Genetics of Cellular Immunodeficiencies

2

Natalie Frede

Abstract

Primary immunodeficiencies (PIDs) constitute a rapidly expanding field. As of 2019 there are 430 distinct human inborn errors of immunity listed in the current IUIS classification (International Union of Immunological Societies) (Tangye et al., J Clin Immunol 40:24–64, 2020).

Within the last 20 years, hundreds of genetic defects underlying PID have been identified. Historically, these were classified into humoral immunodeficiencies, on the one hand, and cellular immunodeficiencies, on the other hand. However, as our understanding has been evolving, these categories have become increasingly complex: genetic defects, for example, can affect multiple cell populations simultaneously, may affect communication between different cell types, may constitute failure of the bone marrow or other compartments to support and regulate certain cell populations, or may be associated with syndromic features. Within this chapter, we aim to provide an overview over the basic concepts of genetics as well as genetics of cellular immunodeficiencies.

Keywords

Genetics · Gene defect · Mutation · Variant · Sequencing · Primary immunodeficiencies · Cellular immunodeficiencies

N. Frede (✉)
Department of Rheumatology and Clinical Immunology, Medical Center—University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
e-mail: natalie.frede@uniklinik-freiburg.de

M. M. D'Elios et al. (eds.), *Cellular Primary Immunodeficiencies*, Rare Diseases of the Immune System, https://doi.org/10.1007/978-3-030-70107-9_2

2.1 Introduction

Primary immunodeficiencies (PIDs) constitute a rapidly expanding field. As of 2019 there are 430 distinct human inborn errors of immunity listed in the current IUIS classification (International Union of Immunological Societies) [1].

Within the last 20 years, hundreds of genetic defects underlying PID have been identified. Historically, these were classified into humoral immunodeficiencies, on the one hand, and cellular immunodeficiencies, on the other hand. However, as our understanding has been evolving, these categories have become increasingly complex: genetic defects, for example, can affect multiple cell populations simultaneously, may affect communication between different cell types, may constitute failure of the bone marrow or other compartments to support and regulate certain cell populations, or may be associated with syndromic features. Within the IUIS classification, this is illustrated by the notion to classify PIDs into ten general categories, each with numerous distinct subcategories [1].

2.2 Importance of Establishing a Genetic Diagnosis

Establishing a genetic diagnosis is essential for confirming a suspected diagnosis and counseling PID patients and their family members regarding treatment options, prognosis, as well as family planning. The in-depth molecular characterization of underlying genetic defects and involved pathways has greatly advanced our knowledge of immunology, in general, and furthered our understanding of specific molecular principles of infection control but also of the termination of immune responses once the infection is controlled. For invasive therapies, such as hematopoietic stem cell transplantation or gene therapy approaches, the prior identification of a molecular defect is an essential prerequisite. A better understanding of defective or involved pathways may furthermore help to identify novel therapeutic targets, both for aiding patients with primary immunodeficiencies and for the development of novel immunosuppressive or antiproliferative drugs used in autoimmune diseases and malignancies, as experienced with JAK inhibitors or the BTK antagonist ibrutinib [2].

2.3 A Brief History of Genetics

The origins of genetics can be found within Augustinian monk Gregor Mendel's experiments on plant hybridization published in 1866, in which he described his observations on the heritability of different traits in peas [3]. However, it took until the early 1900s for the chromosomal theory of heredity to develop, combining Mendel's laws with the idea of chromosomes as the carriers of hereditary information [4]. Nucleic acid was first discovered in 1869 by the Swiss doctor Friedrich Miescher, though its significance was unclear at the time. It took until 1944 for Avery, McLeod, and McCarthy to show that, in fact, DNA constitutes the genetic material of a cell [5]. In 1953 Watson and Crick famously discovered the structure

of DNA and are since often hailed as the founders of modern genetics [6]. Meanwhile the notion of gene mutations as local alterations of a chromosome was derived in the 1920s from Hermann J. Muller's mutagenesis experiments on drosophila. The first description of a human mutation is generally attributed to V.M. Ingram, who described an amino acid exchange as a cause for sickle cell anemia in 1956 [7]. The first molecular defect underlying a primary immunodeficiency was identified by E.R. Giblett with ADA-SCID in 1972 [8], when she described two young girls with recurrent infections without measurable adenosine deaminase enzyme activity in their erythrocytes while their parents had approximately half-normal levels, from which an autosomal recessive inheritance pattern was correctly assumed.

Initial sequencing approaches constituted protein sequencing methods, allowing for the identification of the amino acid sequence of proteins from the 1950s on. At this time the exact mechanisms of transcription and translation had not yet been identified. Subsequently early RNA sequencing methods were developed, while the classical Sanger DNA sequencing method of dideoxy-chain termination was finally published in 1977 by Fred Sanger [9]. The era of PID genetics finally started in the early 1990s with the discovery of BTK mutations as the genetic cause of X-linked agammaglobulinemia more than 40 years after Bruton's first report of the disease in 1952 [10, 11]. Since then, advances in technology including automated sequencing and next-generation sequencing have enabled a rapid growth of the field with numerous exciting discoveries.

2.4 Patterns of Inheritance

Human inborn errors of immunity constitute by definition germline monogenic defects and thus follow Mendelian rules of inheritance. Specifically, monogenic defects can follow an autosomal dominant, autosomal recessive, or X-linked pattern of inheritance.

In the case of an autosomal dominant mode of inheritance, only one affected allele is sufficient to cause disease (i.e., a heterozygous mutation). Males and females will be affected at equal frequency, and the disease will not skip any generations. Autosomal dominant mutations can affect the expression and function of the gene product in different ways, namely, may cause haploinsufficiency, may result in gain of function, or may have a dominant negative effect. Haploinsufficiency arises when a single wild-type allele will not lead to sufficient expression of the gene product, usually a protein, and thus leads to a phenotypic effect. Gain-of-function mutations lead to an increased level of activity, novel function, or prolonged life span of the gene product, thus leading to a gain of function. Gain-of-function mutations are generally less common than loss-of-function mutations. Dominant negative mutations lead to a gene product, which will act antagonistically to the wild-type gene product within the same cell, also referred to as antimorphic mutations. Dominant negative mutations will thus reduce effective expression or function by greater than 50%. A dominant negative effect commonly arises in proteins, which form polymeric structures.

