DEVELOPMENT OF MULTI-LOCI SEQUENCE-BASED GENETIC DIAGNOSTIC TESTING STRATEGY FOR SEVERE COMBINED IMMUNODEFICIENCY (SCID) PATIENTS IN MALAYSIA

SYARIFAH MASYITAH BINTI HABIB DZULKARNAIN

UNIVERSITI SAINS MALAYSIA

2024

DEVELOPMENT OF MULTI-LOCI SEQUENCE-BASED GENETIC DIAGNOSTIC TESTING STRATEGY FOR SEVERE COMBINED IMMUNODEFICIENCY (SCID) PATIENTS IN MALAYSIA

by

SYARIFAH MASYITAH BINTI HABIB DZULKARNAIN

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

February 2024

ACKNOWLEDGEMENT

In the name of Allah, The Most Gracious, The Most Merciful. I testify that there is none worthy of worship except Allah and that Muhammad is His slave and Messenger. Alhamdulillah, first and foremost all perfect praise be to Allah, the Lord of the worlds, for allowing me to embark on my PhD and for completing this journey successfully. I am incredibly thankful for the support and guidance of my wonderful supervisor, Dr Intan Juliana Abd Hamid for her dedication and limitless support and guidance to help me grow into an independent researcher. Her assistance, unconditional support, guidance, and facilitation throughout this process have been incredible. It is with sincerity to acknowledge my co-supervisor as well, Dr. Ilie Fadzilah Hashim, Dr Zarina Thasneem Zainudeen and Professor Dr Narazah Mohd Yusoff for their continuous support and guidance throughout the duration of my study.

The magnitude and scope of works involved in this research requires the support of many people to whom I indebted, through great or small their contributions. Special thanks to my beloved parents, husband, children, sisters and in-laws, for their patience, understanding, encouragement, love and prayer throughout my PhD journey. I sincerely thank my colleagues and staff of IPPT, USM and Universiti Teknologi MARA (UiTM) for their contributions in the accomplishment of this thesis.

Millions of thanks to the financial contributors: UiTM and Ministry of Higher Education Malaysia for providing the staff scholarship and USM for providing Short-Term Research Grant (USM: PO5053 2019/0724) in supporting the funding of my research work.

TABLE OF CONTENTS

APPENDICES

LIST OF PUBLICATIONS

LIST OF PRESENTATIONS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ACT	Beta-actin
AD	Autosomal dominant
ADA	Adenosine deaminase
ALB	Albumin
ALPS	Autoimmune Lymphoproliferative Syndrome
APC	Antigen presenting cell
AK2	Reticular dysgenesis
AR	Autosomal recessive
ATP	Adenosine triphosphate
BCR	B-cell receptor
BRCT1	BRCA1 C-terminus domain
CD	Cluster of differentiation
cDNA	Complementary DNA
CID	Combined Immunodeficiency
CLP	Common lymphoid progenitors
CMP	Common myeloid progenitors
CMV	Cytomegalovirus
Cq	Quantification cycle
Ct	Cycle threshold
CTD	Carboxy-terminal domain
CVID	Common variable immunodeficiency
dATP	Deoxyadenosine triphosphate
dGTP	Deoxyguanosine triphostate
DBD	DNA binding domain
DBS	Double strand break
DC	Dendritic cell
DCLRE1C	DNA cross-link repair enzyme 1c
DN	Double-negative
DdNTP	Dideoxynucleotide triphosphate
DP	Double-positive

EBV	Epstein-Barr virus
EDA-ID	Anhidrotic Ectodermodysplasia with Immunodeficency
EDTA	Ethylene diamine tetraacetic acid
ExAC	Exome Aggregation Consortium
FERM	Protein 4.1R, ezrin, radixin, moesin
FHL	Familial Hemophagocytic
FTT	Failure to thrive
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gDNA	Genomic DNA
GOI	Gene of interest
GOF	Gain of function
HIES	Hyper IgE Syndromes
HLA	Human leukocyte antigen
HPV	Human papillomaviruse
HSC	Haematopoietic stem cell
HSCT	Haematopoietic stem cells transplant
HSE	Herpes Simplex Encephalitis
IEI	Inborn errors of immunity
Ig	Immunoglobulin
IL2RG	Interleukin common gamma (γ) chain
IL7Rα	Interleukin-7 receptor alpha
IUIS	International Union of Immunological Societies
JAK3	Janus tyrosine kinase 3
JH	Janus homology
JMF	Jeffrey Modell Foundation
LIG4	DNA ligase 4
LOF	Loss of function
MCO	Movement Control Order
MHC	Major histocompatibility complex
mRNA	Messenger RNA
MSMD	Mendelian Susceptibility to mycobacterial disease
NBD	Nonamer-binding domain
NMD	Nonsense mediated decay

NBS	Newborn screening
NCBI	National Center for Biotechnology Information
NGS	Next-generation sequencing
NHEJ1	Non-homologous end-joining factor 1
NK	Natural killer
NTD	Nucleotidyltransferase
OBD	Oligonucleotide/oligosaccharide-fold
OD	Optical density
OS	Omenn syndrome
PAXX	Paralog of XRCC4 and XLF
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PHA	Phytohaemagglutinin
PID	Primary immunodeficiency disease
PIDTC	Primary Immune Deficiency Treatment Consortium
PRKDC	Protein kinase DNA-activated catalytic polypeptide
PNP	Purine nucleoside phosphorylase
qPCR	Quantitative polymerase chain reaction
RAG	Recombinase-activating genes
RAPID	Resource of Asian Primary Immunodeficiency Diseases
RBC	Red blood cell
RFU	Relative fluorescence units
RPII	RNA polymerase II
Rpm	Revolutions per minute
SCID	Severe combined immunodeficiency
SP	Single-positive
TCR	T cell receptor
TLR	Toll-like receptors
Tm	Melting temperature
TM	Transmembrane domain
TME	Transplacental maternal engraftment
TREC	T cell receptor excision circles
TTC7A	Tetratricopeptide repeat domain-7A

TYK2	Tyrosine kinase 2
USM	Universiti Sains Malaysia
WBC	White blood cell
WES	Whole exome sequencing
XRCC4	X-ray repair cross-complementing protein 4
ZAP70	Zeta-associated protein 70
ZBD	Zinc-binding domain

LIST OF APPENDICES

PEMBANGUNAN STRATEGI PENGUJIAN DIAGNOSTIK GENETIK BERASASKAN MULTI-LOKUS UNTUK PESAKIT IMUNODEFISIENSI GABUNGAN TERUK (SCID) DI MALAYSIA

