev Immunol* *14*, 631–639.

Gámez-Díaz, L., August, D., Stepensky, P., Revel-Vilk, S., Seidel, M.G., Noriko, M., Morio, T., Worth, A.J.J., Blessing, J., Van de Veerdonk, F., et al. (2016). The extended phenotype of LPS-responsive beige-like anchor protein (LRBA) deficiency. *Journal of Allergy and Clinical Immunology* *137*, 223–230.

Gennery, A.R., Slatter, M.A., Grandin, L., Taupin, P., Cant, A.J., Veys, P., Amrolia, P.J., Gaspar, H.B., Davies, E.G., Friedrich, W., et al. (2010). Transplantation of hematopoietic stem cells and long-term survival for primary immunodeficiencies in Europe: Entering a new century, do we do better? *Journal of Allergy and Clinical Immunology* *126*, 602–610.e11.

Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *New England Journal of Medicine* *371*, 2477–2487.

Godley, L.A., and Shimamura, A. (2017). Genetic predisposition to hematologic malignancies: Management and surveillance. *Blood* *130*, 424–432.

Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics* *17*, 333–351.

Graur, D., Zheng, Y., Price, N., Azevedo, R.B.R., Zufall, R.A., and Elhaik, E. (2013). On the Immortality of Television Sets: “Function” in the Human Genome According to the Evolution-Free Gospel of ENCODE. *Genome Biol Evol* 5, 578–590.

Gudbjartsson, D.F., Helgason, H., Gudjonsson, S.A., Zink, F., Oddson, A., Gylfason, A., Besenbacher, S., Magnusson, G., Halldorsson, B.V., Hjartarson, E., et al. (2015). Large-scale whole-genome sequencing of the Icelandic population. *Nat Genet* 47, 435–444.

Hahn, C.N., Chong, C.-E., Carmichael, C.L., Wilkins, E.J., Brautigan, P.J., Babic, M., Lin, M., Carmagnac, A., Lee, Y.K., Kok, C.H., et al. (2011). Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet* 43, 1012–1017.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674.

Hartz, B., Marsh, R., Rao, K., Henter, J.-I., Jordan, M., Filipovich, L., Bader, P., Beier, R., Burkhardt, B., Meisel, R., et al. (2016). The minimum required level of donor chimerism in hereditary hemophagocytic lymphohistiocytosis. *Blood* 127, 3281–3290.

Hasle, H., Kerndrup, G., and Jacobsen, B.B. (1995). Childhood myelodysplastic syndrome in Denmark: Incidence and predisposing conditions. *Leukemia* 9, 1569–1572.

Henter, J.-I., Elinder, G., Söder, O., and Öst, Å. (1991a). Incidence in Sweden and Clinical Features of Familial Hemophagocytic Lymphohistiocytosis. *Acta Paediatrica* 80, 428–435.

Henter, J.-I., Horne, A., Aricò, M., Egeler, R.M., Filipovich, A.H., Imashuku, S., Ladisch, S., McClain, K., Webb, D., Winiarski, J., et al. (2007). HLH-2004: Diagnostic and therapeutic guidelines for hemophagocytic lymphohistiocytosis. *Pediatric Blood & Cancer* 48, 124–131.

Henter, J.-I., Elinder, G., Söder, O., Hansson, M., Andersson, B., and Andersson, U. (1991b). Hypercytokinemia in familial hemophagocytic lymphohistiocytosis. *Blood* 78, 2918–2922.

Henter, J.-I., Ehrnst, A., Andersson, J., and Elinder, G. (1993). Familial hemophagocytic lymphohistiocytosis and viral infections. *Acta Paediatr.* 82, 369–372.

Henter, J.-I., Aricò, M., Egeler, R.M., Elinder, G., Favara, B.E., Filipovich, A.H., Gadner, H., Imashuku, S., Janka-Schaub, G., Komp, D., et al. (1997). HLH-94: A treatment protocol for hemophagocytic lymphohistiocytosis. HLH study Group of the Histiocyte Society. *Med. Pediatr. Oncol.* 28, 342–347.

Henter, J.-I., Carlson, L.A., Söder, O., Nilsson-Ehle, P., and Elinder, G. (1991c). Lipoprotein alterations and plasma lipoprotein lipase reduction in familial hemophagocytic lymphohistiocytosis. *Acta Paediatr Scand* 80, 675–681.

Hershkovitz, D., Gross, Y., Nahum, S., Yehezkel, S., Sarig, O., Uitto, J., and Sprecher, E. (2011). Functional Characterization of SAMD9, a Protein Deficient in Normophosphatemic Familial Tumoral Calcinosis. *Journal of Investigative Dermatology* 131, 662–669.

Hirschhorn, R. (2003). In vivo reversion to normal of inherited mutations in humans. *J Med Genet* *40*, 721–728.

Horne, A., Janka, G., Maarten Egeler, R., Gadner, H., Imashuku, S., Ladisch, S., Locatelli, F., Montgomery, S.M., Webb, D., Winiarski, J., et al. (2005). Haematopoietic stem cell transplantation in haemophagocytic lymphohistiocytosis. *British Journal of Haematology* *129*, 622–630.

Horne, A., Trottestam, H., Aricò, M., Egeler, R.M., Filipovich, A.H., Gadner, H., Imashuku, S., Ladisch, S., Webb, D., Janka, G., et al. (2008). Frequency and spectrum of central nervous system involvement in 193 children with haemophagocytic lymphohistiocytosis. *British Journal of Haematology* *140*, 327–335.

Horne, A., Wickström, R., Jordan, M.B., Yeh, E.A., Naqvi, A., Henter, J.-I., and Janka, G. (2017). How to Treat Involvement of the Central Nervous System in Hemophagocytic Lymphohistiocytosis? *Curr Treat Options Neurol* *19*, 3.