In autosomal recessive inheritance, two copies of an affected allele are required for an individual to express the disease phenotype (i.e., a homozygous or compound heterozygous mutation). Males and females will be similarly affected. Typically, both parents are heterozygous carriers, and statistically one-fourth of their children will be affected. In consanguineous unions, conditions with autosomal recessive inheritance will appear with an increased frequency due to common ancestry alleles. Autosomal recessive mutations generally result in loss of function.

In the case of X-linked inheritance usually only males are affected as they only possess one gene copy. Transmission occurs through female carriers, who are mostly phenotypically unaffected (no father-to-son transmission). The disease often occurs in multiple generations.

Whereas human inborn errors of immunity are generally defined through an underlying germline monogenic defect, some primary immunodeficiencies seem to have a complex pattern of inheritance, also referred to as polygenic inheritance. In particular, IgA deficiency and some forms of common variable immunodeficiency, both, however, not constituting cellular immunodeficiencies, seem to follow a complex inheritance pattern with an increased incidence in some families. The genetics of antibody deficiency was described in detail in the first book of this series, Humoral Primary Immunodeficiencies [12]. A list with explanations of genetic terms can be found in (Table 2.1).

Table 2.1 Genetic defects leading to cellular immunodeficiency, modified from Tangye et al. [1] (Springer OA BY CC License 4.0)

	Defect	Gene	Inheritance
Severe combined immunodeficiencies (see Chap. 6 of this book)			
T-B+ SCID	Gamma chain deficiency	*IL2RG*	XL
	JAK3 deficiency	*JAK3*	AR
	IL7Ra deficiency	*IL7R*	AR
	CD45 deficiency	*PTPRC*	AR
	CD3D deficiency	*CD3D*	AR
	CD3E deficiency	*CD3E*	AR
	CD3Z deficiency	*CD3Z*	AR
	Coronin 1A deficiency	*CORO1A*	AR
	LAT deficiency	*LAT*	AR
T-B− SCID	RAG1 deficiency	*RAG1*	AR
	RAG2 deficiency	*RAG2*	AR
	Artemis deficiency	*DCLRE1C*	AR
	DNA PKcs deficiency	*PRKDC*	AR
	Cernunnos/XLF deficiency	*NHEJ1*	AR
	DNA ligase 4 deficiency	*LIG4*	AR
	Adenosine deaminase (ADA) deficiency	*ADA*	AR
	AK2 defect	*AK2*	AR
	Activated RAC2 defect	*RAC2*	AD
Combined immunodeficiencies less severe than SCID (see Chap. 7 of this book)			

Table 2.1 (continued)

	Defect	Gene	Inheritance
	CD40 ligand deficiency	*CD40LG*	XL
	CD40 deficiency	*CD40*	AR
	ICOS deficiency	*ICOS*	AR
	ICOSL deficiency	*ICOSLG*	AR
	CD3γ deficiency	*CD3G*	AR
	CD8 deficiency	*CD8A*	AR
	ZAP-70 deficiency	*ZAP70*	AR (LOF or GOF)
	MHC class I deficiency	*TAP1*	AR
		TAP2	AR
		TAPBP	AR
		B2M	AR
	MHC class II deficiency	*CIITA*	AR
		RFXANK	AR
		RFX5	AR
		RFXAP	AR
	IKAROS deficiency	*IKZF1*	AD
	DOCK8 deficiency	*DOCK8*	AR
	DOCK2 deficiency	*DOCK2*	AR
	Polymerase D deficiency	*POLD1*	AR
		POLD2	AR
	RHOH deficiency	*RHOH*	AR
	STK4 deficiency	*STK4*	AR
	TCRα deficiency	*TRAC*	AR
	LCK deficiency	*LCK*	AR
	ITK deficiency	*ITK*	AR
	MALT1 deficiency	*MALT1*	AR
	CARD11 deficiency	*CARD11*	AR
	BCL10 deficiency	*BCL10*	AR
	IL-21 deficiency	*IL21*	AR
	IL-21 receptor deficiency	*IL21R*	AR
	OX40 deficiency	*TNFRSF4*	AR
	IKBKB deficiency	*IKBKB*	AR
	NIK deficiency	*MAP3K14*	AR
	RelB deficiency	*RELB*	AR
	RelA haploinsufficiency	*RELA*	AD
	Moesin deficiency	*MSN*	XL
	TFRC deficiency	*TFRC*	AR
	c-Rel deficiency	*REL*	AR
	FCHO1 deficiency	*FCHO1*	AR

(continued)

Table 2.1 (continued)