ABSTRAK

Penyakit imunodefisiensi gabungan teruk (SCID) adalah penyakit yang jarang berlaku dan paling kritikal di antara semua jenis penyakit imunodefisiensi primer (PID) yang lain, yang menyebabkan pesakit terdedah kepada jangkitan yang serius dan risiko kematian. Penyakit ini disebabkan oleh beberapa faktor genetik yang berkaitan dengan kekukrangan sel T limfosit, sel B limfosit dan sel NK limfosit. Sebanyak 22 pesakit telah dilaporkan menghidap SCID di Malaysia, justeru menjadikan SCID antara pesakit PID terbanyak di Malaysia dengan 7 pesakit dilaporkan mempunyai diagnosis genetik. Diagnosis genetik amat penting dan merupakan piawaian utama bagi diagnosis SCID. Permasalahan utama di Malaysia adalah kekurangan kapasiti untuk mendiagnos penyakit SCID. Kekurangan kesedaran di kalangan pengamal perubatan menyebabkan diagnosis SCID kekal sebagai cabaran walaupun data diagnosis dan rawatan awal penyakit SCID menunjukkan keputusan yang memberangsangkan. Objektif utama tesis ini adalah untuk menghasilkan ujian yang berasaskan qPCR untuk panel gen SCID ke atas pesakit yang menghidap SCID yang melibatkan sebanyak 9 gen: *RAG1, RAG2, DCLRE1C, LIG4, ADA, IL2RG, JAK3, IL7Rα* dan *PNP*. Pembangunan ujian ini seterusnya akan disahkan melalui teknik berasaskan strategi jujukan Sanger untuk mengkaji variasi yang menyebabkan kerosakan gen PNP-SCID untuk satu kes kajian tesis ini yang mempunyai mutasi pada gen PNP. Matlamat ini terhasil dengan ujian berasaskan qPCR. Semua gen yang terlibat telah dinilai mengikut garis panduan MIQE yang melibatkan pengekstrakan DNA daripada sampel kawalan, penentuan kecekapan primer, lengkung piawaian, produk spesifik primer bagi gen panel SCID dan pemilihan gen rujukan melalui analisa qPCR menggunakan teknik celupan SYBR green dan juga perisian Normfinder dan perbandingan teknik ΔCt sebagai teknik analisa keputusan. Objektif kedua tesis ini adalah mengkaji kajian fenotip dan ujian fungsian keluarga pesakit PNP-SCID. Huraian lanjut turut dilakukan bagi memahami dengan lebih mendalam berkenaan spektrum PNP-SCID seterusnya teknik berasaskan jujukan Sanger yang digunakan secara menyeluruh dan menjimatkan untuk mengenalpasti gen yang novel dan mutasi yang terlibat dalam pesakit indeks PNP-SCID dan ahli keluarga terdekat berdasarkan perolehan yang terhasil pada objektif pertama. Tesis ini telah menghasilkan tiga dapatan utama. Pertama, hasil daripada objektif pertama yang bertujuan menghasilkan dan mengoptimumkan set pasangan primer terhadap gen yang terlibat. Perbezaan kepekatan genomik DNA dan primer yang menghasilkan lengkung amplikon tunggal telah dikenalpasti tanpa menghasilkan primer-dimer menunjukkan set primer qPCR sensitif terhadap setiap set primer. Sebagai tambahan, dengan menggunakan kondisi qPCR yang telah dioptimumkan melalui keputusan awal, ekspresi gen rujukan telah dikenalpasti dan gen GAPDH telah dipilih sebagai gen yang paling stabil di mana gen ini menunjukkan mempunyai nilai stabil terendah hasil daripada kajian gen rujukan. Kedua, pesakit juga telah dikenalpasti mempunyai sel T yang amat kurang, autoimun anemia hemolisis dan hipourisemia, di mana semuanya konsisten dengan diagnosis penyakit PNP-SCID. Ketiga, amplifikasi gen telah dibandingkan dan mendapati perbezaan ketara antara amplifikasi pesakit dan ibu bapanya menggunakan teknik Pfaffl. Seterusnya pengecaman varian mutasi homozigot pesakit dan mutasi heterozigot bagi pembawa iaitu ibu bapa pesakit telah dikenalpasti menggunakan analisa berasaskan jujukan Sanger di mana teknik ini mengesahkan teknik yang diaplikasi daripada objektif 1. Jujukan gen PNP menunjukkan novel mutasi homozigot, NM_000270.4: c.547dup (p.Glu183Glyfs*4), yang menyebabkan mutasi anjakan yang menghasilkan translasi penamatan kodon pra matang. Oleh itu, ujian hijau SYBR yang berasaskan qPCR yang menggunakan set primer daripada objektif pertama adalah cara yang efektif untuk mengenalpasti gen yang novel melalui ujian jujukan Sanger untuk pesakit SCID. Kesemua keputusan daripada kedua-dua objektif boleh digunakan untuk diagnosis penyakit SCID untuk bayi yang baru dilahirkan disebabkan pengesanan awal membawa kepada hasil yang memberangsangkan. Dengan kejayaan pengenalpastian mutasi novel gen yang melibatkan penyakit ini, keputusan kajian ini diharap dapat membongkar dan memberi pencerahan kepada kefahaman mengenai pembangunan strategi pengujian diagnostik genetik berasaskan multi-lokus bagi penyakit SCID yang membolehkan rawatan awal dan pengurusan klinikal segera dapat dilakukan justeru mengurangkan morbiditi dan mortaliti pesakit SCID di Malaysia.

DEVELOPMENT OF MULTI-LOCI SEQUENCE-BASED GENETIC DIAGNOSTIC TESTING STRATEGY FOR SEVERE COMBINED IMMUNODEFICIENCY (SCID) PATIENTS IN MALAYSIA