Huck, K., Feyen, O., Niehues, T., Ruschendorf, F., Hubner, N., Laws, H.-J., Telieps, T., Knapp, S., Wacker, H.-H., Meindl, A., et al. (2009). Girls homozygous for an IL-2-inducible T cell kinase mutation that leads to protein deficiency develop fatal EBV-associated lymphoproliferation. *J Clin Invest* *119*, 1350–1358.

Hussein, A.A., Hamadah, T., Qandeel, M., Sughayer, M., Amarin, R., Mansour, A., Chiang, S.C., Al-Zaben, A., Meeths, M., and Bryceson, Y.T. (2014). Hematopoietic stem cell transplantation of an adolescent with neurological manifestations of homozygous missense PRF1 mutation. *Pediatr Blood Cancer* *61*, 2313–2315.

Imai, K., Matsuyama, S., Miyake, S., Suga, K., and Nakachi, K. (2000). Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: An 11-year follow-up study of a general population. *The Lancet* *356*, 1795–1799.

Ivashkiv, L.B., and Donlin, L.T. (2014). Regulation of type I interferon responses. *Nat Rev Immunol* *14*, 36–49.

Jacobs, K.B., Yeager, M., Zhou, W., Wacholder, S., Wang, Z., Rodriguez-Santiago, B., Hutchinson, A., Deng, X., Liu, C., Horner, M.-J., et al. (2012). Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet* *44*, 651–658.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine* *371*, 2488–2498.

Jaiswal, S., Natarajan, P., Silver, A.J., Gibson, C.J., Bick, A.G., Shvartz, E., McConkey, M., Gupta, N., Gabriel, S., Ardissino, D., et al. (2017). Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *New England Journal of Medicine* *377*, 111–121.

Janka, G. (2012). Familial and Acquired Hemophagocytic Lymphohistiocytosis. *Annual Review of Medicine* *63*, 233–246.

- Janka, G.E. (1983). Familial hemophagocytic lymphohistiocytosis. *Eur J Pediatr* *140*, 221–230.
- Janka, G.E., and Lehmberg, K. (2014). Hemophagocytic syndromes — An update. *Blood Reviews* *28*, 135–142.
- Johnson, M.B., De Franco, E., Lango Allen, H., Al Senani, A., Elbarbary, N., Siklar, Z., Berberoglu, M., Imane, Z., Haghghi, A., Razavi, Z., et al. (2017). Recessively Inherited LRBA Mutations Cause Autoimmunity Presenting as Neonatal Diabetes. *Diabetes* *66*, 2316–2322.
- Johnson, T.S., Terrell, C.E., Millen, S.H., Katz, J.D., Hildeman, D.A., and Jordan, M.B. (2014). Etoposide selectively ablates activated T cells to control the immunoregulatory disorder hemophagocytic lymphohistiocytosis. *J. Immunol.* *192*, 84–91.
- Jordan, M.B. (2004). An animal model of hemophagocytic lymphohistiocytosis (HLH): CD8⁺ T cells and interferon gamma are essential for the disorder. *Blood* *104*, 735–743.
- Jouanguy, E., Dupuis, S., Pallier, A., Doffinger, R., Fondaneche, M.-C., Fieschi, C., Lamhamdi-Cherradi, S., Altare, F., Emile, J.-F., Lutz, P., et al. (2000). In a novel form of IFN-γ receptor 1 deficiency, cell surface receptors fail to bind IFN-γ. *J Clin Invest* *105*, 1429–1436.
- Jung, D., Giallourakis, C., Mostoslavsky, R., and Alt, F.W. (2006). Mechanism and Control of V(d)j Recombination at the Immunoglobulin Heavy Chain Locus. *Annual Review of Immunology* *24*, 541–570.
- Kang, E., and Gennery, A. (2014). Hematopoietic Stem Cell Transplantation for Primary Immunodeficiencies. *Hematology/Oncology Clinics of North America* *28*, 1157–1170.
- Katsanis, N. (2016). The continuum of causality in human genetic disorders. *Genome Biology* *17*, 233.
- Kaufman, K.M., Linghu, B., Szustakowski, J.D., Husami, A., Yang, F., Zhang, K., Filipovich, A.H., Fall, N., Harley, J.B., Nirmala, N.R., et al. (2014). Whole-exome sequencing reveals overlap between macrophage activation syndrome in systemic juvenile idiopathic arthritis and familial hemophagocytic lymphohistiocytosis. *Arthritis & Rheumatology (Hoboken, N.J.)* *66*, 3486–3495.
- Kawai, T., and Akira, S. (2006). TLR signaling. *Cell Death Differ* *13*, 816–825.
- Keel, S.B., Scott, A., Sanchez-Bonilla, M., Ho, P.A., Gulsuner, S., Pritchard, C.C., Abkowitz, J.L., King, M.-C., Walsh, T., and Shimamura, A. (2016). Genetic features of myelodysplastic syndrome and aplastic anemia in pediatric and young adult patients. *Haematologica* *101*, 1343–1350.

Kiessling, R., Klein, E., and Wigzell, H. (1975). „Natural” killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype. *Eur. J. Immunol.* *5*, 112–117.

Kircher, M., Witten, D.M., Jain, P., O’Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* *46*, 310–315.

Knudson, A.G. (1971). Mutation and Cancer: Statistical Study of Retinoblastoma. *PNAS* *68*, 820–823.

Kono, H., and Rock, K.L. (2008). How dying cells alert the immune system to danger. *Nat Rev Immunol* *8*, 279–289.

Krummel, M.F., and Allison, J.P. (1996). CTLA-4 engagement inhibits IL-2 accumulation and cell cycle progression upon activation of resting T cells. *J. Exp. Med.* *183*, 2533–2540.