	Defect	Gene	Inheritance
Combined immunodeficiencies with associated or syndromic features (see Chap. 8 of this book)			
Immunodeficiency with congenital thrombocytopenia	Wiskott-Aldrich syndrome	*WAS*	XL
	WIP deficiency	*WIPF1*	AR
	Arp2/3-mediated filament branching defect	*ARPC1B*	AR
Other DNA repair defects	Ataxia-telangiectasia	*ATM*	AR
	Nijmegen breakage syndrome	*NBS1*	AR
	Bloom syndrome	*BLM*	AR
	Immunodeficiency with centromeric instability and facial anomalies	*DNMT3B*	AR
		ZBTB24	AR
		CDCA7	AR
		HELLS	AR
	PMS2 deficiency	*PMS2*	AR
	RIDDLE syndrome/RNF168 deficiency	*RNF168*	AR
	MCM4 deficiency	*MCM4*	AR
	Polymerase E subunit 1 deficiency (FILS syndrome)	*POLE1*	AR
	Polymerase E subunit 2 deficiency	*POLE2*	AR
	Ligase I deficiency	*LIG1*	AR
	NSMCE3 deficiency	*NSMCE3*	AR
	ERCC6L2/Hebo deficiency	*ERCC6L2*	AR
	GINS1 deficiency	*GINS1*	AR
Thymic defects with additional congenital anomalies	DiGeorge/chromosome 22q11.2 deletion syndrome	*Deletion in 22q11.2 (typically spanning TBX1)*	AD
	TBX1 deficiency	*TBX1*	AD
	CHARGE syndrome	*CHD7*	AD
		SEMA3E	AD
	Winged helix nude FOXN1 deficiency/haploinsufficiency	*FOXN1*	AR/AD
	Chromosome 10p13-p14 deletion syndrome	*Del10p13-p14*	AD
	Chromosome 11q deletion syndrome (Jacobsen syndrome)	*11q23del*	AD
Immuno-osseous dysplasias	Cartilage hair hypoplasia	*RMRP*	AR
	Schimke immuno-osseous dysplasia	*SMARCAL1*	AR
	MYSM1 deficiency	*MYSM1*	AR
	MOPD1 deficiency (Roifman syndrome)	*RNU4ATAC*	AR
	Immunoskeletal dysplasia with neurodevelopmental abnormalities (EXTL3 deficiency)	*EXTL3*	AR

Table 2.1 (continued)

	Defect	Gene	Inheritance
Hyper-IgE syndromes (HIES)	STAT3 deficiency (job syndrome)	*STAT3*	AD
	IL6 receptor deficiency	*IL6R*	AR
	IL6 signal transducer (IL6ST) deficiency	*IL6ST*	AR
	ZNF341 deficiency	*ZNF341*	AR
	ERBIN deficiency	*ERBB2IP*	AD
	Loeys-Dietz syndrome (TGFBR deficiency)	*TGFBR1*	AD
		TGFBR2	AD
	Comel-Netherton syndrome	*SPINK5*	AR
	PGM3 deficiency	*PGM3*	AR
	CARD11 deficiency	*CARD11*	AD
Defects of vitamin B12 and folate metabolism	Transcobalamin 2 deficiency	*TCN2*	AR
	SLC46A1/PCFT deficiency	*SLC46A1*	AR
	Methylene-tetrahydrofolate dehydrogenase 1 (MTHFD1) deficiency	*MTHFD1*	AR
Anhidrotic ectodermal dysplasia with immunodeficiency	NEMO/IKBKG deficiency	*IKBKG*	XL
	IKBA GOF mutation	*NFKBIA*	AD GOF
	IKBKB GOF mutation	*IKBKB*	AD GOF
Calcium channel defects	ORAI-1 deficiency	*ORAI1*	AR
	STIM1 deficiency	*STIM1*	AR
Other defects	Purine nucleoside phosphorylase (PNP) deficiency	*PNP*	AR
	Immunodeficiency with multiple intestinal atresias	*TTC7A*	AR
	Tricho-hepato-enteric syndrome	*TTC37*	AR
		SKIV2L	AR
	Hepatic veno-occlusive disease with immunodeficiency (VODI)	*SP110*	AR
	BCL11B deficiency	*BCL11B*	AD
	EPG5 deficiency (Vici syndrome)	*EPG5*	AR
	HOIL1 deficiency	*RBCK1*	AR
	HOIP deficiency	*RNF31*	AR
	Hennekam lymphangiectasia-lymphedema syndrome	*CCBE1*	AR
		FAT4	AR
	Activating de novo mutations in nuclear factor, erythroid 2 like (NFE2L2)	*NFE2L2*	AD
	STAT5b deficiency	*STAT5B*	AR/AD
	Kabuki syndrome	*KMT2D*	AD
		KDM6A	XL
	KMT2A deficiency (Wiedemann-Steiner syndrome)	*KMT2A*	AD

(continued)

Table 2.1 (continued)

	Defect	Gene	Inheritance
Congenital defects of phagocytes (see Chap. 9 of this book)			
Congenital neutropenias	Elastase deficiency/severe congenital neutropenia 1(SCN1)	*ELANE*	AD
	GFI 1 deficiency (SCN2)	*GFI1*	AD
	HAX1 deficiency (Kostmann disease) (SCN3)	*HAX1*	AR
	G6PC3 deficiency (SCN4)	*G6PC3*	AR
	VPS45 deficiency (SCN5)	*VPS45*	AR
	Glycogen storage disease type 1b	*G6PT1*	AR
	X-linked neutropenia/ myelodysplasia	*WAS*	XL GOF
	P14/LAMTOR2 deficiency	*LAMTOR2*	AR
	Barth syndrome (3-methylglutaconic aciduria type II)	*TAZ*	XL
	Cohen syndrome	*VPS13B*	AR
	Clericuzio syndrome (Poikiloderma with neutropenia)	*USB1*	AR
	JAGN1 deficiency	*JAGN1*	AR
	3-Methylglutaconic aciduria	*CLPB*	AR
	G-CSF receptor deficiency	*CSF3R*	AR
	SMARCD2 deficiency	*SMARCD2*	AR
	Specific granule deficiency	*CEBPE*	AR
	Shwachman-diamond syndrome	*SBDS*	AR
		DNAJC21	AR
		EFL1	AR
	HYOU1 deficiency	*HYOU1*	AR
	SRP54 deficiency	*SRP54*	AD
Defects of motility	Leukocyte adhesion deficiency type 1 (LAD1)	*ITGB2*	AR
	Leukocyte adhesion deficiency type 2 (LAD2)	*SLC35C1*	AR
	Leukocyte adhesion deficiency type 3 (LAD3)	*FERMT3*	AR
	Rac2 deficiency	*RAC2*	AD LOF
	β actin deficiency	*ACTB*	AD
	Localized juvenile periodontitis	*FPR1*	AR
	Papillon-Lefèvre syndrome	*CTSC*	AR
	WDR1 deficiency	*WDR1*	AR
	Cystic fibrosis	*CFTR*	AR
	MKL1 deficiency	*MKL1*	AR
Defects of respiratory burst	X-linked chronic granulomatous disease (CGD), gp91phox	*CYBB*	XL
	Autosomal recessive CGD	*CYBA*	AR
		CYBC1	AR
		NCF1	AR
		NCF2	AR
		NCF4	AR
	G6PD deficiency class I	*G6PD*	XL

