GENOTYPE AND ALLELE FREQUENCY OF HUMAN NEUTROPHIL ANTIGENS IN ACHEH, KEDAH, MANDAILING, MINANGKABAU AND PATTANI MALAYS

NUR HAFIZA BINTI MD YUSOP

UNIVERSITI SAINS MALAYSIA

2020

GENOTYPE AND ALLELE FREQUENCY OF HUMAN NEUTROPHIL ANTIGENS IN ACHEH, KEDAH, MANDAILING, MINANGKABAU AND PATTANI MALAYS

By

NUR HAFIZA BINTI MD YUSOP

Thesis submitted in fulfilment of the requirements

for the degree of

Master of Science

November 2020

ACKNOWLEDGEMENTS

In the name of Allah, the Most Beneficent, the Most Merciful.

First and foremost, I am grateful to Allah for giving me the strength and excellent health to complete this thesis for the requirements for the degree of Master of Science.

I wish to take this opportunity to express my deep appreciation to my supervisor, Dr Edinur Hisham Atan for his expert guidance, understanding and encouragement during my research and study. I would also like to thank Associate Professor Dr Zafarina Zainuddin and Mr S. Panneerchelvam (Co Supervisors) for their valuable guidance and encouragement extended to me.

My sincere thanks also go to the Director General of Department of Chemistry Malaysia for the permission to pursue my research. Special gratitude also to the Vice Chancellor, Director of Institute of Graduate Studies, Dean of School of Health Sciences and the staff of Universiti Sains Malaysia for their support, help and cooperation throughout this academic journey. Many thanks go to my numerous friends and colleagues who helped and gave me their full cooperation in completing this thesis.

Finally, I would like to thank my husband, parents and family for their unconditional love and support during my many days of absence in completing this study.

TABLE OF CONTENT

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMNS

%	: Percent
^{o}C	: Degree celsius
<	: Less than
X	: Times
µg	: Microgram
µl	: Microliter
µM	: Micromolar
χ^2	: Chi-squared test
AE	: Elution buffer
a.k.a	: Also known as
AL	: Lysis buffer
Arg	: Arginine
AW 1	: Washing buffer 1
AW 2	: Washing buffer 2
bp	: Base pair
Chr	: Chromosome
CD	: Cluster of differentiation
dH_2O	: Distilled water
dNTPs	: Deoxynucleoside triphosphates
DNA	: Deoxyribonucleic acid
e.g.	: For example
g	: Gram
GAT	: Granulocyte agglutination test
GIFT	: Granulocyte immunofluorescence test
Gln	: Glutamine
GPI	: Glycosyl-phosphatidylinositol
HNA	: Human neutrophil antigen
HREC	: Human Research Ethical Committee
HWE	: Hardy-Weinberg equilibrium
i.e.	: In other words
ISBT	: International Society of Blood Transfusion

kDa	: Kilodalton
km^2	: Square kilometre
LFA-1	: Leukocyte function antigen-1
MAIGA	: Monoclonal antibody-specific immobilization of granulocyte antigen
ml	: Millilitre
mg	: Milligram
n	: Number of subjects
ng	: Nanogram
n.a	: Not applicable
NA_2EDTA	: Ethylenediaminetetraacetic acid disodium salt
NAN	: Neonatal alloimmune neutropenia
P	: Significant value
PCR	: Polymerase chain reaction
PCO	: Principle coordinate
pH	: Power of hydrogen
Phe	: Phenylalanine
PNH	: Paroxysmal nocturnal haemoglobinuria
psi	: Pound-force per square inch
RFLP	: Restriction fragment length polymorphism
SBT	: Sequence-based typing
SNP	: Single nucleotide polymorphism
SSP	: Sequence specific primer
TBE	: Tris borate ethylenediaminetetraacetic
Thr	: Threonine
TRALI	: Transfusion-related acute lung injury
UK	: United Kingdom
USA	: United States of America
UV	: Ultraviolet
V	: Volts
vs	: Versus

GENOTIP DAN FREKUENSI ALEL ANTIGEN NEUTROFIL MANUSIA BAGI KAUM MELAYU ACHEH, KEDAH, MANDAILING, MINANGKABAU DAN PATTANI

ABSTRAK

Pengetahuan berkaitan alel antigen neutrofil manusia (HNA) memainkan peranan penting dalam kajian penyakit memandangkan HNA telah dilaporkan terlibat dalam patogenesis keadaan klinikal seperti *alloimmune neonatal neutropenia* (NAN) dan kecederaan paru-paru akut yang berkaitan transfusi (TRALI). Kajian ini melibatkan kaedah tindak balas rantai polimerase dengan jujukan primer tertentu bagi pemprofilan lokus genetik HNA-1, -3, -4 dan -5 ke atas 194 sampel DNA genomik dengan berat molekul tinggi yang diekstrak daripada spesimen darah daripada kaum Melayu Acheh (n = 35), Kedah (n = 30), Mandailing (n = 47), Minangkabau (n = 47) dan Pattani (n = 35) dengan sasaran bagi mendapatkan data untuk kajian dalam populasi serta risiko ketidaksesuaian HNA antara individu. Secara umum, set data yang diperoleh dalam kajian ini menunjukkan HNA-1a/1a, -3a/3a, -4a/4a dan -5a/5a sebagai genotip yang paling kerap diperhatikan dalam kumpulan sub-etnik Melayu. Walau bagaimanapun, HNA-1a/1b dicatatkan sebagai genotip yang paling dominan untuk sistem HNA-1 bagi Melayu Kedah, Mandailing dan Pattani manakala HNA-5a/5b dicatat sebagai genotip unggul bagi Melayu Acheh dan Pattani untuk sistem HNA-5. Data HNA bagi kumpulan populasi penduduk yang diperoleh dalam kajian ini adalah serupa dengan kajian terdahulu yang melibatkan kumpulan sub-etnik Melayu (Kelantan, Banjar, Jawa, Banjar dan Bugis Melayu). Walau bagaimanapun, terdapat perbezaan yang ketara antara kumpulan data HNA