Kuehn, H.S., Ouyang, W., Lo, B., Deenick, E.K., Niemela, J.E., Avery, D.T., Schickel, J.-N., Tran, D.Q., Stoddard, J., Zhang, Y., et al. (2014). Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. *Science* *345*, 1623–1627.

Kwan, A., Abraham, R.S., Currier, R., Brower, A., Andruszewski, K., Abbott, J.K., Baker, M., Ballow, M., Bartoshesky, L.E., Bonilla, F.A., et al. (2014). Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. *JAMA* *312*, 729–738.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* *409*, 860–921.

Laurie, C.C., Laurie, C.A., Rice, K., Doheny, K.F., Zelnick, L.R., McHugh, C.P., Ling, H., Hetrick, K.N., Pugh, E.W., Amos, C., et al. (2012). Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet* *44*, 642–650.

Law, R.H.P., Lukyanova, N., Voskoboinik, I., Caradoc-Davies, T.T., Baran, K., Dunstone, M.A., D’Angelo, M.E., Orlova, E.V., Coulibaly, F., Verschoor, S., et al. (2010). The structural basis for membrane binding and pore formation by lymphocyte perforin. *Nature* *468*, 447–451.

Leechawengwongs, E., and Shearer, W.T. (2012). Lymphoma complicating primary immunodeficiency syndromes. *Current Opinion in Hematology* *19*, 305–312.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O’Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* *536*, 285–291.

- Lelieveld, S.H., Spielmann, M., Mundlos, S., Veltman, J.A., and Gilissen, C. (2015). Comparison of Exome and Genome Sequencing Technologies for the Complete Capture of Protein-Coding Regions. *Hum. Mutat.* *36*, 815–822.
- Levy, S., Sutton, G., Ng, P.C., Feuk, L., Halpern, A.L., Walenz, B.P., Axelrod, N., Huang, J., Kirkness, E.F., Denisov, G., et al. (2007). The Diploid Genome Sequence of an Individual Human. *PLOS Biology* *5*, e254.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* *25*, 1754–1760.
- Li, F.P., Potter, N.U., Buchanan, G.R., Vawter, G., Whang-Peng, J., and Rosen, R.B. (1978). A family with acute leukemia, hypoplastic anemia and cerebellar ataxia. *The American Journal of Medicine* *65*, 933–940.
- Li, F.P., Hecht, F., Kaiser-McCaw, B., Baranko, P.V., and Upp Potter, N. (1981). Ataxia-pancytopenia: Syndrome of cerebellar ataxia, hypoplastic anemia, monosomy 7, and acute myelogenous leukemia. *Cancer Genetics and Cytogenetics* *4*, 189–196.
- Li Y., Oosting, M., Smeekens, S.P., Jaeger, M., Aguirre-Gamboa, R., Le, K.T.T., Deelen, P., Ricaño-Ponce, I., Schoffelen, T., Jansen, A.F.M., et al. (2016). A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell* *167*, 1099–1110.e14.
- Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., et al. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* *478*, 476–482.
- Liu, J., and McFadden, G. (2015). SAMD9 is an innate antiviral host factor with stress response properties that can be antagonized by poxviruses. *J. Virol.* *89*, 1925–1931.
- Ljunggren, H.G., and Kärre, K. (1990). In search of the 'missing self': MHC molecules and NK cell recognition. *Immunol. Today* *11*, 237–244.
- Lo, B., Zhang, K., Lu, W., Zheng, L., Zhang, Q., Kanelloupolou, C., Zhang, Y., Liu, Z., Fritz, J.M., Marsh, R., et al. (2015). Patients with LRBA deficiency show CTLA4 loss and immune dysregulation responsive to abatacept therapy. *Science* *349*, 436–440.
- Longerich, S., Basu, U., Alt, F., and Storb, U. (2006). AID in somatic hypermutation and class switch recombination. *Current Opinion in Immunology* *18*, 164–174.
- Lopez-Herrera, G., Tampella, G., Pan-Hammarström, Q., Herholz, P., Trujillo-Vargas, C.M., Phadwal, K., Simon, A.K., Moutschen, M., Etzioni, A., Mory, A., et al. (2012). deleterious Mutations in LRBA Are Associated with a Syndrome of Immune Deficiency and Autoimmunity. *The American Journal of Human Genetics* *90*, 986–1001.

Löfstedt, A., Chiang, S.C.C., Onelöv, E., Bryceson, Y.T., Meeths, M., and Henter, J.-I. (2015). Cancer risk in relatives of patients with a primary disorder of lymphocyte cytotoxicity: A retrospective cohort study. *The Lancet Haematology* 2, e536–e542.

Ma, Q., Yu, T., Ren, Y.-Y., Gong, T., and Zhong, D.-S. (2014). Overexpression of SAMD9 suppresses tumorigenesis and progression during non small cell lung cancer. *Biochemical and Biophysical Research Communications* 454, 157–161.

MacArthur, D. (2016). Superheroes of disease resistance. *Nat Biotech* 34, 512–513.

MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature* 508, 469–476.

Machaczka, M., Vaktnäs, J., Klimkowska, M., Hägglund, H. (2011). Malignancy-associated hemophagocytic lymphohistiocytosis in adults: A retrospective population-based analysis from a single center. *Leuk. Lymphoma* 52, 613–619.

Machaczka, M., Klimkowska, M., Chiang, S.C.C., Meeths, M., Muller, M.-L., Gustafsson, B., Henter, J.-I., and Bryceson, Y.T. (2012). Development of classical Hodgkin's lymphoma in an adult with biallelic STXBP2 mutations. *Haematologica* 98, 760–764.

Mahlaoui, N., Ouachee-Chardin, M., de Saint Basile, G., Neven, B., Picard, C., Blanche, S., and Fischer, A. (2007). Immunotherapy of Familial Hemophagocytic Lymphohistiocytosis With Antithymocyte Globulins: A Single-Center Retrospective Report of 38 Patients. *PEDIATRICS* 120, e622–e628.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F., Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.Z., and Tainsky, M.A. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* 250, 1233–1238.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753.