Table 2.1 (continued)

	Defect	Gene	Inheritance
Other nonlymphoid defects	GATA2 deficiency	*GATA2*	AD
	Pulmonary alveolar proteinosis	*CSF2RA*	XL
		CSFR2B	AR
Defects of innate and intrinsic immunology (see Chap. 10 of this book)			
Mendelian susceptibility to mycobacterial disease	IL-12 and IL-23 receptor β1 chain deficiency	IL12RB1	AR
	IL-12p40 (IL-12 and IL-23) deficiency	IL12B	AR
	IL-12Rβ2 deficiency	IL12RB2	AR
	IL-23R deficiency	IL23R	AR
	IFN-γ receptor 1 deficiency	IFNGR1	AR/AD
	IFN-γ receptor 2 deficiency	IFNGR2	AR
	STAT1 deficiency	STAT1	AD LOF
	Macrophage gp91 phox deficiency	CYBB	XL
	IRF8 deficiency	IRF8	AD
	SPPL2a deficiency	SPPL2A	AR
	Tyk2 deficiency	TYK2	AR
	ISG15 deficiency	ISG15	AR
	RORγt deficiency	RORC	AR
	JAK1 deficiency	JAK1	AR
Epidermodysplasia verruciformis (HPV)	EVER1 deficiency	TMC6	AR
	EVER2 deficiency	TMC8	
	CIB1 deficiency	CIB1	
	WHIM (warts, hypogammaglobulinemia, infections, myelokathexis) syndrome	CXCR4	AD GOF
Predisposition to severe viral infection	STAT1 deficiency	STAT1	AR LOF
	STAT2 deficiency	STAT2	AR
	IRF9 deficiency	IRF9	AR
	IRF7 deficiency	IRF7	AR
	IFNAR1 deficiency	IFNAR1	AR
	IFNAR2 deficiency	IFNAR2	AR
	CD16 deficiency	FCGR3A	AR
	MDA5 deficiency	IFIH1	AR
	RNA polymerase III deficiency	POLR3A	AD
		POLR3C	AD
		POLR3F	AD
Predisposition to herpes simplex encephalitis	TLR3 deficiency	TLR3	AR/AD
	UNC93B1 deficiency	UNC93B1	AR
	TRAF3 deficiency	TRAF3	AD
	TRIF deficiency	TICAM1	AD/AR
	TBK1 deficiency	TBK1	AD
	IRF3 deficiency	IRF3	AD
	DBR1 deficiency	DBR1	AR

(continued)

Table 2.1 (continued)

	Defect	Gene	Inheritance
Predisposition to recurrent fungal infections	CARD9 deficiency	CARD9	AR
	IL-17RA deficiency	IL17RA	AR
	IL-17RC deficiency	IL17RC	AR
	IL-17F deficiency	IL17F	AD
	STAT1 GOF	STAT1	AD GOF
	ACT1 deficiency	TRAF3IP2	AR
TLR signaling pathway deficiency with bacterial susceptibility	IRAK4 deficiency	IRAK4	AR
	MyD88 deficiency	MYD88	AR
	IRAK1 deficiency	IRAK1	XL
	TIRAP deficiency	TIRAP	AR
Other inborn errors of immunity related to non-hematopoietic tissues	Isolated congenital asplenia (ICA)	RPSA	AD
		HMOX	AR
	Trypanosomiasis	APOL1	AD
	Acute liver failure due to NBAS deficiency	NBAS	AR
	Acute necrotizing encephalopathy	RANBP2	AR
	Osteopetrosis	CLCN7	AR
		SNX10	AR
		OSTM1	AR
		PLEKHM1	AR
		TCIRG1	AR
		TNFRSF11A	AR
		TNFSF11	AR
	Hidradenitis suppurativa	NCSTN	AD
		PSEN	AD
		PSENEN	AD
Other inborn errors of immunity related to leukocytes	IRF4 haploinsufficiency	IRF4	AD
	IL-18BP deficiency	IL18BP	AR

Abbreviations: *XL* X-linked, *AR* Autosomal recessive, *AD* Autosomal dominant, *GOF* Gain of function, *LOF* Loss of function

2.5 Germline Versus Somatic Mutations

A **germline** mutation is defined as a heritable mutation, which occurred originally in a germ cell or the zygote at single-cell stage, and thus will be present in all cells in the offspring, including the germ cells. Thus, a germline mutation can be passed on from generation to generation.

In contrast to germline mutations, somatic mutations occur when a mutation arises postzygotically, which will lead to mosaicism. The phenotype of mosaicism depends upon the developmental stage at which the mutation arises. A mutation in early embryonic development will likely affect many different tissues. If only

somatic cells are affected, this will be called somatic mosaicism. Somatic mosaicism will not be transmitted to the next generation. If both somatic cells and germ cells are affected, this is defined as gonosomal mosaicism, which will cause symptoms depending on the affected somatic tissues and may be transmitted on to the next generation. In contrast, gonadal mosaicism affects only the germ cells and, thus, typically has no phenotype in the carrier but will be transmitted to the offspring as a germline mutation. Somatic mutations are generally considered phenocopies of PID within the IUIS classification, although we want to point out that mosaicism may in theory cause a phenotype indistinguishable from germline mutations and, if including the germ cells, can lead to full germline mutations in the offspring. Mosaicism should be suspected in case of marked intergenerational phenotypical differences, unexpected intrafamilial reoccurrence in children of seemingly healthy parents without mutation in Sanger sequencing, or unequal height/intensity of sequencing peaks in Sanger chromatograms [13].