ABSTRACT

Severe combined immunodeficiencies (SCID) is a rare and the most critical form of primary immunodeficiencies (PID), which leads to extreme susceptibility to serious infections and lethal disorder due to the absence of T lymphocyte, B lymphocyte, and in many cases natural killer lymphocyte function. Twenty-two of the PID patients reported in Malaysia had SCID, making it the most prevalent recorded PID diagnosis with seven of these patients had a genetic diagnosis of SCID. Genetic testing is essential and represents the gold standard in the diagnosis of SCID. The primary gap in Malaysia is the inadequate genetic diagnostic capacity for SCID diagnosis. Due to the lack of awareness among healthcare professionals, SCID diagnosis remains a challenge despite the availability of data showing favourable outcomes for those who are diagnosed and treated early. The central objective of this thesis was to establish a qPCR-based assay of SCID gene panel based on patients investigated for SCID that involving 9 genes: *RAG1*, *RAG2*, *DCLRE1C*, *LIG4*, *ADA*, *IL2RG*, *JAK3*, *IL7R*α and *PNP* by using qPCR-based assay. The establishment of this assay would in turn validated by using Sanger sequencing-based strategy for investigating the variant that causing the defective gene of PNP-SCID from our case study of a female child who has a mutation of *PNP*. This goal was achieved by using qPCR-based assay. The targeted genes were evaluated following MIQE guidelines which involved DNA extraction from control sample, the determination of primer efficiency, standard curve, and product specificity of the primers of the SCID gene panel and the selection of candidate reference genes through qPCR analysis by using intercalating SYBR green dyes, Normfinder and comparative ΔCt method as the analysis methods. The second objective is to analyse the phenotype, variants and functional assay of family with PNP- SCID to better understand the spectrum of the disease followed by the analysis of Sanger sequencing that could be used as a comprehensive and cost-effective tool to discover novel gene and mutation in PNP-SCID of the index patient and the immediate family members using the findings established from the first objective. This thesis has achieved and reports upon three key findings. Firstly, findings from objective 1 that was aimed at establishing and optimising primer pairs set of all respective genes. Different DNA and primer concentrations were evaluated to obtain single curve amplicon without primer dimer formation for each respected gene indicated that the qPCR primer sets were sensitive to the respective primer set. In addition, using the optimised qPCR conditions obtained from previous finding, the amplification candidate reference genes were identified and the *GAPDH* gene was chosen as the most stable genes in this method as the gene showed to have the lowest stability value following the analysis of the reference gene. Secondly, the patient was found to have profound T cell lymphopenia, autoimmune haemolytic anaemia and hypouricemia, all consistent with PNP-SCID diagnosis. Thirdly, the gene amplification was compared and the result from this analysis showed significant differences in *PNP* amplification of the index patients compared to her parents using Pfaffl method. Subsequent identification of variant of the homozygous mutation of the index patient and heterozygous carrier state of the parents were described from Sanger sequencing analysis that validated the method in objective 1. Sequencing of *PNP* gene in the patient harboured a novel homozygous mutation, NM_000270.4: c.547dup (p.Glu183Glyfs*4), which leads to frameshift mutation that creates a premature translational stop codon. Therefore, the

SYBR green-base qPCR analysis using the primer pair obtained from objective 1 was the effective approach for the identification of novel candidate genes through this method of sequencing analysis. By successfully identifying novel candidate gene mutations which may be involved in the disease, it is hoped that the results of this work will improved the understanding of the multi-loci sequence-based genetic diagnostic testing strategy of the SCID gene that enables early diagnosis and prompts clinical management thereby decreasing the morbidity and mortality of SCID patients in Malaysia.

CHAPTER 1

INTRODUCTION

1.1 Background of study

Primary immunodeficiency diseases (PID) are immune system disorders caused by genetic abnormalities that disrupt the development, differentiation, and/or functioning of immune cells or signals. The hallmark of the diseases is their susceptibility to infections and is usually presented with recurrent infections. PID, also known as inborn errors of immunity (IEI), are heterogeneous group of genetic disorders that increase infection susceptibility, immunological dysregulation, autoimmunity, inflammation, allergy, and cancer (Abd Hamid *et al.*, 2020; Tangye *et al.*, 2020). The most severe type of PID is severe combined immunodeficiency (SCID), which is characterised by impaired T and/or B cell production or differentiation. The disease causes decreased lymphocyte function and significant immune system abnormalities (Gaspar *et al.*, 2013; Shearer *et al.*, 2014; Bonilla *et al.*, 2015); depending on the level of defect, the cells, tissues and proteins of the immune system.

SCID patients usually manifest with lymphopenia, recurrent and opportunistic infections, which may be fatal if untreated by the age of 1 year old. Most SCID infants are normal at birth, thus identifying a SCID newborn is difficult and this contributes to late diagnosis. Early detection of patients with PID is critically important as curative procedures such as haematopoietic stem cell transplantation (HSCT) can be performed.

Incidence of SCID reported globally ranged from 1 in 10,000 to 1 in 50,000 live births, with interleukin common gamma (γ) chain (IL2RG), an X-linked inheritance with a high prevalence in males, as the most prevalent mutation (Aloj *et* *al.*, 2012; Zhang *et al.*, 2017). The prevalence of SCID may be affected by a country's immunodiagnostic laboratory capacity and level of expertise. In Malaysia, a systematic review study revealed that the most frequently reported PID cases was SCID, affecting 22 patients, with 10 genetically diagnosed SCID patients, that includes common gamma chain cytokine receptor (*IL2RG*), purine nucleoside phosphorylase (*PNP*), adenosine deaminase (*ADA*), zeta-associated protein 70 (*ZAP70*) and tetratricopeptide repeat domain-7A (*TTC7A*) (Abd Hamid *et al.*, 2020). Apart from that, others include recombinase-activating genes 1 (*RAG1), RAG2*, DNA cross-link repair enzyme 1c (*DCLRE1C*), DNA Ligase IV (*LIG4*), Janus kinase 3 (*JAK3*) and Interleukine-7 receptor subunit alpha (*IL7R*α) that were mostly reported. Recently, a newborn screening programme for SCID has been introduced in the USA. The striking information gained from the programme was the SCID detection, with one in every 58 000 live births, which suggests that SCID might not be as rare as anticipated (Kwan *et al.*, 2014). To date, there is no national registry for PID in the country. Thus, the information on the actual incidence of PID or SCID in Malaysia is lacking.

Genetic diagnosis contributes significantly to determine the treatment option of SCID. The study of the molecular aetiology of SCID has provided important insight into the development and regulation of immune cells. SCID is caused by several genetic mutations affecting T and B cell activities, including those in recombinase-activating genes (*RAG1* and *RAG2*), interleukin-7 receptor alpha (*IL7Rα*), DNA cross-link repair enzyme 1c (*DCLRE1C*) genes, *IL2RG*, *ADA*, and Janus tyrosine kinase 3 (*JAK3*) genes (Al-Herz *et al.*, 2011; Gaspar *et al.*, 2013; Notarangelo, 2014). In this research, a multi-locus sequence based on genetic diagnostic testing of SCID was developed to understand the genotype profiles involving nine genes of SCID patients in Malaysia using quantitative polymerase chain reaction (qPCR) and Sanger sequencing.

1.2 Problem statement

Besides being under-reported in Malaysia, the problems of SCID are further complicated by the absence of a national patient registry for PID. Thus, the true incidence and disease burden of SCID remained unknown in Malaysia. The main gap in Malaysia is the limited immunodiagnostic capabilities for diagnosing SCID and remains a challenge.

Diagnostics are essential and central to provide high-quality healthcare. Diagnostic tests of SCID are crucial since early diagnosis of SCID resulting in prompt treatment and improves survival rates. Diagnostic tests through genetic testing can be done to confirm SCID through full gene sequencing. Given the rising prevalence of SCID, there is an essential need to the develop a multi-locus sequence based on genetic diagnostic testing of SCID in Malaysia that will help in the identification of the type of SCID based on the mutation that is present. The identification of other family members who may be at risk for the disease can also be done since genetic testing may reveal more information about the carrier status or other affected family member.

Consequently, increased diagnostic capability and prompt treatment are possible to reduce SCID-related morbidity and mortality. In addition, early diagnosis of SCID in affected people requires consistent efforts to raise awareness among healthcare providers.