Marsh, R.A., Vaughn, G., Kim, M.-O., Li, D., Jodele, S., Joshi, S., Mehta, P.A., Davies, S.M., Jordan, M.B., Bleesing, J.J., et al. (2010). Reduced-intensity conditioning significantly improves survival of patients with hemophagocytic lymphohistiocytosis undergoing allogeneic hematopoietic cell transplantation. *Blood* 116, 5824–5831.

Marsh, R.A., Allen, C.E., McClain, K.L., Weinstein, J.L., Kanter, J., Skiles, J., Lee, N.D., Khan, S.P., Lawrence, J., Mo, J.Q., et al. (2013). Salvage therapy of refractory hemophagocytic lymphohistiocytosis with alemtuzumab. *Pediatr. Blood Cancer* 60, 101–109.

Maschalidi, S., Sepulveda, F.E., Garrigue, A., Fischer, A., and Basile, G. de S. (2016). Therapeutic effect of JAK1/2 blockade on the manifestations of hemophagocytic lymphohistiocytosis in mice. *Blood* 128, 60–71.

- McGee, R.B., and Nichols, K.E. (2016). Introduction to cancer genetic susceptibility syndromes. *Hematology* *2016*, 293–301.
- McKusick, V.A. (2007). Mendelian Inheritance in Man and Its Online Version, OMIM. *The American Journal of Human Genetics* *80*, 588–604.
- McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flliceck, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* *26*, 2069–2070.
- McVean, G.A., Altshuler (Co-Chair), D.M., Durbin (Co-Chair), R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flliceck, P., et al. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* *491*, 56–65.
- Meeths, M., and Bryceson, Y.T. (2016). HLH susceptibility: Genetic lesions add up. *Blood* *127*, 2051–2052.
- Meeths, M., Bryceson, Y.T., Rudd, E., Zheng, C., Wood, S.M., Ramme, K., Beutel, K., Hasle, H., Heilmann, C., Hultenby, K., et al. (2010a). Clinical presentation of Griscelli syndrome type 2 and spectrum of RAB27A mutations. *Pediatr. Blood Cancer* *54*, 563–572.
- Meeths, M., Entesarian, M., Al-Herz, W., Chiang, S.C.C., Wood, S.M., Al-Ateeqi, W., Almazan, F., Boelens, J.J., Hasle, H., Ifversen, M., et al. (2010b). Spectrum of clinical presentations in familial hemophagocytic lymphohistiocytosis type 5 patients with mutations in STXBP2. *Blood* *116*, 2635–2643.
- Meeths, M., Chiang, S.C.C., Wood, S.M., Entesarian, M., Schlums, H., Bang, B., Nordenskjold, E., Bjorklund, C., Jakovljevic, G., Jazbec, J., et al. (2011). Familial hemophagocytic lymphohistiocytosis type 3 (FHL3) caused by deep intronic mutation and inversion in UNC13D. *Blood* *118*, 5783–5793.
- Meeths, M., Chiang, S.C.C., Löfstedt, A., Müller, M.-L., Tesi, B., Henter, J.-I., and Bryceson, Y.T. (2014). Pathophysiology and spectrum of diseases caused by defects in lymphocyte cytotoxicity. *Experimental Cell Research* *325*, 10–17.
- Meeths, M., Horne, A., Sabel, M., Bryceson, Y.T., and Henter, J.-I. (2015). Incidence and clinical presentation of primary hemophagocytic lymphohistiocytosis in Sweden. *Pediatr. Blood Cancer* *62*, 346–352.
- Metzker, M.L. (2010). Sequencing technologies — the next generation. *Nat Rev Genet* *11*, 31–46.
- Meyts, I., Bosch, B., Bolze, A., Boisson, B., Itan, Y., Belkadi, A., Pedergnana, V., Moens, L., Picard, C., Cobat, A., et al. (2016). Exome and genome sequencing for inborn errors of immunity. *Journal of Allergy and Clinical Immunology* *138*, 957–969.

Ménasché, G., Pastural, E., Feldmann, J., Certain, S., Ersoy, F., Dupuis, S., Wulffraat, N., Bianchi, D., Fischer, A., Le Deist, F., et al. (2000). Mutations in RAB27A cause Griscelli syndrome associated with haemophagocytic syndrome. *Nat Genet* 25, 173–176.

Mombaerts, P., Iacomini, J., Johnson, R.S., Herrup, K., Tonegawa, S., and Papaioannou, V.E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* 68, 869–877.

Nagamachi, A., Matsui, H., Asou, H., Ozaki, Y., Aki, D., Kanai, A., Takubo, K., Suda, T., Nakamura, T., Wolff, L., et al. (2013). Haploinsufficiency of SAMD9L, an Endosome Fusion Facilitator, Causes Myeloid Malignancies in Mice Mimicking Human Diseases with Monosomy 7. *Cancer Cell* 24, 305–317.

Nagasaki, M., Yasuda, J., Katsuoka, F., Narai, N., Kojima, K., Kawai, Y., Yamaguchi-Kabata, Y., Yokozawa, J., Danjoh, I., Saito, S., et al. (2015). Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. *Nature Communications* 6, ncomms9018.

Nalls, M.A., Couper, D.J., Tanaka, T., Rooij, F.J.A. van, Chen, M.-H., Smith, A.V., Toniolo, D., Zakai, N.A., Yang, Q., Greinacher, A., et al. (2011). Multiple Loci Are Associated with White Blood Cell Phenotypes. *PLOS Genetics* 7, e1002113.

Narumi, S., Amano, N., Ishii, T., Katsumata, N., Muroya, K., Adachi, M., Toyoshima, K., Tanaka, Y., Fukuzawa, R., Miyako, K., et al. (2016). SAMD9 mutations cause a novel multisystem disorder, MIRAGE syndrome, and are associated with loss of chromosome 7. *Nat Genet* 48, 792–797.