2.6 Types of Mutations

Mutations can be classified either by their impact on DNA or by their impact on the respective protein.

Regarding the impact of a mutation on the DNA, a mutation can further be classified as a substitution, i.e., one or multiple bases are substituted by an equal number of other bases; an insertion, where additional bases are gained; or a deletion, i.e., a loss of bases. Inversions and translocations may also be listed in this category and constitute structural rearrangements of DNA. An inversion is defined as an end-to-end reversion of a piece of DNA, whereas a translocation describes the integration of a piece of DNA or a piece of a chromosome at another position. Translocations can be balanced, i.e., an even exchange of DNA, or unbalanced, thus leading to loss or gain of genetic material in daughter cells.

Regarding the impact on the resulting protein, mutations can be classified as silent mutations, missense mutations, or frameshift mutations. A silent mutation is a mutation which does not lead to a change in the amino acid sequence of the protein. A missense mutation is defined as an exchange of a single amino acid within the protein sequence. A nonsense mutation constitutes a change of one amino acid within the protein into a premature stop codon. In contrast, a readthrough mutation constitutes the change of a stop codon into an amino acid. A frameshift mutation arises from a deletion or insertion of bases in the DNA sequence of a number not divisible by three, thus changing the reading frame as the genetic code is organized in base triplets, i.e., codons, each coding for a specific amino acid. Also mutations outside of the coding DNA sequence may have an impact on the protein. Splice site mutations have long been recognized as disease causing; however, deep intronic mutations may have an impact on the protein sequence or expression through inclusion of pseudo-exons due to creation of alternative splice sites or may affect expression when located within regulatory regions, such as promoter or enhancer sequences [14].

2.7 Pleiotropy of PIDs

Many different genetic defects have overlapping phenotypes. Genetic heterogeneity can generally be classified into allelic heterogeneity and locus heterogeneity. Allelic heterogeneity is defined as different mutations (alleles) within the same gene producing a similar phenotype. In contrast, locus heterogeneity implies that a similar phenotype may be caused by mutations in different genes. Similarly, the term genocopy refers to a genotype or mutation resulting in a similar phenotype to another genotype or mutation at a different locus. In contrast, a phenocopy is defined as environmental factors producing the same phenotype as a specific genetic mutation, thus mimicking the phenotype [15]. A phenocopy by definition is not a genetic trait and thus is not hereditary in a strict Mendelian sense, though epigenetic changes may count as phenocopies and can be passed on to daughter cells.

Differences in phenotype despite the same genotype may be caused by a variable expressivity of a trait or phenotype. Expressivity thus constitutes a measure for the extent of phenotypic expression. In contrast, penetrance refers to the proportion of individuals with a certain genotype, who exhibit the associated phenotype. With complete penetrance all individuals with a certain genotype, i.e., a certain mutation, show the associated symptoms/trait, whereas reduced penetrance means that some individuals who carry a genetic defect may in fact be phenotypically healthy.

2.8 Sequencing Technologies

In the past Sanger sequencing constituted the gold standard for genetic diagnostics. Depending on the clinical phenotype, the most likely candidate genes needed to be identified and sequenced sequentially exon by exon. However, this constituted a laborious time- and resource-consuming process and often did not lead to success due to atypical presentations and obvious limitations due to being a hypothesis-driven approach (i.e., the candidate gene needed to be known). In recent years, great advances have been made with the help of next-generation sequencing techniques, which have substituted Sanger sequencing in the diagnostic workup process of PIDs in many places.

Next-generation sequencing technologies allow for the simultaneous massive parallel sequencing of thousands of genes at dramatically reduced costs. With many novel sequencing techniques, several patients can be multiplexed and thus sequenced in the same sequencing run. The introduction of next-generation sequencing methods thus has greatly facilitated the identification of novel genetic defects, which is reflected in the rising number of novel defects described every year. In general, next-generation sequencing methods consist of the following three steps: The first step is the preparation of a library, in which the DNA is fragmented (usually through either restriction enzymes or sonication) and fragments are ligated with custom linkers or sequencing adapters. In panel and exome sequencing, there is an amplification step relying on clonal amplification/PCR, whereas whole genome sequencing, in general, does not necessitate amplification. Lastly, the fragments are

sequenced. Depending on the sequencer, different technologies are employed for this step.

Next-generation sequencing approaches can be divided into panel sequencing approaches, whole exome sequencing, and whole genome sequencing, each with their distinct advantages and disadvantages.

Gene panel sequencing provides a high coverage of sequenced regions, which is one of its distinct advantages over whole exome and especially whole genome sequencing and is crucial to reduce errors, particularly in a diagnostic setting. Since less regions are sequenced, panel sequencing results in less variants of unknown significance. Limitations include that panel sequencing is a biased approach (variants can only be detected in sequenced regions, i.e., in identified target genes), preassembled panels are rigid and may not contain all genes of interest, and continuous redesigning and revalidation of the panel may be necessary.

Whole exome sequencing constitutes an unbiased approach, allowing the identification of novel defects. However, in practice, complete coverage of all coding exons is impracticable, and a significant proportion of regions will have a low read depth, necessitating resequencing for use in a clinical context.

Whole genome sequencing is the only method which also allows for the identification of deep intronic variants. However, data interpretation is still difficult and hampered by detection of vast numbers of variants of unknown significance.

All next-generation sequencing approaches necessitate bioinformatic analysis in order to align the sequenced fragments correctly to the reference genome and identify variants. Detected variants subsequently need to be compared with reference databases and evaluated for harmfulness through either prediction tools or experimental validation.

2.9 Interpretation of Sequencing Results

A drawback of the novel sequencing technologies is the detection of numerous variants of unclear significance. Sometimes it may be difficult to establish whether a variant is in fact disease causing or not, which may pose clinical as well as at times ethical and legal challenges.