1.3 Significance of study

The findings from the characterisation of the genotypes of SCID patients in Malaysia with a multi-locus sequence-based strategy involving the defect genes associated with SCID may provide fundamental knowledge and new insight into the establishment of genetic testing for SCID patients in Malaysia. A better understanding of the SCID genotype profiles for patients in Malaysia with recurrent infection through establishing screening protocols based on genetic diagnostic strategy will provide new knowledge and contribute towards better diagnostic and clinical management guidelines for SCID patients in Malaysia.

1.4 Research question

The study was established based on the following question:

What is the characterisation of the genetic defect in SCID patients from Malaysia?

1.5 Objectives

1.5.1 General objective

The purpose is to establish a multi-locus sequence-based genetic diagnostic (SCID Sanger gene panels) strategy for diagnosing patients investigated for SCID.

1.5.2 Specific objectives

1) To establish a qPCR-based assay of SCID gene panel based on the general classification of SCID phenotypes of patients investigated for PID involving 9 genes: *RAG1, RAG2, DCLRE1C, LIG4, ADA, IL2RG, JAK3, IL7R*α and *PNP* gene.
2) To analyse the clinical phenotype, variants, and functional assay of family with PNP- SCID.
3) To establish a Sanger sequencing-based strategy for patients investigated for SCID by validating the qPCR-based assay of SCID gene panel through the testing of mutations causing genes of PNP-SCID.

1.6 Scope and limitation

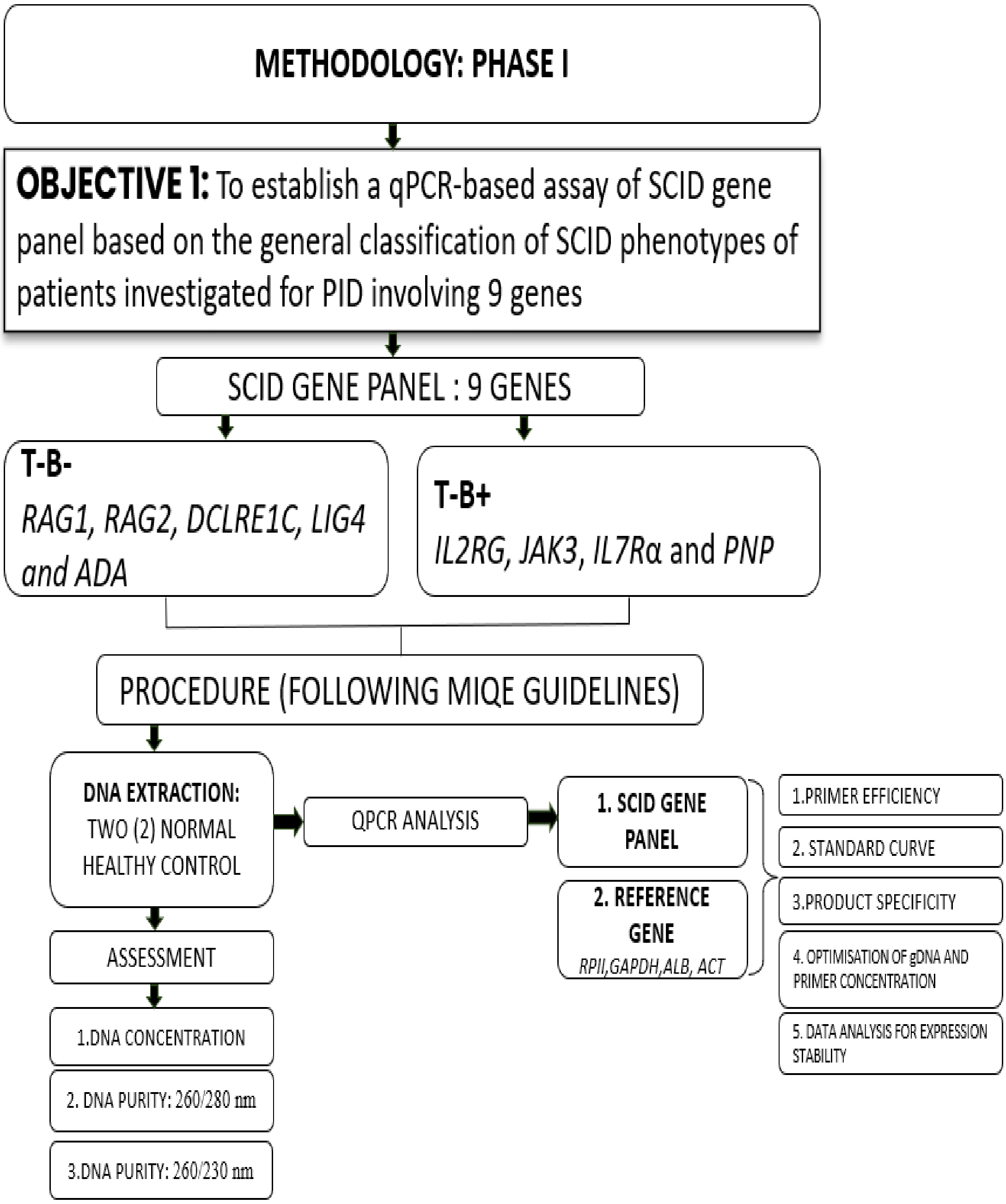
The study of SCID is rarely reported with one out of every 58,000 live births is affected. At the time of undertaking the current study, there was only a case of PNP-SCID female patient being referred for molecular analysis from the public hospital. The key idea in this work is to use a positive SCID case to validate the methodology used. Therefore, *PNP* gene was chosen for further Sanger sequencing validation of the method development.