Neefjes, J., Jongsma, M.L.M., Paul, P., and Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*.

Netter, P., Chan, S.K., Banerjee, P.P., Monaco-Shawver, L., Noroski, L.M., Hanson, I.C., Forbes, L.R., Mace, E.M., Chinen, J., Gaspar, H.B., et al. (2016). A novel Rab27a mutation binds melanophilin, but not Munc13-4, causing immunodeficiency without albinism. *J. Allergy Clin. Immunol.* 138, 599–601.e3.

Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucl. Acids Res.* 31, 3812–3814.

Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., et al. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461, 272–276.

Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K., Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., et al. (2010). Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 42, 30–35.

Nijman, I.J., van Montfrans, J.M., Hoogstraat, M., Boes, M.L., van de Corput, L., Renner, E.D., van Zon, P., van Lieshout, S., Elferink, M.G., van der Burg, M., et al. (2014). Targeted

next-generation sequencing: A novel diagnostic tool for primary immunodeficiencies. *Journal of Allergy and Clinical Immunology* *133*, 529–534.e1.

Notarangelo, L.D., and Casanova, J.-L. (2009). Primary immunodeficiencies: Increasing market share. *Current Opinion in Immunology* *21*, 461–465.

Notarangelo, L.D., and Fleisher, T.A. (2017). Targeted strategies directed at the molecular defect: Toward precision medicine for select primary immunodeficiency disorders. *Journal of Allergy and Clinical Immunology* *139*, 715–723.

Orgogozo, V., Peluffo, A.E., and Morizot, B. (2016). The “Mendelian Gene” and the “Molecular Gene”. *Current Topics in Developmental Biology* *119*, 1–26.

Orrù, V., Steri, M., Sole, G., Sidore, C., Virdis, F., Dei, M., Lai, S., Zoledziewska, M., Busonero, F., Mulas, A., et al. (2013). Genetic Variants Regulating Immune Cell Levels in Health and Disease. *Cell* *155*, 242–256.

Paila, U., Chapman, B.A., Kirchner, R., and Quinlan, A.R. (2013). GEMINI: Integrative Exploration of Genetic Variation and Genome Annotations. *PLoS Comput Biol* *9*, e1003153.

Pang, W.W., Schrier, S.L., and Weissman, I.L. (2017). Age-associated changes in human hematopoietic stem cells. *Semin. Hematol.* *54*, 39–42.

Pappas, D.J., Coppola, G., Gabatto, P.A., Gao, F., Geschwind, D.H., Oksenberg, J.R., and Baranzini, S.E. (2009). Longitudinal system-based analysis of transcriptional responses to type I interferons. *Physiol. Genomics* *38*, 362–371.

Parvaneh, N., Filipovich, A.H., and Borkhardt, A. (2013). Primary immunodeficiencies predisposed to Epstein-Barr virus-driven haematological diseases. *British Journal of Haematology* *162*, 573–586.

Peltonen, L., and McKusick, V.A. (2001). Dissecting Human Disease in the Postgenomic Era. *Science* *291*, 1224–1229.

Philippakis, A.A., Azzariti, D.R., Beltran, S., Brookes, A.J., Brownstein, C.A., Brudno, M., Brunner, H.G., Buske, O.J., Carey, K., Doll, C., et al. (2015). The Matchmaker Exchange: A Platform for Rare Disease Gene Discovery. *Human Mutation* *36*, 915–921.

Picard, C., Al-Herz, W., Bousfiha, A., Casanova, J.-L., Chatila, T., Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Holland, S.M., Klein, C., et al. (2015). Primary Immunodeficiency Diseases: An Update on the Classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency 2015. *J. Clin. Immunol.* *35*, 696–726.

Plagnol, V., Curtis, J., Epstein, M., Mok, K.Y., Stebbings, E., Grigoriadou, S., Wood, N.W., Hambleton, S., Burns, S.O., Thrasher, A.J., et al. (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* *28*, 2747–2754.

Quintana-Murci, L., and Clark, A.G. (2013). Population genetic tools for dissecting innate immunity in humans. *Nat Rev Immunol* *13*, 280–293.

Ramos-Casals, M., Brito-Zerón, P., López-Guillermo, A., Khamashta, M.A., and Bosch, X. (2014). Adult haemophagocytic syndrome. *The Lancet* *383*, 1503–1516.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., et al. (2006). Global variation in copy number in the human genome. *Nature* *444*, 444–454.

Reese, M.G., Eeckman, F.H., Kulp, D., and Haussler, D. (1997). Improved Splice Site Detection in Genie. *Journal of Computational Biology* *4*, 311–323.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* *17*, 405–423.

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* *518*, 317–330.

Rodriguez-Abreu, D., Bordoni, A., and Zucca, E. (2007). Epidemiology of hematological malignancies. *Ann Oncol* *18*, i3–i8.

Roederer, M., Quaye, L., Mangino, M., Beddall, M.H., Mahnke, Y., Chattopadhyay, P., Tosi, I., Napolitano, L., Terranova Barberio, M., Menni, C., et al. (2015). The Genetic Architecture of the Human Immune System: A Bioresource for Autoimmunity and Disease Pathogenesis. *Cell* *161*, 387–403.

Rohr, J., Beutel, K., Maul-Pavicic, A., Vraetz, T., Thiel, J., Warnatz, K., Bondzio, I., Gross-Wielisch, U., Schundeln, M., Schutz, B., et al. (2010). Atypical familial hemophagocytic lymphohistiocytosis due to mutations in UNC13D and STXBP2 overlaps with primary immunodeficiency diseases. *Haematologica* *95*, 2080–2087.