Sequence variants are common, often benign, and the source of genetic variation. In fact, each genome is thought to have approximately 4 million sequence variants, which mostly constitute single-nucleotide polymorphisms (SNP, by definition frequency above 1%) but may also encompass, e.g., structural variants [16]. The relative occurrence of sequence variants varies between regions and genes, with important functional domains of genes often showing evolutionary conservation through many different species.

To establish whether a variant may be benign or disease causing, several methods can be employed. Firstly, to establish whether a variant has been reported before or is listed as a SNP, bioinformatic analysis including queries of population databases can

be employed; useful resources include, e.g., dbSNP, Human Gene Mutation Database (HGMD), ClinVar, and gnomAD (formerly ExAc) [17–20]. If the variant has not been reported as disease associated before, in silico prediction tools, such as CADD, PolyPhen2, or SIFT, may provide helpful insights [21–23]. Additionally, genetic analysis of affected and unaffected family members as well as functional testing including cloning of the mutation may be performed but constitute laborious processes.

To facilitate the classification and interpretation of sequencing results, the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology have published recommendations [24]. In particular, they recommend to classify variants into the following categories: (i) pathogenic, (ii) likely pathogenic, (iii) uncertain significance, (iv) likely benign, or (v) benign. Furthermore, they advise to use a standardized nomenclature of variants, as published and regularly updated by the Human Genome Variation Society (HGVS) [25].

As a general rule, variants should always be described at the most basic level possible, which will normally be the DNA level (e.g., c.1650C>T, where ">" is used to label a base substitution and "del," "ins," and "dup" label deletions, insertions, and duplications, respectively). However, in the case of, e.g., frameshifts, the protein sequence may be used, such as p.R908Kfs*15, indicating that the variant leads to an amino acid exchange from R to K in position 908 with a subsequent frameshift (fs) and termination after 15 amino acids. Importantly, a reference sequence or transcript should always be provided. The online tool Mutalyzer (https://mutalyzer.nl) provides both a name generator producing a valid HGVS variant description and a syntax checker to assess existing variant descriptions regarding compatibility with HGVS nomenclature [26].

Furthermore, variants described in only a single patient should be taken with caution. Casanova et al. published criteria aiding the decision whether data establish a causal relationship between phenotype and genotype for this case [27]. In brief, the variant in question should not occur in healthy individuals, there need to be experimental data indicating that the gene product is altered in function or expression by the variant, and a causal relationship between genotype and phenotype must be confirmed in an animal model or a relevant cellular phenotype.

2.10 Genetics of Combined Immunodeficiencies

Combined immunodeficiencies can be classified into severe combined immunodeficiencies (SCID) and combined immunodeficiencies less severe than SCID (CID). In total, the current IUIS classification lists a total number of 50 disorders caused by 58 distinct genetic defects.

Clinically, SCID defects can be classified by presence or absence of T, B, and NK cells. Genetically, SCID defects can be grouped by pathogenic mechanism into VDJ recombination and T cell receptor defects, cytokine signaling defects, defects with toxic metabolite accumulation, and defects with defective survival of hematopoietic precursors. SCID usually follows an autosomal recessive or X-linked pattern of inheritance.

VDJ recombination defects constitute mainly enzymatic defects, which are inherited in an autosomal recessive pattern and include RAG1/RAG2 deficiency, DNA cross-link repair enzyme 1c (DCLRE1C)/Artemis deficiency, nonhomologous end-joining enzyme (NHEJ1)/Cernunnos deficiency, DNA ligase IV (LIG4) deficiency, and DNA PKcs deficiency (PRKDC). Recombination-activating proteins (RAG) 1 and 2 initiate recombination of immunoglobulin and T cell receptor genes in B and T cells to diversify the repertoire through rearrangement of variable (V), diversity (D), and joining (J) segments. Failure to induce VDJ recombination will lead to apoptosis. It is essential that the induced double-strand breaks subsequently are repaired correctly to produce a recombined continuous DNA strand. Defects in nonhomologous end-joining DNA repair mechanisms, such as mutations in DCLRE1C, DNA PKcs, NHEJ1, and LIG4, will thus lead to SCID with increased radiosensitivity as DNA damage cannot be sufficiently repaired. Other DNA repair defects leading to radiosensitivity and immunodeficiency regularly include syndromic features and thus belong to syndromic combined immunodeficiencies.

T cell receptor defects affecting CD3D, CD3E, and CD3Z lead to T-B+NK+ SCID, whereas CD3G deficiency leads to combined immunodeficiency, generally less profound than SCID. Also, CD45 deficiency may be counted toward the T cell receptor defects in that CD45 encodes a receptor-associated tyrosine phosphatase essential for the activation of T cells through the T cell receptor, its deficiency also leads to T-B + NK+ SCID. T cell receptor defects generally follow an autosomal recessive inheritance pattern.

Cytokine signaling defects causing SCID include common gamma chain deficiency (IL2RG), JAK3 deficiency, and IL7Rα deficiency (IL7RA). The common gamma chain is a receptor chain shared by the receptors of interleukins IL2, IL4, IL7, IL9, IL15, and IL21. As IL7 and IL15 are essential in the development of T and NK cells, common gamma chain deficiency leads to T-B + NK- SCID. Since the *IL2RG* gene is located on the X chromosome, common gamma chain deficiency follows an X-linked inheritance pattern and thus is also referred to as X-SCID. The Janus kinase 3 mediates the signal transduction downstream of common gamma chain cytokines; therefore, deficiency also leads to T-B+NK− SCID, however, following an autosomal recessive inheritance pattern. IL7 signaling is essential in the early development of lymphocytes but also proliferation and survival of T cells peripherally; thus, deficiency of the IL7 receptor alpha chain also causes T-B+NK+ SCID. Multiple other cytokine and signaling defects lead to less severe primary immunodeficiencies.