bagi sub-etnik Melayu dengan kumpulan data HNA yang dilaporkan untuk Orang Asli Semang (Bateq, Kensiu dan Lanoh) dan Senoi (Che Wong dan Semai). Pemerhatian ini menggambarkan terdapatnya kumpulan genetik yang unik bagi populasi yang berbeza di Semenanjung Malaysia, yang sebahagian besarnya dikaitkan dengan perbezaan asal-usul mereka. Analisis statistik menunjukkan potensi risiko imunisasi bagi transfusi dan kehamilan berkaitan HNA adalah rendah untuk HNA-1c, -3a, -4a dan -4b dan penilaian risiko ini melibatkan pasangan ibu/bapa dan penderma/pesakit daripada kumpulan populasi yang sama atau berbeza. Walau bagaimanapun, HNA lain (contohnya HNA-1a, -1b, -5a dan -5b) menunjukkan risiko imunisasi yang tinggi walaupun terdiri daripada pasangan atau penderma dan penerima daripada latar belakang etnik yang sama. Kesimpulannya, kajian ini telah berjaya melaksanakan pemprofilan genotip bagi empat sistem HNA dari lima kumpulan sub-etnik Melayu dan kumpulan data HNA ini menyediakan sumber maklumat yang berharga untuk mengkaji sejarah penduduk dan kesihatan di Semenanjung Malaysia.

GENOTYPE AND ALLELE FREQUENCY OF HUMAN NEUTROPHIL ANTIGENS IN ACHEH, KEDAH, MANDAILING, MINANGKABAU AND PATTANI MALAYS

ABSTRACT

Human neutrophil antigens (HNAs) have been reported to involve in pathogenesis of clinical conditions such as neonatal alloimmune neutropenia (NAN) and transfusion-related acute lung injury (TRALI), thus knowledge on local distribution of HNA alleles play significant roles in disease studies. The present study involves polymerase chain reaction (PCR) using sequence specific primer (SSP) typing of HNA-1, -3, -4 and -5 loci in 194 high molecular weight genomic DNA samples extracted from blood samples of Acheh (n=35), Kedah (n=30), Mandailing (n=47), Minangkabau (n=47) and Pattani (n=35) Malay individuals with the target of obtaining data for population review along with HNA incompatibility risk investigation. In general, datasets obtained in this study shows HNA-1a/1a, -3a/3a, -4a/4a and -5a/5a were the most frequent genotypes observed in the Malay sub-ethnic groups. However, HNA-1a/1b was recorded as the most frequent genotype for HNA-1 system in Kedah, Mandailing and Pattani Malays while HNA-5a/5b was logged as the most common HNA-5 genotype in Acheh and Pattani Malays. The HNA population datasets collected in the present survey are similar to those from the earlier study on Malay sub-ethnic groups (Kelantan, Banjar, Jawa, Banjar and Bugis Malays), yet, significant differences were observed between Malay sub-ethnic groups and HNA datasets reported for Orang Asli Semang (Bateq, Kensiu and Lanoh) and Senoi (Che Wong and Semai). This observation reflects the unique

gene pools of different population groups in Peninsular Malaysia, which largely associated with their different origins. Statistical analyses on the HNA datasets also showed a low risk of HNA-related transfusion and gestation alloimmunizations for HNA-1c, -3a, -4a and -4b and these risk assessments applied to any pairs of mother/father and donor/patient population groups. However, probability of alloimmunization for other HNAs (e.g. HNA-1a, -1b, -5a and -5b) is high even for couples or donor and recipient of similar ethnic background. In conclusion, the present study has successfully genotyped four HNA systems in five Malay sub-ethnic groups and these HNA datasets provide valuable source of information for studying population history and health in Peninsular Malaysia.

CHAPTER 1

INTRODUCTION

1.1 Background of study

Human neutrophil antigens (HNAs) are immunogenic molecules and have been reported to involve in pathogenesis of several clinical conditions which includes neonatal alloimmune neutropenia (NAN) and transfusion-related acute lung injury (TRALI) (Bayat *et al*., 2012 and Bowens *et al*., 2012). Genes code for HNAs (e.g. HNA-1, HNA-3, HNA-4 and HNA-5 systems) have been characterized and reported for several European and Asian populations and their allele and genotype frequency spectra reflect population history and origins (Hauck *et al*., 2011; Xia *et al*., 2011; Matsuhashi *et al*., 2012; Nielsen *et al*., 2012; Cardoso *et al*., 2013; Changsri *et al*., 2013 and Manaf *et al*., 2015 and 2016).

The present study involves polymerase chain reaction (PCR) using sequence specific primer (SSP) typing of HNA-1, -3, -4 and -5 loci in Acheh, Kedah, Mandailing, Minangkabau and Pattani Malays and is compliment to the previous studies (Manaf *et al*., 2015 and 2016) on various sub-populations in Peninsular Malaysia. Benefits of the present study to the medical genetics and ancestry study are discussed in Chapter 5.

1.2 Problem statements

HNA were recorded to be polymorphic and unique between people of different ethnicities (Xia *et al*., 2017). Thus, differences in HNA allelic distributions between people of different origins in Malaysia may lead to various medical conditions including transfusion and gestation incompatibilities. Previous HNA survey (Manaf *et al*., 2015 and Manaf *et al*., 2016) only reported HNA data for Orang Asli and few Malay subethnic groups. The latter include Banjar, Bugis, Champa, Jawa and Kelantan Malays but not for the five Malay sub-ethnic groups screened in this study (Acheh, Kedah, Mandailing, Minangkabau and Pattani Malays).

1.3 Rational of study

HNA molecules are important determinants in transfusion and gestation and their allelic frequency spectra at population level reflect ancestry. Previous studies of HNA on Malays (Banjar, Bugis, Champa, Jawa and Kelantan) and Orang Asli (Semang, Senoi and Proto-Malays) reported by Manaf *et al.* (2015, 2016) showed highly diverse of HNA distributions in Peninsular Malaysia, especially between unrelated ethnic groups (e.g. Semang vs Malay sub-ethnic groups). These earlier studies also reported a potential risk of transfusion and gestation alloimmunizations between donor and patient or couples of different ethnical backgrounds. Therefore, the present study on the remaining five Malay sub-ethnic groups (Acheh, Kedah, Mandailing, Minangkabau and Pattani) is compliment to those reported by Manaf *et*

al. (2015, 2016) and provide valuable HNA datasets for health analysis and population genetic study in Peninsular Malaysia.

1.4 Objectives of study

1.4.1 General objective

To genotype HNA-1, -3, -4 and -5 loci in five Malay sub-ethnic groups namely Acheh, Kedah, Mandailing, Minangkabau and Pattani Malays using polymerase chain reaction with sequence specific primers.