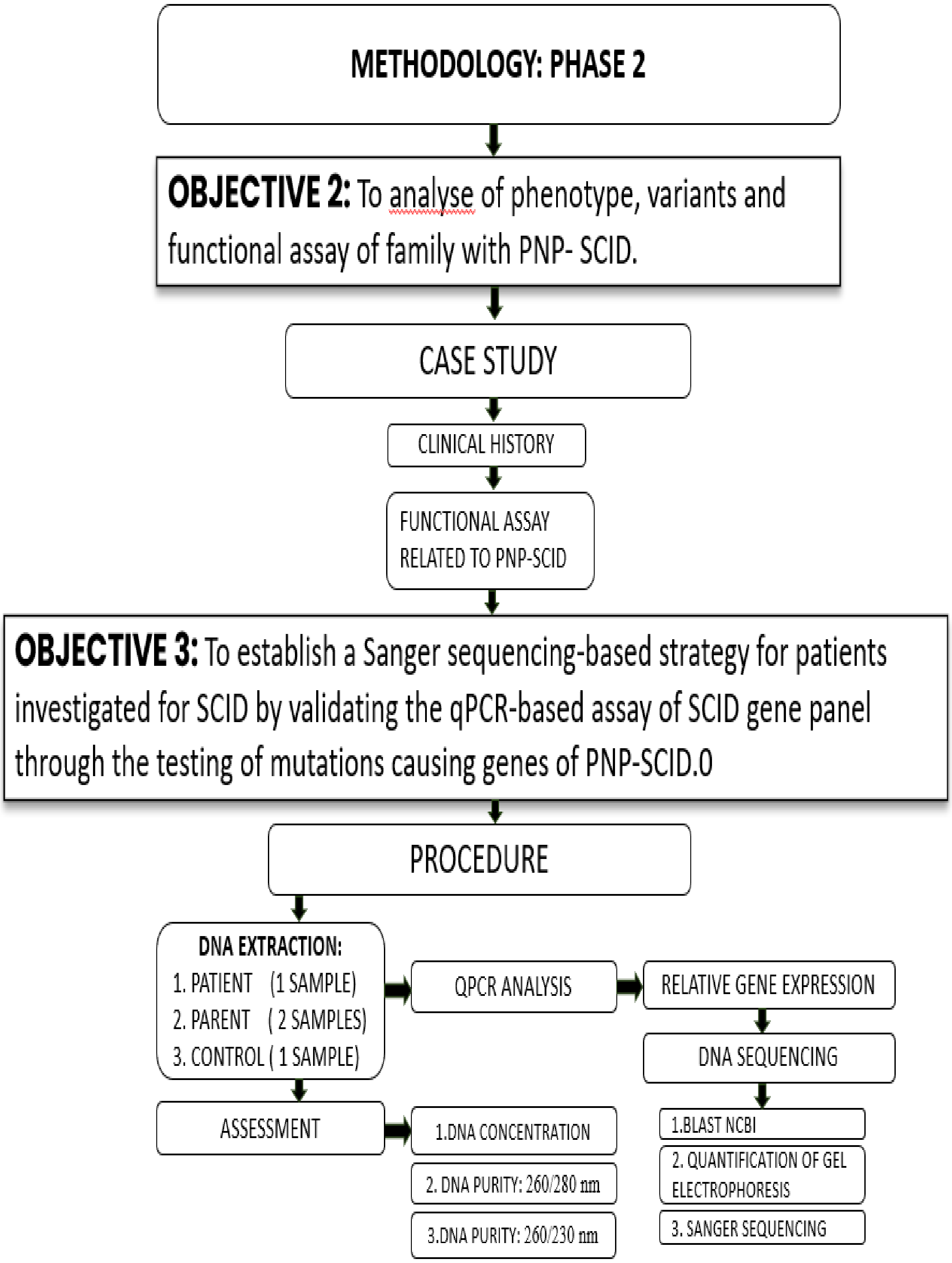
In addition, on March 18th, 2020, the Malaysian government enacted Movement manage Order (MCO) to flatten the curve and manage the pandemic of the COVID-19. During the MCO, a nationwide prohibition on public gatherings and movements was imposed on the entire nation and following a few series of MCO. This situation presents an even greater challenge to the research execution in relation to budget, logistic issue, the rotation scheduled in the laboratory and the kit expiry date. The problem worsened with time as multiple experiment trials need to be done to determine the conditions of the kit used. These issues compounded with delays in getting the new DNA kit extraction due to the defect of the previous kit. The blood specimen collection expired, necessitating the need to repurchase it, which consumed a significant amount of time. In addition, the patient samples were transported from Kelantan to AMDI, USM and in order to keep the samples fresh, we worked till late at night.

1.7 Thesis outline

There are five parts to this thesis. An introduction to PID and SCID is provided in Chapter 1. This work aimed at establishing a multi-locus sequence-based genetic diagnostic testing for SCID. Chapter 2 focuses on SCID immune responses involving innate and adaptive immunity. The classification of phenotypic characteristics of SCID by known genetic abnormalities is also presented. A basic discussion of normal T and B lymphocyte growth and function is explained before describing the variety forms of SCID. Moreover, Chapter 2 discusses genomic diagnostic approaches used in identifying SCID and the importance of identifying the molecular defect for the patients and family members. This research was divided into Phase I (Chapter 3) and Phase II (Chapter 4). Chapter 3 presents the comprehensive development of SCID gene panel testing. Chapter 4 highlights the validated gene amplification of the PNP-SCID gene panel established from the findings of Chapter 3 with analysis of phenotype, variants, and functional assays of family with PNP SCID mutation. Chapter 5 discusses the summary of this research, followed by the bibliography and appendix.

1.8 Flow chart of the study





OBJECTIVE 2: To analyse of phenotype, variants and functional assay of family with PNP- SCID.

OBJECTIVE 3: To establish a Sanger sequencing-based strategy for patients investigated for SCID by validating the qPCR-based assay of SCID gene panel through the testing of mutations causing genes of PNP-SCID.0

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of primary immunodeficiency diseases (PID)

PIDs include over 485 potentially serious disorders, that result in the specific impairment of normal immune development and function which can lead to immune dysfunction with autoimmunity, autoinflammation, allergy, or predisposition to malignancy (Tangye *et al.*, 2022). PID may be diagnosed as an immune cell abnormality, such as suppression of immune cell growth in patients (Chen *et al.*, 2020). PID can affect a single immune system component or multiple cells and proteins.

It is important to know about the characteristic and development of immunological responses, i.e., the innate and adaptive immunity, for better understanding of the various types of PID. The first line of defence against pathogens is innate immunity, while adaptive immunity develops after innate immune responses. Together, these mechanisms effectively protect vertebrates from the numerous diseases that pose a threat to their survival (Rosenzweig and Holland, 2011). Most PID are monogenic disorders which follow the rule of Mendelian inheritance and are inherited in the form of autosomal recessive (AR; (Ochs and Hagin, 2014; Erman *et al.*, 2017). Many defective genes with AR inheritance that underlie PID were first described in patients from the Middle East, with a high parental consanguinity rate (Abolhassani *et al.*, 2018). Therefore, the molecular research of PID is particularly relevant in consanguineous populations (Barbouche *et al.*, 2017).

Although PID varies from secondary immunodeficiencies caused by other reasons such as bacterial or viral infections, malnutrition, immunoglobulin (Ig) loss, malignancy, or the use of immune-suppressing medications (Duraisingham *et al.*, 2014a; Duraisingham *et al.*, 2014b; Srivastava and Wood, 2016), it might result in a developmentally compromised immune system in a similar fashion. Most PID is attributed to the loss of function (LOF) alleles inherited in an AR fashion. However, several have been implicated in autosomal dominant (AD) gain of function (GOF) alleles (Boisson *et al.*, 2015).