Defects with toxic metabolite accumulation include ADA and PNP deficiency. The adenosine deaminase (ADA) is an essential enzyme of the purine salvage pathway, mediating the deamination of adenosine and 2-deoxyadenosine into inosine and deoxyinosine. Deficiency of ADA leads to toxic accumulation of these metabolites, which results in lymphocyte apoptosis (T-B−NK− SCID). The purine nucleoside phosphorylase (PNP) constitutes another key enzyme within the purine salvage pathway downstream of ADA, deribosylating inosine to hypoxanthine and guanosine to guanine. PNP deficiency leads to toxic accumulation of deoxyguanosine and deoxyguanosine triphosphate, leading to apoptosis, mainly affecting the T cells

(T-B+NK− SCID). As PNP deficiency leads to neurological symptoms and autoimmunity, the IUIS classification lists it as a combined immunodeficiency with associated features.

Hypomorphic mutations in any of these SCID genes may lead to less severe phenotypes of leaky SCID/combined immunodeficiency. Detailed descriptions of severe combined immunodeficiencies and combined immunodeficiencies may be found in Chaps. 6 and 7 of this book, whereas Table 2.2 gives an overview of genetic defects associated with cellular immunodeficiencies (Table 2.2).

Table 2.2 Definitions of genetic terms

Patterns of inheritance	
Autosomal recessive inheritance	Two copies of an affected allele are required for an individual to express the disease phenotype (i.e., a homozygous or compound heterozygous mutation); may occur if gene is located on autosomal chromosome
Autosomal dominant inheritance	Only one affected allele is sufficient to cause disease (i.e., a heterozygous mutation); may occur if gene is located on autosomal chromosome
X-linked inheritance	Usually only males are affected as they only possess one gene copy. Transmission occurs through female carriers, who are mostly phenotypically unaffected; only occurs in genes located on X chromosome
Zygosity	
Homozygosity	Cell possesses two identical alleles of a particular gene, one inherited from each parent
Heterozygosity	Cell possesses two different alleles (one wild-type allele and one variant allele) of a particular gene
Compound heterozygosity	Cell possesses two different variant alleles but no wild-type allele of a particular gene
Hemizygosity	Cell possesses only one copy of a particular gene, either through location of that gene on sex chromosome (X,Y) or through loss of homologous chromosome
Types of mutations	
Substitution	Mutation, in which one or multiple bases are replaced with an equal number of other bases
Deletion	Mutation which leads to loss of bases
Insertion	Mutation which leads to gain of additional bases
Inversion	End-to-end reversion of a piece of DNA
Translocation	Integration of a piece of DNA or piece of a chromosome at another position
Silent mutation	Mutation which does not change the amino acid sequence of the protein
Missense mutation	Mutation which leads to exchange of a single amino acid within a protein sequence
Nonsense mutation	Mutation which leads to formation of premature stop codon
Readthrough mutation	Mutation which changes the stop codon into an amino acid
Frameshift mutation	Deletion or insertion of a base number not divisible by three, thus changing the reading frame

Table 2.2 (continued)

Patterns of inheritance	
Mosaicism and germline mutations	
Germline mutation	Heritable mutation, which occurred originally in a germ cell or zygote at single-cell stage and thus will be present in all cells in the offspring
Somatic mutation	Postzygotic mutation, which will lead to mosaicism of one of the three types listed below:
Somatic mosaicism	Affects only somatic cells, thus not transmitted to the next generation; will produce phenotype depending on affected cells/tissues
Gonosomal mosaicism	Affects both somatic cells and germ cells, may thus cause symptoms depending on affected somatic tissues, and may be transmitted to the next generation
Gonadal mosaicism	Affects only germ cells, thus, typically has no phenotype in the carrier but may be transmitted to the offspring as a germline mutation
Pleiotropy and heterogeneity	
Allelic heterogeneity	Different mutations (alleles) within the same gene produce a similar phenotype
Locus heterogeneity	Mutations in different genes produce a similar phenotype
Genocopy	Genotype or mutation, which produces a similar phenotype to another genotype or mutation at a different locus
Phenocopy	Environmental factors produce the same phenotype as a specific genetic mutation, thus mimicking the phenotype
Expressivity	Measure for the extent of phenotypic expression
Penetrance	Proportion of individuals with a certain genotype, who exhibit the associated phenotype

2.11 Genetics of Combined Immunodeficiencies with Associated or Syndromic Features

In most primary immunodeficiencies, the immunodeficiency is the most prominent clinical finding. In contrast, syndromic immunodeficiencies are characterized by associated syndromes or clinical findings taking a front role. Associated features may commonly affect the skeletal, nervous, or ectodermal development or function but may include almost any organ system. In contrast to most other classes of immunodeficiencies, not all syndromic immunodeficiencies are typically caused by a genetic defect in a single gene but may result from underlying cytogenetic abnormalities, i.e., abnormalities of chromosomal number or structure. DiGeorge syndrome caused by 22q11 deletions is the most well-known example. Cytogenetic abnormalities may be detected by karyotyping or fluorescence in situ hybridization (FISH).

As of 2019, there are a total number of 58 combined immunodeficiencies with associated or syndromic features comprising 62 distinct genetic defects [1]. These include immunodeficiencies with congenital thrombocytopenia (*WAS, WIPF1, ARPC1B*), other DNA repair defects (*ATM, NBS1, BLM, DNMT3B, CDCA7, HELLS, PMS2, RNF168, MCM4, POLE1, POLE2, LIG1, NSMCE3, ERCC6L2,*

GINS1), thymic defects with congenital abnormalities (22q11.2DS, TBX1, CHD7, SEMA3E, FOXN1, 10p13-p14DS, 11q23del), immuno-osseous dysplasias (*RMRP, SMARCAL1, MYSM1, RNU4ATAC, EXTL3*), hyper-IgE syndromes (*STAT3, IL6R, IL6ST, ZNF341, ERBB2IP, TGFBR1, TGFBR2, SPINK5, PGM3, CARD11*), defects in vitamin B12 and folate metabolism (*TCN2, SLC46A1, MTHFD1*), anhidrotic ectodermal dysplasia with immunodeficiency (*IKBKG, NFKBIA, IKBKB*), calcium channel defects (*ORAI1, STIM1*), and other defects (*PNP, TTC7A, TTC37, SKIV2L, SP110, BCL11B, EPG5, RBCK1, RNF31, CCBE1, FAT4, NFE2L2, STAT5B, KMT2D, KDM6A, KMT2A*). A detailed description of combined immunodeficiencies with associated or syndromic features may be found in Chap. 8 of this book.