1.4.2 Specific objective

i. To use HNA-1, -3, -4 and -5 allele frequencies from Acheh, Kedah, Mandailing, Minangkabau and Pattani Malays and other reference HNA datasets to study genetic relationships between various sub-populations in Peninsular Malaysia and Asia Pacific region.

ii. To use HNA datasets of Acheh, Kedah, Mandailing, Minangkabau and Pattani Malays and those reported for other sub-populations in Peninsular Malaysia for estimating risk associated with HNA incompatibility in transfusion and gestation.

CHAPTER 2

LITERATURE REVIEW

2.1 Human neutrophil antigen

Neutrophil is an important cell in immune responses and make-up the largest proportion (40 – 70%) of the white blood cells (Fung and Minchinton, 2011). The neutrophil cells in human body play a crucial role in innate immune response and destroy foreign microorganisms by phagocytosis. Phagocytosis activity of neutrophil on pathogenic cell is shown in Figure 2.1.

The immunogenic and polymorphic glycoproteins on neutrophil cell surface were first discovered in a case of NAN (Lalezari *et al*., 1960). In 1998, a new nomenclature for well-defined glycoproteins on neutrophil membrane has been adopted by Granulocyte Antigen Working Party of the International Society of Blood Transfusion (ISBT) during their meeting in S' Agaró, Spain (Bux, 1999). These glycoproteins were then referred as human neutrophil antigens (HNAs) by the Granulocyte Antigen Working Party (Silliman *et al*., 2003). According to this nomenclature, each HNA system is indicated by integer and specific antigen and assigned alphabetically by the date of publication (Bux, 1999). At present, eleven HNAs were reported and are assigned into five HNA systems (Table 2.1). These five HNA systems are known as HNA-1, HNA-2, HNA-3, HNA-4 and HNA-5.

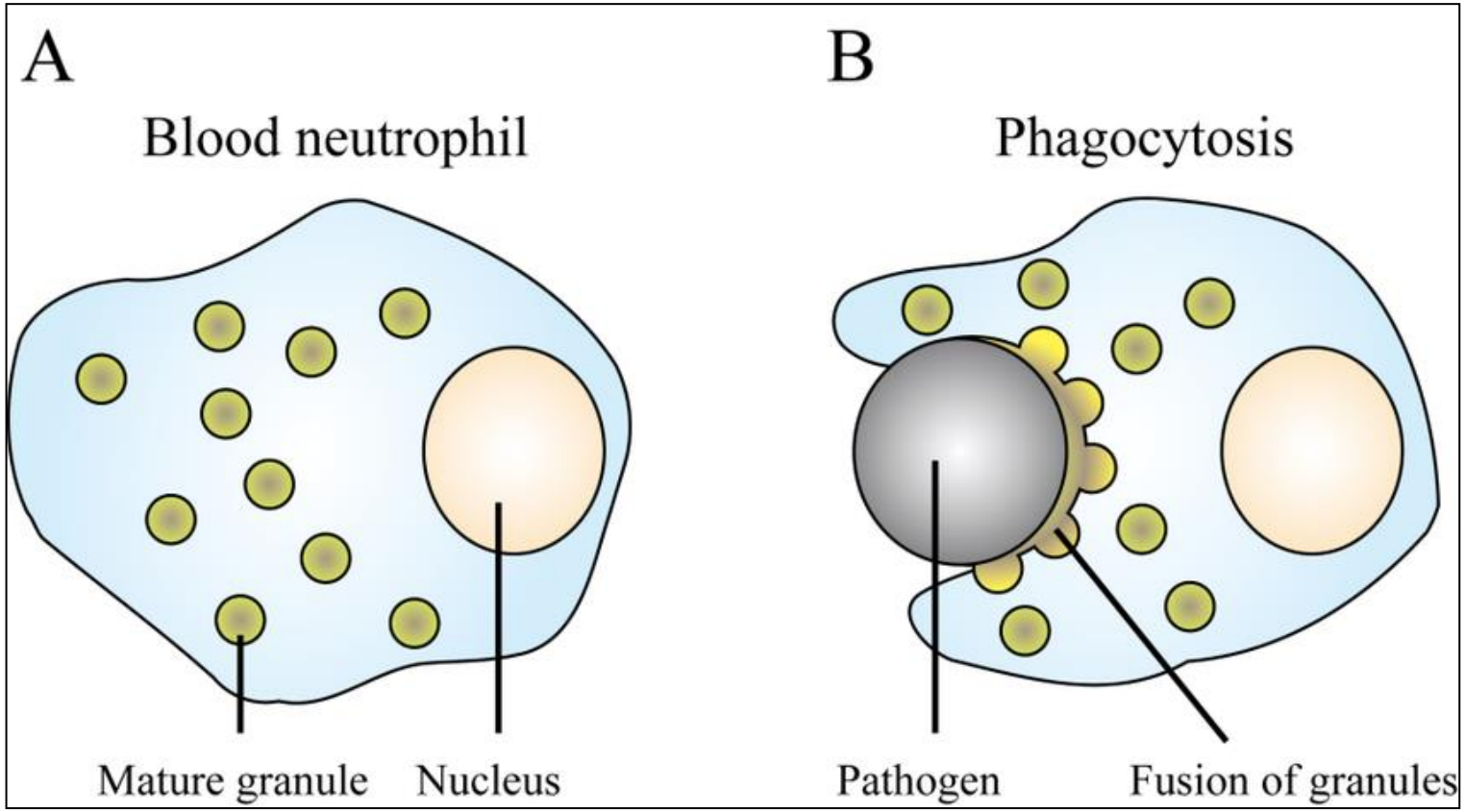

Figure 2.1: Phagocytosis of pathogen by neutrophils. (A) Neutrophil with mature azurophil granules (B) Phagocytosis process of pathogen. When a pathogen being ingested, azurophil granules will fused together around the pathogen which will exert the antimicrobial activity in a limited space. This figure was adapted from Aoki and Euda (2013).