Globally, the incidence of PID is predicted to range from 1 in 10,000 to 1 in 100,000 as the diseases vary greatly from country to country (Notarangelo, 2010; van Zelm *et al.*, 2019; Bousfiha *et al.*, 2020; Meyts *et al.*, 2020; Tangye *et al.*, 2020). In Malaysia, the prevalence of PID was 0.37 per 100,000 population from 1979 to 2019 (Abd Hamid *et al.*, 2020). Despite the fact that the prevalence of PID is low in general and is classified as a rare disease, its global incidence is more prevalent than generally thought (Abd Hamid *et al.*, 2020). According to a recent study, PID affects at least 1%-2% of the world's population (Modell *et al.*, 2018).

The clinical manifestation of PID is highly variable, i.e., from the first few days of infancy until adulthood. Although PIDs have a wide range of clinical presentations, many conditions have a high susceptibility to infection. (McCusker *et al.*, 2018) therefore predisposing them to severe opportunistic infections (Al Sukaiti *et al.*, 2021). In primary care, many PIDs may go unnoticed since the disease may manifest as regular infections of the sinuses, ears, and lungs at any age, and accurate and prompt identification of these conditions necessitates particular testing as well as a high level of suspicion (McCusker *et al.*, 2018). The need for a specialised consultation with an expert clinical immunologist in evaluating and managing immunodeficiencies is crucial because early detection and treatment are critical for reducing disease-related morbidity and improving patient outcomes (Shehata *et al.*, 2010; Bonilla *et al.*, 2015).

Since early diagnosis has been demonstrated to show a positive impact on outcomes for many PIDs, the Jeffrey Modell Foundation (JMF) has developed a list of ten warning indicators to assist physicians in its early recognition, as shown in Table 2.1. The JMF warning signs are important for early diagnosis of PID to increase physicians' awareness of PID (Eldeniz *et al.*, 2022), with the overall goal of reducing associated morbidity and mortality (Modell *et al.*, 2018). For children admitted with repeated infections, a complete history should be collected, and the ten warning symptoms of PID defined by the JMF should be reviewed in addition to a full physical examination (Eldeniz *et al.*, 2022). In addition to the ten signs, the following criteria may be considered warning indicators for early diagnosis of PID: chronic diarrhoea, parental consanguinity, and a history of tuberculosis in the family (Eldeniz *et al.*, 2022).

Table 2.1 10 warning signs of immune deficiency in paediatrics suggested by JMF.

Source: (JMF-Primary immunodeficiency resource centre)

	10 Warning Signs of Primary Immunodeficiency
1.	Four or more new ear infections within 1 year.
2.	Two or more serious sinus infections within 1 year.
3.	Two or more months on antibiotics with little effect.
4.	Two or more pneumonias within 1 year.
5.	Failure of an infant to gain weight or grow normally.
6.	Recurrent, deep skin or organ abscesses.
7.	Persistent thrush in mouth or fungal infection on skin.
8.	Need for intravenous antibiotics to clear infections.
9.	Two or more deep-seated infections including septicaemia.
10.	A family history of primary immunodeficiency.

There has been a rise in PID-related molecular defects every year, and the advances in genetic discoveries of human immune diseases continue to remarkably deepen the understanding of cellular, molecular, and immune mechanisms of the pathogenesis of the disease (Tangye *et al.,* 2022). Therefore, the classifications of PID are revised and updated with an extensive list of recognised immunodeficiencies by the International Union of Immunological Societies (IUIS) expert committee (Tangye *et al.,* 2020). The committee provides a document to the research and clinical communities involving a brief description of the immunologic and clinical features, the genetic abnormalities, and the type of inheritance to facilitate genetic diagnoses of IEI by guiding the design of panels used for targeted gene sequencing (Tangye *et al.,* 2022).

In the latest updates of the IUIS in 2022, 10 groups are used to categorise PIDs: immunodeficiencies affecting cellular and humoral immunity (Table 1: 3 sub-tables), combined immunodeficiencies with associated or syndromic features (Table 2: 9 sub-tables), predominantly antibody deficiencies (Table 3: 4 sub-tables); diseases of immune dysregulation (Table 4: 7 sub-tables), congenital defects of phagocyte number or function (Table 5: 4 sub-tables) defects in intrinsic and innate immunity (Table 6: 9 sub-tables), autoinflammatory disorders (Table 7: 3 sub-tables), complement deficiencies (Table 8), bone marrow failure (Table 9), and phenocopies of inborn errors of immunity (Table 10; (Tangye *et al.*, 2022), as shown in Table 2.2.

An understanding of PID is vital for advancing medical knowledge and therapies to provide patients with a better quality of life (Meyts *et al.*, 2020). According to the IUIS updates in 2022, 485 genetic defects were identified as causing IEI (Tangye *et al.*, 2022). Moreover, similar clinical symptoms can result from mutations in several genes, and vice versa, various mutations in the same gene might result in various disease phenotypes (Erman *et al.*, 2017), complicating both diagnosis and disease classification.

Thus, the IUIS classification serves as an important reference for researchers and immunologists and researchers globally concerning the recent genetic cause of immune deficiency (Picard *et al.*, 2018) and the classification is expected to undergo major revision and update.

Table 2.2 The International Union of Immunological Societies (IUIS) Classification

Table	Subtable	Description
1		**Immunodeficiencies affecting cellular and humoral immunity**
	1	T-B+(SCID)
	2	T-B- (SCID)
	3	Combined Immunodeficiency (CID), Generally Less Profound than SCID
2		**Combined immunodeficiencies with associated or syndromic features**
	1	Immunodeficiency with Congenital Thrombocytopenia
	2	DNA Repair Defects Other Than Those Listed in Table 1
	3	Thymic Defects with Additional Congenital Anomalies
	4	Immuno-osseous Dysplasias
	5	Hyper IgE Syndromes (HIES)
	6	Defects of Vitamin B12 and Folate Metabolism
	7	Anhidrotic Ectodermodysplasia with Immunodeficiency (EDA-ID)
	8	Calcium Channel Defects
	9	Other Defects
3		**Predominantly antibody deficiencies**
	1	Severe Reduction in All Serum Immunoglobulin Isotypes with Profoundly Decreased or Absent B Cells, Agammaglobulinemia
	2	Severe Reduction in at Least 2 Serum Immunoglobulin Isotypes with Normal or Low Number of B Cells, CVID Phenotype
	3	Severe Reduction in Serum IgG and IgA with Normal/Elevated IgM and Normal Numbers of B cells, Hyper IgM
	4	Isotype, Light Chain, or Functional Deficiencies with Generally Normal Numbers of B Cells
4		**Diseases of immune dysregulation**
	1	Familial Hemophagocytic Lymphohistiocytosis (FHL syndromes)
	2	FHL Syndromes with Hypopigmentation
	3	Regulatory T Cell Defects
	4	Autoimmunity with or without Lymphoproliferation
	5	Immune Dysregulation with Colitis
	6	Autoimmune Lymphoproliferative Syndrome (ALPS, Canale-Smith syndrome)
	7	Susceptibility to EBV and Lymphoproliferative Conditions
5		**Congenital defects of phagocyte number or function**
	1	Congenital Neutropenias
	2	Defects of Motility
	3	Defects of Respiratory Burst
	4	Other Non-Lymphoid Defects