2.12 Genetics of Defects in Intrinsic and Innate Immunity

Defects of innate immunity comprise multiple heterogeneous groups of defects, out of which some can be counted toward cellular immunodeficiency, others constitute defects of soluble factors such as complement factors, and yet other defects derive from a defective barrier function, which are defects of nonimmune cells, however, predisposing to infection. Furthermore, defects of innate immunity may have an impact on the adaptive immune system through impaired (co-)stimulation or antigen presentation.

In recent years many novel inborn errors of innate immunity have been described, often leading to an increased susceptibility to a narrow range of pathogens. These include Mendelian susceptibility to mycobacterial disease, predisposition to chronic mucocutaneous candidiasis or invasive fungal infections, predisposition to herpes simplex encephalitis, and other severe viral infections. A detailed description of defects of intrinsic and innate immunity may be found in Chap. 10 of this book.

A major subgroup within the defects of innate immunity are the congenital defects of phagocytes; thus, they are often listed separately, and also this book dedicated a separate chapter to them (for congenital defects of phagocytes, see Chap. 9). Congenital defects of phagocytes comprise congenital neutropenias (i.e., defects with reduced neutrophil numbers); defects of neutrophil function including motility, chemotaxis, and adhesion; defects of respiratory burst (chronic granulomatous disease); and other nonlymphoid defects. As of 2019, the IUIS recognizes 41 distinct phagocyte defects [1].

2.13 Outlook

While so far more than 430 genetic defects causing primary immunodeficiencies have been identified, there may be many more novel defects left to discover. In theory, the number of potential human inborn errors of immunity is only limited by the number of genes related to the immune system. The gene ontology (GO) database lists 2782 genes within the category “immune system process”; thus, there may be many discoveries of novel primary immunodeficiencies in the years to come.

As our understanding of molecular as well as regulatory processes evolves, phenocopies of PID with somatic mutations, polygenic traits, and also epigenetic changes, all not constituting PID in the narrower sense, might gain further importance, which may ultimately lead to changes in our understanding of the concept of what constitutes a primary immunodeficiency or an inborn error of immunity.

References

1. Tangye SG, Al-Herz W, Bousfiha A, Chatila T, Cunningham-Rundles C, Etzioni A et al (2020) Human inborn errors of immunity: 2019 update on the classification from the International Union of Immunological Societies Expert Committee. J Clin Immunol 40(1):24–64
2. Herman SE, Gordon AL, Hertlein E, Ramanunni A, Zhang X, Jaglowski S et al (2011) Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. Blood 117(23):6287–6296
3. Mendel G (1996) Experiments in plant hybridization (1865). Verhandlungen des naturforschenden Vereins Brünn) Available online
4. Morgan TH, Sturtevant AH, Bridges CB, Muller HJ (1915) The mechanism of Mendelian heredity. H. Holt, New York
5. Avery OT, Macleod CM, McCarty M (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J Exp Med 79(2):137–158
6. Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171(4356):737–738
7. Ingram VM (1956) A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. Nature 178(4537):792–794
8. Giblett ER, Anderson JE, Cohen F, Pollara B, Meuwissen HJ (1972) Adenosine-deaminase deficiency in two patients with severely impaired cellular immunity. Lancet 2(7786):1067–1069
9. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74(12):5463–5467
10. Vetrie D, Vorechovský I, Sideras P, Holland J, Davies A, Flinter F et al (1993) The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. Nature 361(6409):226–233
11. Tsukada S, Saffran DC, Rawlings DJ, Parolini O, Allen RC, Klisak I et al (1993) Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. Cell 72(2):279–290
12. D'Elios MM, Rizzi M (2019) Humoral Primary Immunodeficiencies. Springer, Cham
13. Mensa-Vilaró A, Bravo García-Morato M, de la Calle-Martin O, Franco-Jarava C, Martínez-Saavedra MT, González-Granado LI et al (2019) Unexpected relevant role of gene mosaicism in patients with primary immunodeficiency diseases. J Allergy Clin Immunol 143(1):359–368
14. Vaz-Drago R, Custódio N, Carmo-Fonseca M (2017) Deep intronic mutations and human disease. Hum Genet 136(9):1093–1111
15. Goldschmidt R (1935) Gen und Außeneigenschaft. Z Indukt Abstamm Vererbungsl 69(1):38–69
16. Marian AJ (2010) DNA sequence variants and the practice of medicine. Curr Opin Cardiol 25(3):182–185
17. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM et al (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29(1):308–311
18. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S et al (2018) ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46(D1):D1062–D10D7
19. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D et al (2017) The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res 45(D1):D840–D8D5

20. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q et al (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581(7809):434–443
21. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46(3):310–315
22. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P et al (2010) A method and server for predicting damaging missense mutations. Nat Methods 7(4):248–249
23. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4(7):1073–1081
24. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J 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(5):405–424
25. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J et al (2016) HGVS recommendations for the description of sequence variants: 2016 update. Hum Mutat 37(6):564–569
26. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PE (2008) Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. Hum Mutat 29(1):6–13
27. Casanova JL, Conley ME, Seligman SJ, Abel L, Notarangelo LD (2014) Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. J Exp Med 211(11):2137–2149