Table 2.1: Specificities of the five HNA systems

System	Former name	Antigen	Chromosome location	Carrier glycoprotein / CD classification	Allele	Expression
HNA-1	NA1 NA2 SH	HNA-1a HNA-1b HNA-1c	Chr. 1	Fcγ Receptor IIIb / CD16b	*FCGR3B*01/04* *FCGR3B*02/05* *FCGR3B*03*	Neutrophils
HNA-2	NB1	HNA-2a	Chr. 19	NB1 glycoprotein / CD177	*CD177*01*	Neutrophils, Monocytes
HNA-3	5b 5a	HNA-3a HNA-3b HNA-3ab	Chr. 19	CTL-2 (Choline transporter-like) / undefined	*SLC44A2*1* *SLC44A2*2* *SLC44A2*1:2*	Neutrophils, granulocytes, monocytes, lymphocytes, platelets, kidney and placental, spleen, lymph node and endothelial cells
HNA-4	MART MART negative	HNA-4a HNA-4b	Chr. 19	MAC1; CR3; $\alpha_M\beta_2$-integrin / CD11b	*ITGAM*01 (230G)*	Granulocytes, monocytes, T-lymphocytes
HNA-5	OND OND negative	HNA-5a HNA-5b	Chr. 16	LFA-1; $\alpha_L\beta_2$-integrin / CD11a	*ITGAL*01 (2372G)*	Granulocytes, monocytes, T-lymphocytes

Data obtained from Bux (2008), Moritz *et al*. (2009), Fung and Minchinton (2011), Reil *et al*. (2011) and Huvard *et al*. (2012).
Abbreviation: CD=cluster of differentiation, Chr=chromosome.

2.1.1 HNA-1 system

The neutrophil antigens within HNA-1 system are HNA-1a, HNA-1b, HNA-1c and HNA-1d. These HNA-1 molecules are located on Fc gamma-receptor IIIb that attached to neutrophil membrane via glycosyl-phosphatidylinositol (GPI) anchor protein (Fung and Minchinton, 2011) and are coded by the five *FCGR3B* allelic variants (*FCGR3B*01, 02, 03, 04* and *05)* – see Table 2.1.

The *FCGR3B*04* and *FCRG3B*05* alleles are only differentiated from the *FCRG3B*01* and *FCRG3B*02* alleles by a single nucleotide substitution (G316A and A244G, respectively) but both were shown to be reactive with HNA-1b and HNA-1a antibodies, respectively. Therefore, products of *FCGR3B*01* and *FCGR3B*04* alleles are assigned as HNA-1a and those by *FCGR3B*02* and *FCGR3B*05* as HNA-1b. However, another single nucleotide substitution on the background of the *FCGR3B*02* allele cause an additional polymorphism to *FCGR3B* gene, the *FCGR3B*03* allele code for HNA-1c molecule (Bux, 2008).

The molecular background of HNA-1d is yet to be identified but individuals with this granulocyte antigen was reported to be reactive with antibodies that recognized HNA-1b epitopes and alloimmunization occurs in individuals with HNA-1c phenotype (Reil *et al*., 2013). In contrast, *FCGR3B* gene deletion resulted the absent of HNA-1 molecule on the surface of neutrophil granulocytes. Individuals with low *FCGR3B* affinity were identified as HNA-1 null phenotype and they may develop alloantibodies once exposed to foreign HNA antigens (Norcia *et al*., 2009).

2.1.2 HNA-2 system

The HNA-2 is glycosyl-phosphatidylinositol linked to CD77 glycoprotein and was first described by Lalezari research group (Lalezari *et al*., 1971). This neutrophil-specific antigen was formerly known as neutrophil-specific antigen 'NB1' and was later assigned as HNA-2 by the Granulocyte Immunology Working Party of ISBT (Fung and Minchinton, 2011). The 56-64 kDa HNA-2 molecule is coded by the CD177 gene on chromosome 19 (Bux, 2008). Several single nucleotide polymorphisms (SNPs) were reported on the gene but none of them can stimulate the production of alloantibody. However, splicing defect during transcription resulted null expression of HNA-2 molecule (Kissel *et al*., 2002).

2.1.3 HNA-3 system

Three allelic variants (*SLC44A2*1*, **2* and **1:2*) in *SLC44A2* gene on chromosome 19 code for three HNA-3 antigens (i.e. HNA-3a, -3b and -$3a_{var}$). These antigens are expressed on the choline transporter glycoprotein of human cells (neutrophils and lymphocytes) and organs (kidney and spleen) (Fung and Minchinton, 2011; Huvard *et al*., 2012) – see Table 2.1. The HNA-3a and -3b molecules are differentiated by a 461G and 461A SNPs, respectively. This G461A SNP lead to Arg154Gln amino acid changes (Nielsen *et al*., 2012). Another SNP (C451T) on HNA-3a allele attribute to Phe151Arg amino acid changes and the product, HNA-$3a_{var}$ was reported to have different affinity to antibody against HNA-3a molecule.

2.1.4 HNA-4 system

The HNA-4 system was first reported in 1968 and initially described as 'Mart' antigen by Kline and colleagues (Moritz *et al*., 2009; Bux, 2008). There are two alleles within HNA-4 system, *HNA-4a* and *-4b* and their products are expressed on granulocytes, monocytes and natural killer cells. They are located on the αM sub-unit of C3bi receptor (CD11b) and the two alleles differ by a single nucleotide substitution from G to A at position 302 (Fung and Minchinton, 2011). This substitution resulted in changes of amino acid at position 61 from arginine for HNA-4a to histidine for HNA-4b molecule (Bux, 2008). Three phenotypes were determined from the HNA-4a and -4b molecules; HNA-4a/4a, HNA-4a/4b and HNA-4b/4b.

2.1.5 HNA-5 system

HNA-5 was first described as 'Ond' in 1979. The two different types of HNA-5 alleles are; *HNA-5a* and *HNA-5b*, is due to G2466C nucleotide substitution in the coding sequence which resulted in Arg776Thr amino acid changes (Simsek *et al*., 1996). These alleles are encoded by the *ITGAL* gene located on chromosome 16. The HNA-5a and -5b molecules determined three phenotypes; HNA-5a/5a, HNA-5a/5b and HNA-5b/5b.

2.2 Serological and molecular typing of HNA

2.2.1 Serological typing of HNA

Over the years, many techniques have been developed, tested and validated for typing of the clinically relevant HNA molecules (Bux *et al*., 1995; Moritz *et al*., 2009). Those techniques can be broadly assigned into two large group, serological methods and molecular typing methods.

Neutrophils are known as fragile and short lived cells with lifespan ranges from six to eight hours. Serological examination of neutrophils are quiet difficult as it relies on gathering a representative pure population of neutrophils (Fung and Minchinton, 2011). Among the most commonly used serological typing methods are granulocyte agglutination test (GAT), granulocyte immunofluorescence test (GIFT) and monoclonal antibody-specific immobilization of granulocyte antigen (MAIGA) assay.