Romberg, N., Al Moussawi, K., Nelson-Williams, C., Stiegler, A.L., Loring, E., Choi, M., Overton, J., Meffre, E., Khokha, M.K., Huttner, A.J., et al. (2014). Mutation of NLRC4 causes a syndrome of enterocolitis and autoinflammation. *Nat Genet* *46*, 1135–1139.

Rubin, T.S., Zhang, K., Gifford, C., Lane, A., Choo, S., Bleesing, J.J., and Marsh, R.A. (2017). Perforin and CD107a testing is superior to NK cell function testing for screening patients for genetic HLH. *Blood* *129*, 2993–2999.

Sakaguchi, S., Miyara, M., Costantino, C.M., and Hafler, D.A. (2010). FOXP3⁺ regulatory T cells in the human immune system. *Nature Reviews Immunology* *10*, 490–500.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, C.A., Hutchison, C.A., Slocombe, P.M., and Smith, M. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* *265*, 687–695.

Schirmer, M., Smeekens, S.P., Vlamakis, H., Jaeger, M., Oosting, M., Franzosa, E.A., ter Horst, R., Jansen, T., Jacobs, L., Bonder, M.J., et al. (2016). Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* *167*, 1125–1136.e8.

Schlums, H., Cichocki, F., Tesi, B., Theorell, J., Beziat, V., Holmes, T.D., Han, H., Chiang, S.C.C., Foley, B., Mattsson, K., et al. (2015). Cytomegalovirus Infection Drives Adaptive Epigenetic Diversification of NK Cells with Altered Signaling and Effector Function. *Immunity* *42*, 443–456.

Schmid, J.P., Ho, C.-H., Chrétien, F., Lefebvre, J.M., Pivert, G., Kosco-Vilbois, M., Ferlin, W., Geissmann, F., Fischer, A., and de Saint Basile, G. (2009). Neutralization of IFN γ defeats haemophagocytosis in LCMV-infected perforin- and Rab27a-deficient mice. *EMBO Molecular Medicine* *1*, 112–124.

Schrader, K.A., Cheng, D.T., Joseph, V., Prasad, M., Walsh, M., Zehir, A., Ni, A., Thomas, T., Benayed, R., Ashraf, A., et al. (2016). Germline Variants in Targeted Tumor Sequencing Using Matched Normal DNA. *JAMA Oncol* *2*, 104–111.

Schubert, D., Bode, C., Kenefech, R., Hou, T.Z., Wing, J.B., Kennedy, A., Bulashevska, A., Petersen, B.-S., Schäffer, A.A., Grüning, B.A., et al. (2014). Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. *Nat Med* *20*, 1410–1416.

Schulert, G.S., Zhang, M., Fall, N., Husami, A., Kissell, D., Hanosh, A., Zhang, K., Davis, K., Jentzen, J.M., Napolitano, L., et al. (2016). Whole-Exome Sequencing Reveals Mutations in Genes Linked to Hemophagocytic Lymphohistiocytosis and Macrophage Activation Syndrome in Fatal Cases of H1N1 Influenza. *J. Infect. Dis.* *213*, 1180–1188.

Seidel, M.G., Hirschnugl, T., Gamez-Diaz, L., Schwinger, W., Serwas, N., Deutschmann, A., Gorkiewicz, G., Zenz, W., Windpassinger, C., Grimbacher, B., et al. (2015). Long-term remission after allogeneic hematopoietic stem cell transplantation in LPS-responsive beige-like anchor (LRBA) deficiency. *Journal of Allergy and Clinical Immunology* *135*, 1384–1390.e8.

Seidel, M.G., Böhm, K., Dogu, F., Worth, A., Thrasher, A., Florkin, B., İkinciogullari, A., Peters, A., Bakhtiar, S., Meeths, M., et al. (2017). Treatment of severe forms of LPS-responsive beige-like anchor protein deficiency with allogeneic hematopoietic stem cell transplantation. *J. Allergy Clin. Immunol.*

Seo, J.Y., Song, J.-S., Lee, K.-O., Won, H.-H., Kim, J.-W., Kim, S.-H., Lee, S.-H., Yoo, K.-H., Sung, K.-W., Koo, H.H., et al. (2013). Founder effects in two predominant intronic mutations of UNC13D, c.118-308C>T and c.754-1G>C underlie the unusual predominance of type 3 familial hemophagocytic lymphohistiocytosis (FHL3) in Korea. *Ann. Hematol.* *92*, 357–364.

Sepulveda, F.E., Maschalidi, S., Vosshenrich, C.A.J., Garrigue, A., Kurowska, M., Ménasche, G., Fischer, A., Santo, J.P.D., and Basile, G. de S. (2015). A novel immunoregulatory role for NK-cell cytotoxicity in protection from HLH-like immunopathology in mice. *Blood* *125*, 1427–1434.

Sepulveda, F.E., Garrigue, A., Maschalidi, S., Garfa-Traore, M., Ménasché, G., Fischer, A., and Basile, G. de S. (2016). Polygenic mutations in the cytotoxicity pathway increase susceptibility to develop HLH immunopathology in mice. *Blood* *127*, 2113–2121.

Sieni, E., Cetica, V., Hackmann, Y., Coniglio, M.L., Da Ros, M., Ciambotti, B., Pende, D., Griffiths, G., and Aricò, M. (2014). Familial hemophagocytic lymphohistiocytosis: When rare diseases shed light on immune system functioning. *Front. Immunol.* *5*, 167.

Smyth, M.J., Thia, K.Y.T., Street, S.E.A., MacGregor, D., Godfrey, D.I., and Trapani, J.A. (2000). Perforin-Mediated Cytotoxicity Is Critical for Surveillance of Spontaneous Lymphoma. *J Exp Med* *192*, 755–760.