6		**Defects in intrinsic and innate immunity**
	1	Mendelian Susceptibility to mycobacterial disease (MSMD)
	2	Epidermodysplasia verruciformis (HPV)
	3	Predisposition to Severe Viral Infection
	4	Herpes Simplex Encephalitis (HSE)
	5	Predisposition to INVASIVE Fungal Diseases
	6	Predisposition to Mucocutaneous Candidiasis
	7	TLR Signalling Pathway Deficiency with Bacterial Susceptibility
	8	Other Inborn Errors of Immunity Related to Non-Hematopoietic Tissues
	9	Other Inborn Errors of Immunity Related to Leukocytes
7		**Autoinflammatory disorders**
	1	Type 1 Interferonopathies
	2	Defects Affecting the Inflammasome
	3	Non-Inflammasome Related Conditions
8		**Complement deficiencies**
9		**Bone marrow failure**
10		**Phenocopies of inborn errors of immunity**

2.2 Overview of severe combined immunodeficiency disease (SCID)

SCID is the most severe form of PID. It is a heterogeneous group of diseases characterised by hampered T lymphocytes with varying B lymphocytes abnormalities and natural killer (NK) cells (Al Sukaiti *et al.*, 2021), resulting in reduced lymphocyte functioning and significant immune system defects (Gaspar *et al.*, 2013; Shearer *et al.*, 2014; Bonilla *et al.*, 2015). Affected kids appear normal at birth, with the first symptom usually being a potentially fatal opportunistic infection such as *Pneumocystis jirovecii* pneumonia (PJP); (Bateman *et al.*, 2021). In the absence of early diagnosis, infection causes death in these infants, which often occurs during the first two years of life and receives allogeneic HSCT, enabling restoration of a functioning immune system by HSCT (Bateman *et al.*, 2021). Allogeneic HSCT with human leukocyte antigen (HLA)-matched donors have been used effectively for more than 60 years to treat various malignant and non-malignant diseases (Demirci *et al.*, 2020).

After the implementation of newborn screening (NBS), it has been found that in the U.S., SCID occurs at a rate of 1 in 58,000 live births, with 52 cases of classic SCID, leaky SCID, and Omenn syndrome, showing that SCID may not be as uncommon as previously thought (Kwan *et al.*, 2014). SCID incidence and prevalence varies by region (Al Sukaiti *et al.*, 2021) and are reported to be greater in nations with a high consanguinity rate (Aluri *et al.*, 2019). A study in Oman, with a high rate of consanguineous marriage, with approximately half (49%) of all marriages being between blood relatives (Mazharul Islam, 2017), reported an annual SCID incidence rate of 4.5 per 100,000 Omani live births (Al Sukaiti *et al.*, 2021). Meanwhile, the overall incidence of SCID at birth, particularly among the Irish traveller population, is 1/39,248 live births, with a high incidence of ADA SCID (Burns *et al.*, 2021).

Despite the fact that SCID is a particularly severe variant of PID, an early diagnosis can be life-saving (Erman *et al.*, 2017) since early discovery of SCID has improved the outcomes of persons suffering from the disease (Habib Dzulkarnain *et al.*, 2021). These symptoms must be recognised by healthcare providers who treat people with PID to allow an early diagnosis and provide prompt treatment. As the PID phenotype determination is essential for genetic counselling and therapy, especially HSCT and gene therapy for SCID patients, the next gene sequencing (NGS) method will help identify disease phenotypes rapidly (Erman *et al.*, 2017). An investigation of SCID patients using targeted sequencing was done in Turkey, involving 356 PID-related genes and eight disease-causing mutations were identified in 27 SCID patients (Erman *et al.*, 2017) with RAG1 was identified as their commonest SCID gene mutation.

The outcome among infants who received transplants prior to the age of 3.5 months is optimal (94% survival rate), regardless of donor type (Pai *et al.*, 2014). The survival rate for older infants without prior infection is about 90%, or with an infection that had cleared up by the time the transplant was performed is 82% (Pai *et al.*, 2014). In addition, the presence of active viral infections such as cytomegalovirus (CMV) transmission has an impact on the prognosis of patients undergoing HSCT (del Rosal *et al.*, 2022). Reduced time between birth and intervention may prevent CMV transmission (Strand *et al.*, 2020).

According to the general agreement reached by the Primary Immune Deficiency Treatment Consortium (PIDTC) 2014, SCID is further outlined as typical SCID and leaky SCID (Shearer et al., 2014). The failure to produce naive T cells is the underlying problem in SCID, resulting in severe T cell lymphopenia (Bateman *et al.*, 2021). A typical SCID is determined by a lack of or an extremely low concentrations of T cells

(CD3+ T cells <300/uL) and the absence or extremely decreased T cell function (<10% lower range of normal), as determined by proliferation in response to phytohaemagglutinin (PHA) or demonstrable transplacental maternal engraftment (TME) (Richards *et al.*, 2020). Meanwhile, an atypical or leaky SCID is characterised by a low number of CD3+ T cells, i.e., 30% less than the lower limit of normal T cell function, as determined by proliferation with PHA; reduced or absent naive T cells; and the absence of transplacental maternal engraftment (TME) (Richards *et al.*, 2020). This latter type is a less serious type of SCID, which happens when a patient exhibits classic SCID symptoms yet has a low T-cell count, which can be diagnosed by NBS (Crestani *et al.*, 2015). In most cases, the T cell proliferation assay reveals a poor T cell response to PHA (less than 30% of the lower limit of normal T cell activity).

However, a revised version of PIDTC Criteria and Definitions for SCID 2022 updated the criteria for SCID patients, focusing on distinguishing patients with various types of SCID subtypes (Dvorak *et al.*, 2023). A typical SCID is shown by very low T cells ($<0.05 \times 10^9$/L) without TME. For the presence of TME, no alternate explanation for a reduced T-cell count and either undetectable or reduced TRECs or <20% of CD4+ T cells have naive cell surface markers, whereas leaky or atypical SCID is shown by two or more of the following characteristics: reduced T-cell number for age (0.05–1.0 $\times 10^9$/L), oligoclonal T cells or abnormal TRECs or <20% of CD4+ T cells are naive together with low response (<50%) to PHA, anti-CD3, or anti-CD3/CD28. The 2022 definitions classified more patients with pathogenic mutations, which described SCID and its subtypes more precisely and facilitated the SCID characteristics and outcomes.

The implementation of a SCID NBS programme in many countries has been made possible by the invention of a reliable test to detect TREC as a marker of naive T cell production, mostly with advanced healthcare systems (Bateman *et al.*, 2021).