The GAT technique relies on antibody-HNA reactivity after neutrophil enrichment via dextran sedimentation and followed by gradient centrifugation. Ammonium chloride will be used for removing remaining red blood cells in the mixture. The isolated neutrophils are then incubated with known HNA antibodies. The antibody-HNA complexes can then be observed under an inverted-phase microscope (Bux *et al*., 1997). In contrast to GAT, the GIFT technique involves an additional binding of fluorescent-conjugated antibodies (fluorescein isothiocyanate-conjugated rabbit anti-human antibody) to antibody-HNA complexes. These

fluorescence complexes are detected using flow cytometry or can be observed under any fluorescent microscope (Reil *et al*., 2011).

As for MAIGA assay, selected monoclonal antibodies to HNA are added to antibody-HNA complex and then captured and immobilized in goat anti-mouse antibody-coated micro titre wells. The specific reaction can then visualized by the detection of captured human antibody with alkaline phosphatase-labeled anti human immunoglobulin (Fung *et al*., 2011).

2.2.2 Molecular typing of HNA

Except for HNA-2, phenotyping of other neutrophil granulocytes either by GAT or GIFT is slowly being replaced by molecular method. This is due to inadequate supply or unavailability of suitable alloantibodies against HNAs and the short lived of neutrophils (Moritz *et al*., 2009). The most widely applied molecular technique involves oligonucleotide primer binding to a particular HNA gene and followed by amplification of the targeted region (known as polymerase chain reaction; PCR). Most of the PCR based typing of HNA loci are capable of detecting HNA allelic variants during amplification itself (e.g. PCR with sequence specific primer) or involve post PCR amplification process such as using restriction enzyme (i.e. PCR-restriction fragment length polymorphism (RFLP)) and sequencing (PCR-sequence based typing (SBT)) – see Figure 2.2.

In PCR-SSP, two oligonucleotide primers are used to differentiate two HNA alleles that differ by a SNP (Figure 2.2). Perfectly matched oligonucleotide primer

will be amplified and the amplicons can be examined using agarose gel electrophoresis (Hessner *et al*., 1996). Thus, two reaction mixtures are needed for PCR-SSP technique and each reaction mixture contains a pair of primers targeting highly conserved region in human genome (e.g. growth hormone gene) as an internal positive control (Veldhuisen *et al*., 2014).

PCR-SBT (also known as Sanger sequencing) involves amplification of region of interest using forward and reverse primers and followed by cycle sequencing and capillary electrophoresis (He *et al*., 2014). The PCR-SBT technique allows for actual determination of DNA sequence, including new SNPs, deletion and duplication (Chu *et al*., 2013).

PCR-RFLP is also commonly used for HNA typing. The technique started with targeted amplification of region of interest and followed by digestion using restriction enzyme (Narayanan, 1991). In this context, the two HNA alleles are distinguished by the presence and absence of restriction site (Figure 2.2). In this context, the specificity of the restriction enzyme plays a crucial part and should be able to differentiate the two HNA allelic variants (Ota *et al*., 2007).

Above all techniques that are available for HNA typing, molecular PCR-SSP technique was chosen in the present study on the grounds that it is relatively simple, rapid and requiring limited equipment. Although PCR-SBT is the gold standard for genetic screening, the technique is very costly, laborious and requires expensive DNA sequencer machine. Therefore, PCR-SSP technique is more than sufficient for detecting nucleotide differences between HNA allelic variants (He *et al*., 2014).