Speckmann, C., Lehmberg, K., Albert, M., Damgaard, R., Fritsch, M., Gyrd-Hansen, M., Rensing-Ehl, A., Vraetz, T., Grimbacher, B., Salzer, U., et al. (2013). X-linked inhibitor of apoptosis (XIAP) deficiency: The spectrum of presenting manifestations beyond hemophagocytic lymphohistiocytosis. *Clinical Immunology* *149*, 133–141.

Spessott, W.A., Sanmillan, M.L., McCormick, M.E., Patel, N., Villanueva, J., Zhang, K., Nichols, K.E., and Giraudo, C.G. (2015). Hemophagocytic lymphohistiocytosis caused by dominant-negative mutations in STXBP2 that inhibit SNARE-mediated membrane fusion. *Blood* *125*, 1566–1577.

Spinner, M.A., Sanchez, L.A., Hsu, A.P., Shaw, P.A., Zerbe, C.S., Calvo, K.R., Arthur, D.C., Gu, W., Gould, C.M., Brewer, C.C., et al. (2014). GATA2 deficiency: A protean disorder of hematopoiesis, lymphatics, and immunity. *Blood* *123*, 809–821.

Staines-Boone, A.T., Deswarte, C., Venegas Montoya, E., Sánchez-Sánchez, L.M., Campos, G., Alberto, J., Muñiz-Ronquillo, T., Bustamante, J., Espinosa-Rosales, F.J., Reyes, L., et al. (2017). Multifocal Recurrent Osteomyelitis and Hemophagocytic Lymphohistiocytosis in a Boy with Partial Dominant IFN- γ R1 Deficiency: Case Report and Review of the Literature. *Front. Pediatr.* *5*.

Starr, T.K., Jameson, S.C., and Hogquist, K.A. (2003). Positive and Negative Selection of T Cells. *Annual Review of Immunology* *21*, 139–176.

Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P., and Ebert, B.L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* *126*, 9–16.

Stenson, P.D., Mort, M., Ball, E.V., Evans, K., Hayden, M., Heywood, S., Hussain, M., Phillips, A.D., and Cooper, D.N. (2017). The Human Gene Mutation Database: Towards a

comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet* 136, 665–677.

Stepp, S.E., Dufourcq-Lagelouse, R., Deist, F.L., Bhawan, S., Certain, S., Mathew, P.A., Henter, J.-I., Bennett, M., Fischer, A., Basile, G. de S., et al. (1999). Perforin Gene Defects in Familial Hemophagocytic Lymphohistiocytosis. *Science* 286, 1957–1959.

Stray-Pedersen, A., Sorte, H.S., Samarakoon, P., Gambin, T., Chinn, I.K., Coban Akdemir, Z.H., Erichsen, H.C., Forbes, L.R., Gu, S., Yuan, B., et al. (2017). Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. *Journal of Allergy and Clinical Immunology* 139, 232–245.

Stunnenberg, H.G., International Human Epigenome Consortium, and Hirst, M. (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. *Cell* 167, 1897.

Sudmant, P.H., Mallick, S., Nelson, B.J., Hormozdiari, F., Krumm, N., Huddleston, J., Coe, B.P., Baker, C., Nordenfelt, S., Bamshad, M., et al. (2015). Global diversity, population stratification, and selection of human copy-number variation. *Science* 349, aab3761.

Taurisano, R., Maiorana, A., Benedetti, F.D., Dionisi-Vici, C., Boldrini, R., and Deodato, F. (2014). Wolman disease associated with hemophagocytic lymphohistiocytosis: Attempts for an explanation. *Eur J Pediatr* 1–4.

Taylor, J.C., Martin, H.C., Lise, S., Broxholme, J., Cazier, J.-B., Rimmer, A., Kanapin, A., Lunter, G., Fiddy, S., Allan, C., et al. (2015). Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat Genet* 47, 717–726.

Terrell, C.E., and Jordan, M.B. (2013). Mixed hematopoietic or T-cell chimerism above a minimal threshold restores perforin-dependent immune regulation in perforin-deficient mice. *Blood* 122, 2618–2621.

Tesi, B., Schlums, H., Cichocki, F., and Bryceson, Y.T. (2016). Epigenetic Regulation of Adaptive NK Cell Diversification. *Trends in Immunology* 37, 451–461.

The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature* 526, 68–74.

The FANTOM Consortium and the RIKEN PMI and Clst (dgt) (2014). A promoter-level mammalian expression atlas. *Nature* 507, 462–470.

The Genome of the Netherlands Consortium (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet* 46, 818–825.

The UK10K Consortium (2015). The UK10K project identifies rare variants in health and disease. *Nature* 526, 82–90.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief Bioinform* *14*, 178–192.

Tivol, E.A., Borriello, F., Schweitzer, A.N., Lynch, W.P., Bluestone, J.A., and Sharpe, A.H. (1995). Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* *3*, 541–547.

Tjio, J.H., and Levan, A. (1956). The Chromosome Number of Man. *Hereditas* *42*, 1–6.

Tomasetti C, Vogelstein B. 2015). Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* *347*, 78–81.

Tomasetti, C., Li, L., and Vogelstein, B. (2017). Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* *355*, 1330–1334.

Topaz, O., Indelman, M., Chefetz, I., Geiger, D., Metzker, A., Altschuler, Y., Choder, M., Bercovich, D., Uitto, J., Bergman, R., et al. (2006). A Deleterious Mutation in SAMD9 Causes Normophosphatemic Familial Tumoral Calcinosis. *The American Journal of Human Genetics* *79*, 759–764.

Trizzino, A., Stadt, U. zur, Ueda, I., Risma, K., Janka, G., Ishii, E., Beutel, K., Sumegi, J., Cannella, S., Pende, D., et al. (2008). Genotype–phenotype study of familial haemophagocytic lymphohistiocytosis due to perforin mutations. *J Med Genet* *45*, 15–21.