Immunoglobulin levels, lymphocyte subsets, mitogen-induced lymphocyte proliferation, and genetic testing are all necessary to definitively diagnose positive TRECs NBS (Bateman *et al.,* 2021). Within the first three and a half months of a child's life, early intervention through NBS and transplantation has better effects on the outcomes of the disease (Pai *et al.*, 2014).

The NBS programmes have shed light on the prevalence of SCID, which was previously misunderstood to be less common than it actually is (Al Sukaiti *et al.*, 2021). SCID NBS has shown a prevalence of 1 in 131,485 live births in Taiwan (Chien *et al.*, 2017), 1 in 58,000 live births in the USA (Kwan *et al.*, 2014), and 1 in 11,821 live births in China (Chien *et al.*, 2015). However, it has also been investigated that the SCID rates are higher in countries where marriages between family members are common, such as Saudi Arabia, with 1 in 2,906 live births (Al-Mousa *et al.*, 2018), exhibiting a higher incidence than the reported cases from the USA NBS programmes.

2.3 Normal haematopoiesis

The immune system of the human body has evolved a multi-layered defence mechanism that is continually interacts with microorganisms inhabiting the human body to maintain a balance between the quantity and type of non-pathogenic microorganisms and to combat or eliminate potentially pathogenic microorganisms. The immune system is a tightly regulated network that can maintain a balance of immune homeostasis under normal physiological conditions (Crimeen-Irwin *et al.*, 2005). It can be classified into two subsystems: the innate and adaptive immunity. The initiation of the host defence is done by the innate immune system. Innate immunity is often regarded as a nonspecific response, whereas the adaptive immune system is very specific (Vivier and Malissen, 2005).

PID is broadly classified as disorders of innate (e.g., complement and phagocyte disorders) or adaptive (i.e., T-cell, B-cell or combined immunodeficiencies) immunity (McCusker and Warrington, 2011), as indicated in Figure 2.1. Innate immune cell populations such as myeloid cells, NK cells, and innate lymphoid cells mediate innate immunity and ancient humoral systems like defensins and complements (Netea *et al.*, 2019). The adaptive immune system, also known as the acquired immune system, initiates specific immunity against the pathogen through antigen presentation. Adaptive immunity contains specialised immune cells and antibodies that target and eliminate foreign invaders. This adaptive immunity can last a few weeks or endure a long period and are divided into humoral and cell-mediated immunities. Initially, the innate and adaptive immune processes were viewed as temporally compartmentalised responses; however, in the last two decades, research has revealed strong connections and an efficient network between them (Netea *et al.*, 2019).

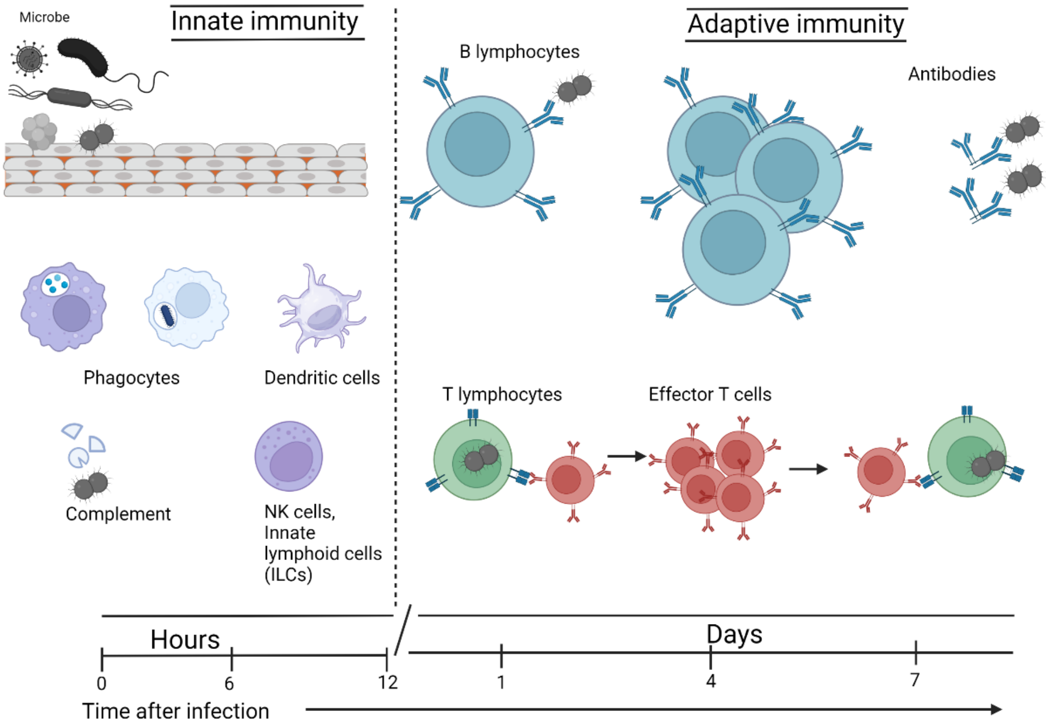


Figure 2.1 Innate and adaptive immune system.

Created with BioRender.com. Source: (Abbas *et al.*, 2014).

In the 1950s, scientists found that the bone marrow has stem cells called hematopoietic stem cells (HSCs). HSCs can self-renew and turn into many different types of cells (Ambreen and Deepa, 2012), as shown in Figure 2.2. HSCs are immature adult stem cells that are multipotent and exhibit long-term engraftment. They are responsible for the development of all mature blood cell lineages and potentially rebuild the complete hematopoietic system in a conditioned recipient after infusion (Demirci *et al.*, 2020). The blood elements include many types of cells or cell fragments, each with its own morphology and function. These elements are produced through haematopoiesis during which HSCs proliferate and undergo self-renewal or differentiation into lineage-committed progenitors, which continue to differentiate into mature blood cells. Self-renewal and differentiation are two essential characteristics of HSCs required for normal haemopoiesis.

The HSC differentiate into two types of progenitors, namely common myeloid progenitors (CMP) and common lymphoid progenitors (CLP), which then differentiate into numerous blood cells, such as platelets, granulocytes, lymphocytes, and macrophages (Ambreen and Deepa, 2012). Myeloid progenitor cells continue to differentiate into erythrocytes, platelets (via megakaryocyte fragmentation), neutrophils, monocytes, basophils, and eosinophils. Monocytes mature into macrophages, which are tissue-resident or infiltrating immune cells important for innate immunity, proper tissue development, homeostasis, and tissue repair (Watanabe *et al.*, 2019). Neutrophils are essential effector cells in the innate immune system that constantly monitor the body for symptoms of microbial infection. Upon encountering microorganism invasion, these cells act fast to trap and eliminate invading germs (Rosales, 2018). Eosinophils are involved in host defence against parasites and promoting allergic reactions (Wen and Rothenberg, 2016). Erythrocytes are the integral