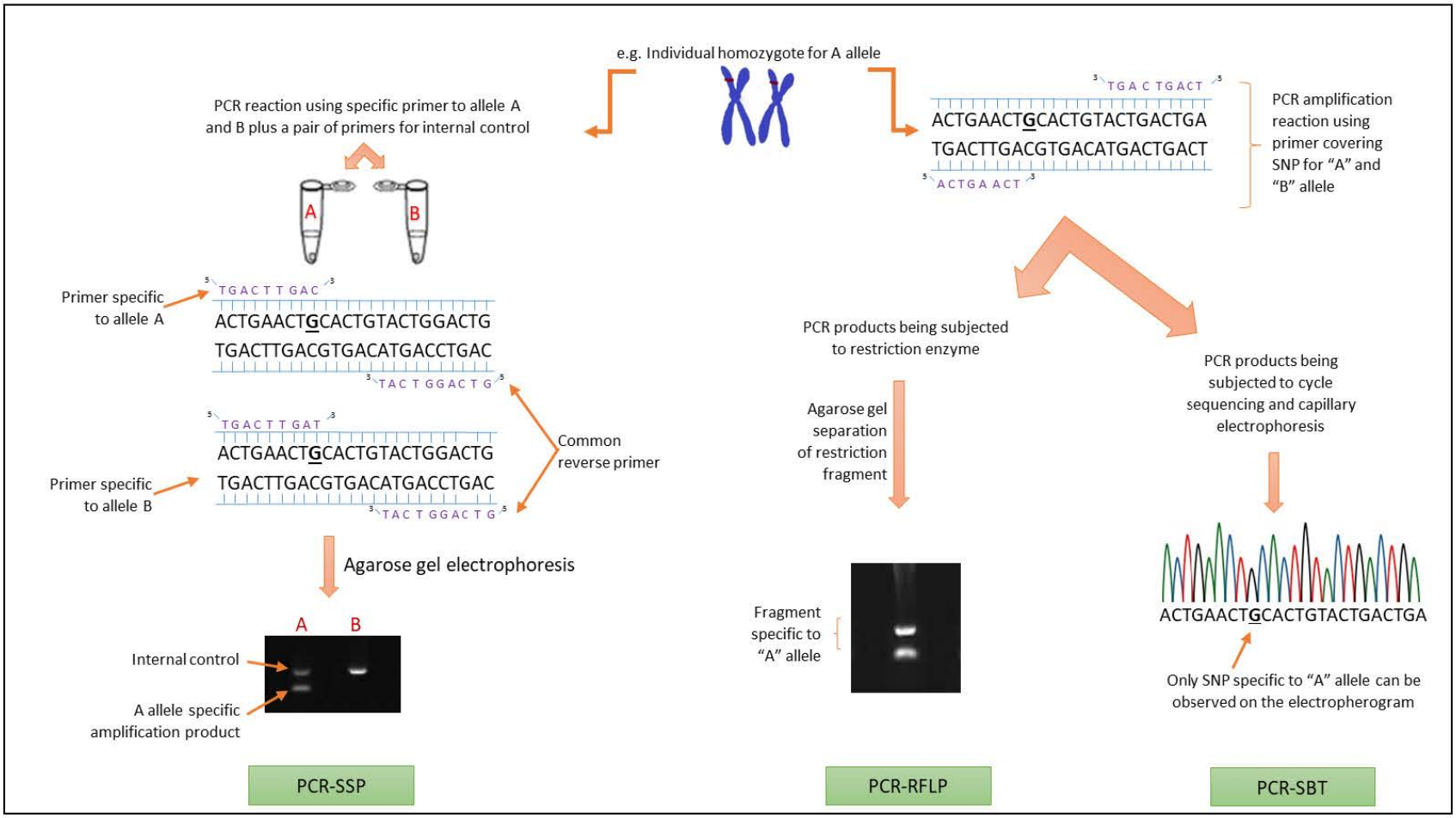

Figure 2.2: Summary of the available molecular HNA typing methods.

2.3 HNA and health

Alloantibodies to HNAs have been reported as contributing factors in certain pathophysiological conditions, including NAN and TRALI (Bayat *et* al., 2012; Cardoso *et al.,* 2013; Davoren *et al*., 2003). These clinical conditions occurred due to HNA incompatibilities in gestation (between incompatible mother and foetus) or HNA mismatched between donor and recipient in transfusion. Details of these clinical disorders are discussed in the following sub-sections.

2.3.1 Neonatal alloimmune neutropenia (NAN)

NAN is a rare disease which account for less than 1% of all birth but an important cause of neonatal neutropenia (Wiedl and Walter, 2010). NAN occurs when a mother become sensitized to a foreign antigen present on foetal granulocytes that is of paternal origin. The mother immune system will produce antibodies against the incompatible paternal HNA types and will lead to the destruction of the foetus's neutrophil granulocytes (Bowens *et al*., 2012). This may lead to severe and sometimes lethal complication in the neonates. However, there are cases where NAN is reported to be caused by isoantibodies when the mother is lacking a complete HNA system-carrying structure (e.g. anti-FcγRIIIb isoantibodies or HNA-2 antibodies) (Porcelijn and de Haas, 2018).

Symptomatic infants with NAN often present with delayed separation of the umbilical cord, skin infection or pneumonia. Most infections are normally mild but severe sepsis is known to occur in some cases (Han *et al*., 2006). However, in some

cases, it may occur asymptomatically and only detected after the first week of life when the newborn being alloimmunized and become febrile (von dem Borne, 1994). Prenatal HNA screening might be useful in order to reduce the occurrence of NAN. For neonatal diagnose with this neutropenia, treatment with antibiotics may be needed to control bacterial infections and increase neutrophil count to normal state. The consequences of NAN to the newborn can be serious even it is a relative rare condition, and timely with good differential diagnosis can prevent more severe complications (Porcelijn and de Haas, 2018).

2.3.2 Transfusion-related acute lung injury (TRALI)

TRALI has been classified as the number one cause of transfusion-related death which occurred when white blood cells antibodies in a transfused blood react with antigens on the recipient's neutrophils and cause lung injury (Bowens *et al*., 2012). Although it is uncommon, the frequency of TRALI has been estimated to be 1 in 5000 unit of transfusion (Popovsky and Moore, 1985). Most of the reported cases of TRALI were those associated with antibody-mediated reaction where alloantibodies against human neutrophil antigens were detected in a transfused blood product of the patient (Juffermans *et al*., 2007). Among the five HNAs system, antibodies against HNA-3a are prone to cause more severe and often fatal cases of TRALI (Bayat *et al*., 2012; Bowens *et al*., 2012). The HNA-3a antibody will bind to neutrophil causing agglutination following by the release of inflammatory cytokines and reactive oxygen species by neutrophils which induce permeability of the lung's endothelial cells (Bux, 2008; Flesch *et al*., 2011).

TRALI should be considered as a differential diagnosis of patients with acute chest syndrome in sickle cell anaemia, diffuse alveolar haemorrhage during bone marrow recovery, lung injury associated with complement-mediated haemolysis in paroxysmal nocturnal haemoglobinuria (PNH) and lung injury associated with granulocyte transfusions (Sanchez and Toy, 2005).

2.4 HNA and human history

Genetics has long been used in ancestry study. It started with phenotyping of serum proteins and globulin genes before slowly been replaced by molecular typing of polymorphic regions in human genome including genetic loci coded for blood groups, human leukocyte antigens, human platelet antigens, HNA, cytokine and killer cell immunoglobulin-like receptors (Norhalifah *et al*., 2016b). Current trend of population genetic research is to maximize the benefit of population datasets especially those that have significant values in health such as human leukocyte antigen and blood groups. This new direction of genetic ancestry study match well with the previous study (Hauck *et al*., 2011; Xia *et al*., 2011; Matsuhashi *et al*., 2012; Nielsen *et al*., 2012; Cardoso *et al*., 2013; He *et al*., 2014; Changsri *et al*., 2013 and Table 2.2) and present survey of HNA allelic variations in Malay sub-ethnic groups. As described in previous sub-sections, HNAs loci are polymorphic (sub-section 2.1) and have been reported as causative agents in NAN and TRALI (sub-section 2.3). Thus, HNA datasets for the Malay sub-ethnic groups from the present study not only increase number of populations being characterized for HNA loci but have significant values for dual analyses of health and ancestry in Peninsular Malaysia.

Table 2.2: HNA allele frequencies in various populations including those reported for Malays and Orang Asli in Peninsular Malaysia.