Trottestam, H., Horne, A., Aricò, M., Egeler, R.M., Filipovich, A.H., Gadner, H., Imashuku, S., Ladisch, S., Webb, D., Janka, G., et al. (2011). Chemoimmunotherapy for hemophagocytic lymphohistiocytosis: Long-term results of the HLH-94 treatment protocol. *Blood* *118*, 4577–4584.

Tsang, J.S., Schwartzberg, P.L., Kotliarov, Y., Biancotto, A., Xie, Z., Germain, R.N., Wang, E., Olnes, M.J., Narayanan, M., Golding, H., et al. (2014). Global Analyses of Human Immune Variation Reveal Baseline Predictors of Postvaccination Responses. *Cell* *157*, 499–513.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* *43*, 11.10.1–33.

van Montfrans, J.M., Hoepelman, A.I.M., Otto, S., van Gijn, M., van de Corput, L., de Weger, R.A., Monaco-Shawver, L., Banerjee, P.P., Sanders, E.A.M., Jol-van der Zijde, C.M., et al. (2012). CD27 deficiency is associated with combined immunodeficiency and persistent symptomatic EBV viremia. *Journal of Allergy and Clinical Immunology* *129*, 787–793.e6.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. *Science* *291*, 1304–1351.

Wang, Q., Zhai, Y.-Y., Dai, J.-H., Li, K.-Y., Deng, Q., and Han, Z.-G. (2014a). SAMD9L inactivation promotes cell proliferation via facilitating G1-S transition in hepatitis B virus-associated hepatocellular carcinoma. *Int. J. Biol. Sci.* *10*, 807–816.

Wang, Y., Wang, Z., Zhang, J., Wei, Q., Tang, R., Qi, J., Li, L., Ye, L., Wang, J., and Ye, L. (2014b). Genetic Features of Late Onset Primary Hemophagocytic Lymphohistiocytosis in Adolescence or Adulthood. *PLoS ONE* *9*, e107386.

Wheeler, D.A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y.-J., Makhijani, V., Roth, G.T., et al. (2008). The complete genome of an individual by massively parallel DNA sequencing. *Nature* *452*, 872–876.

Włodarski, M.W., Hirabayashi, S., Pastor, V., Starý, J., Hasle, H., Masetti, R., Dworzak, M., Schmugge, M., Heuvel-Eibrink, M. van den, Ussowicz, M., et al. (2016). Prevalence, clinical characteristics, and prognosis of GATA2-related myelodysplastic syndromes in children and adolescents. *Blood* *127*, 1387–1397.

Worthey, E.A., Mayer, A.N., Syverson, G.D., Helbling, D., Bonacci, B.B., Decker, B., Serpe, J.M., Dasu, T., Tschannen, M.R., Veith, R.L., et al. (2011). Making a definitive diagnosis: Successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med* *13*, 255–262.

Wu, S., Powers, S., Zhu, W., and Hannun, Y.A. (2016). Substantial contribution of extrinsic risk factors to cancer development. *Nature* *529*, 43–47.

Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* *369*, 1502–1511.

Yasumi, T., Hori, M., Hiejima, E., Shibata, H., Izawa, K., Oda, H., Yoshioka, K., Nakagawa, K., Kawai, T., Nishikomori, R., et al. (2015). Laboratory parameters identify familial haemophagocytic lymphohistiocytosis from other forms of paediatric haemophagocytosis. *Br. J. Haematol.* *170*, 532–538.

Yu, H., Zhang, V.W., Stray-Pedersen, A., Hanson, I.C., Forbes, L.R., de la Morena, M.T., Chinn, I.K., Gorman, E., Mendelsohn, N.J., Pozos, T., et al. (2016). Rapid molecular diagnostics of severe primary immunodeficiency determined by using targeted next-generation sequencing. *Journal of Allergy and Clinical Immunology* *138*, 1142–1151.e2.

Zhang, J., Walsh, M.F., Wu, G., Edmonson, M.N., Gruber, T.A., Easton, J., Hedges, D., Ma, X., Zhou, X., Yergeau, D.A., et al. (2015). Germline Mutations in Predisposition Genes in Pediatric Cancer. *New England Journal of Medicine* *373*, 2336–2346.

Zhang, K., Jordan, M.B., Marsh, R.A., Johnson, J.A., Kissell, D., Meller, J., Villanueva, J., Risma, K.A., Wei, Q., Klein, P.S., et al. (2011). Hypomorphic mutations in PRF1, MUNC13-4, and STXBP2 are associated with adult-onset familial HLH. *Blood* 118, 5794–5798.

Zhang, M., Bracaglia, C., Prencipe, G., Bemrich-Stolz, C.J., Beukelman, T., Dimmitt, R.A., Chatham, W.W., Zhang, K., Li, H., Walter, M.R., et al. (2016). A Heterozygous RAB27A Mutation Associated with Delayed Cytolytic Granule Polarization and Hemophagocytic Lymphohistiocytosis. *J Immunol* 1501284.

zur Stadt, U., Schmidt, S., Kasper, B., Beutel, K., Diler, A.S., Henter, J.-I., Kabisch, H., Schneppenheim, R., Nürnberg, P., Janka, G., et al. (2005). Linkage of familial hemophagocytic lymphohistiocytosis (FHL) type-4 to chromosome 6q24 and identification of mutations in syntaxin 11. *Hum. Mol. Genet.* 14, 827–834.

zur „, U., Rohr, J., Seifert, W., Koch, F., Grieve, S., Pagel, J., Strauß, J., Kasper, B., Nürnberg, G., Becker, C., et al. (2009). Familial Hemophagocytic Lymphohistiocytosis Type 5 (FHL-5) Is Caused by Mutations in Munc18-2 and Impaired Binding to Syntaxin 11. *The American Journal of Human Genetics* 85, 482–492.

(2010). E pluribus unum. *Nat Meth* 7, 331–331.