Populations	HNA allele frequencies									
	HNA-1a	HNA-1b	HNA-1c	HNA-1null	HNA-3a	HNA-3b	HNA-4a	HNA-4b	HNA-5a	HNA-5b
Banjar Malays	0.700	0.300	0.000	0.000	0.867	0.133	1.000	0.000	0.733	0.267
Bugis Malays	0.703	0.293	0.027	0.000	0.676	0.324	1.000	0.000	0.716	0.284
Champa Malays	0.765	0.235	0.039	0.000	0.735	0.265	0.961	0.039	0.657	0.343
Jawa Malays	0.641	0.359	0.000	0.000	0.795	0.205	0.987	0.013	0.910	0.090
Kelantan Malays	0.714	0.286	0.114	0.000	0.743	0.257	0.943	0.057	0.529	0.471
Bateq	0.444	0.556	0.000	0.000	0.926	0.074	1.000	0.000	0.759	0.241
Che Wong	0.556	0.444	0.000	0.000	0.741	0.259	0.870	0.130	0.852	0.148
Kensiu	0.539	0.461	0.000	0.000	0.947	0.053	0.987	0.013	0.789	0.211
Lanoh	0.820	0.140	0.000	0.040	0.860	0.140	0.900	0.100	0.860	0.140
Orang Kanaq	0.773	0.227	0.000	0.000	0.909	0.091	0.864	0.136	1.000	0.000
Semai	0.583	0.345	0.000	0.072	0.940	0.060	0.929	0.071	0.774	0.226
Danish	0.365	0.635	0.060	0.000	0.814	0.186	0.881	0.119	0.724	0.276
Europeans	0.321	0.679	0.029	0.000	0.768	0.232	0.882	0.118	0.736	0.264
Germans	0.391	0.601	0.030	0.008	0.744	0.256	0.908	0.092	0.731	0.269
Han Guangzhou	0.667	0.331	0.000	0.002	0.738	0.262	0.996	0.004	0.854	0.146
Han Zhejiang	0.613	0.387	0.000	0.000	0.654	0.346	1.000	0.000	0.896	0.104
Japanese	0.623	0.377	0.000	0.000	0.654	0.346	1.000	0.000	0.840	0.160
Thais	0.470	0.530	0.005	0.000	0.490	0.510	0.973	0.027	0.790	0.210
Turkish	0.420	0.564	0.030	0.016	0.737	0.263	0.881	0.119	0.754	0.246
Zambians	0.475	0.395	0.390	0.130	0.974	0.026	0.892	0.108	0.500	0.500

HNA datasets were obtained from Manaf *et al*. (2015 and 2016), Nielsen *et al*. (2012), Cardoso *et al*. (2013), Hauck *et al*. (2011), Xia *et al*. (2011), He *et al*. (2014), Matsuhashi *et al*. (2012) and Changsri *et al*. (2013).

2.5 Population structure in Peninsular Malaysia

Malaysia is a country with thirteen states and three federal territories. Eleven states and one federal territory (i.e. Wilayah Persekutuan) are located in East Malaysia (commonly known as Peninsular Malaysia) and others are located in West Malaysia (also known as Borneo). The West and East Malaysia are separated by South China Sea and each has a total land mass of 132,156 km^2 and 198,447 km^2, respectively (see Figure 2.3).

The country of Malaysia is occupied by several ethnicities which can be broadly classified as Malays, Chinese, Indians, Orang Asli and natives of Sabah and Sarawak. The Orang Asli can be further classified into three major groups; namely Semang (a.k.a Negrito), Senoi and Proto-Malays. The Malays, Orang Asli and native of Sabah and Sarawak are known as Bumiputera (original people) and they are either speakers of Aslian (Semang and Senoi) or Malayic (Malays, Proto-Malays and natives of Sabah/Sarawak) languages (Bellwood, 1993).

However, these two large groups of speakers migrated into Peninsular Malaysia at different time during Prehistoric Era. The Semang are the first to populate the Peninsular Malaysia around 50,000 years ago and their other relatives are now scattered as aboriginal people in Australia, Philippines and Papua New Guinea (Norhalifah *et al*., 2016b).

The Aslian language (i.e. Austro-Asiatic language) adopted by the Semang people was introduced by the Orang Asli Senoi who travelled south from Indochina

7,000-6,000 years ago (Bellwood, 1993; Hill *et al*., 2006). Then come Austronesian people (5-3,000 years ago) who migrated from East Asia through Taiwan and Philippines before they finally reach Peninsular Malaysia, Indonesia and become the first settlers in Polynesia (e.g. Samoa, Tokelau and New Zealand) region (Chambers and Edinur, 2013).

In Peninsular Malaysia, these Proto-Austronesian people are known as Proto-Malays but some have admixed with other ethnicities including recent arrivals such as Indians, Chinese and Malay sub-ethnic groups which formed the majority (Deutro-Malays) groups in Peninsular Malaysia (Chambers and Edinur, 2015; Norhalifah *et al*., 2016b). The following sub-sections discuss the origins and migration patterns of the five Malay sub-ethnic groups that being the subjects of the present study.

2.5.1 Acheh Malays

The Acheh Malays of Peninsular Malaysia originated from Aceh, Indonesia. Aceh Empire in Indonesia has once been among the earliest and most successful Malay Empire in South East Asia region by gaining a lot of profit due to its position on the edge of the Malacca Strait (Madjid, 2012). During their golden era, Aceh Empire was a major producer of gold, mine and spices which attract traders from India, Arab and East Asia. However, people from the Aceh Empire were then migrated out including to Peninsular Malaysia due to political struggle with colonial power and internal conflict (Andaya, 2001).

Figure 2.3: Map of Malaysia. This figure is modified from https://www.worldatlas.com/webimage/countrys/asia/lgcolor/mycolor.htm.

2.5.2 Kedah Malays

The Kedah Malays are native to the state in the northern part of Peninsular Malaysia and believed to be among the earliest settlers of the land. In early 16th century, Kedah land was a great store of tin and lead which attracts the British Company. However, the Dutch managed to conquered Kedah between 1650 until 1680 and actively involved in trading of tin, gold and wildlife (Winstedt, 1936). Portuguese and Siamese were then invaded Kedah in the 17th century. However, the Siamese remain as the controller of Kedah Sultanate during the 18th century (Suwannathat-Pian, 1986). History also stated that the Kedah Malays were popular among the Arabian traders which lead to inter racial marriage between them (Teng and Tan, 1979).

2.5.3 Mandailing Malays

The Mandailing Malays were people of the south-western corner of the province of North Sumatera that migrated to the Peninsular Malaysia in the early 18th century (Tugby, 1977). They fled to the west cost of Malay Peninsula after the Padri War via Riau and Jambi. During 1860s, these Mandailing people engaged in mining, trading, mercenary activities and political affairs in Peninsular Malaysia (Lubis, 2003). Most of their settlements located in Selangor, Perak, Negeri Sembilan and Pahang (Tugby, 1977) where they established their networks for migration, kinship, trade, industry and leadership (Lubis, 2005).

2.5.4 Minangkabau Malays

This Malay sub-ethnic group originated from the Minangkabau Highlands of West Sumatera, Indonesia and migrated to Peninsular Malaysia (i.e. Negeri Sembilan) in late 17th and early 18th centuries (Hoh *et al*., 2008). They were then becoming the main sub-ethnic group in Negeri Sembilan. The Minangkabau Malays was a patrilineal society where inheritance is based on the male line. They were then changed to matrilineal system when many men leaving their original homeland for fortune abroad (Madjid, 2012). Since then, the matrilineal society system remains for Minangkabau in Negeri Sembilan where the main lineage of the family is under the mother genealogy and inheritance is passed through maternal lines (Wan Nurhayati *et al*., 2014).

2.5.5 Pattani Malays

The people of Pattani Malays were originated from those reside on the east coast of southern Thailand (Patani) and are much similar to the Malays of Kelantan in term of language and culture (Hoh *et al*., 2008). In the 16th century, East India Company settled in Patani and transformed it to a trading centre with the Siamese Kingdom (Winstedt, 1936). Portuguese, Dutch, English and French were also actively trading textiles, pepper, gold and foodstuffs through the Malay port of Patani. However, due to changes in trading patterns and internal conflict, the importance of Patani as a trading port was declined (Welch and McNeill, 1989) and their people slowly migrated out of Patani.

CHAPTER 3

MATERIALS AND METHODS

3.1 Experimental design

The HNA typing was performed on 194 high molecular weight genomic deoxyribonucleic acid (DNA) samples collected from five Malay sub-ethnic groups (Acheh (n=35), Kedah (n=30), Mandailing (n=47), Minangkabau (n=47) and Pattani (n=35)). This work was approved by Human Research Ethical Committee (HREC) of Universiti Sains Malaysia (USM/JEPeM/1406216 - Appendix A) and was funded by the Short Term Grant Scheme (no: 304/PPSK/61313062), Universiti Sains Malaysia.

Figure 3.1 shows the work flow of the present study which consist of genomic DNA extraction from whole blood samples, DNA quantification, polymerase chain reaction using sequence specific primer typing of HNA loci, amplicons separation using agarose gel electrophoresis and statistical analyses of the generated HNA datasets. All work related to this study was conducted internally at the School of Health Sciences, Health Campus, Universiti Sains Malaysia.

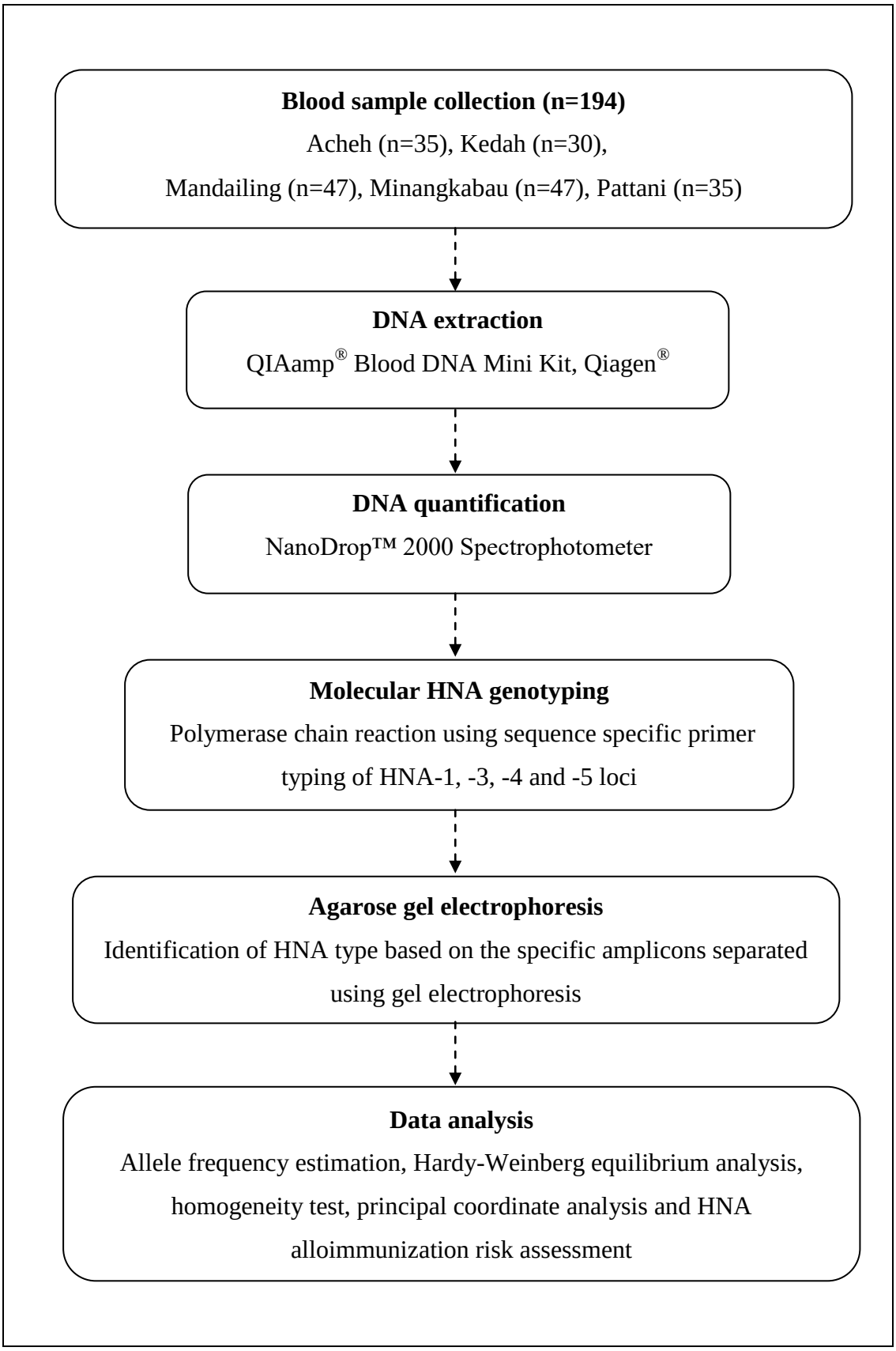

Figure 3.1: The work flow of HNA profiling in five Malay sub-ethnic groups in Peninsular Malaysia.