Study of Human Leukocyte Antigen gene complex and ABO System among the Indo-Aryan Kami population

A Thesis Submitted

To

Sikkim University

In Partial Fulfilment of the Requirement for the

Degree of Doctor of Philosophy

By

Dependra Chamlagai

Department of Zoology

School of Life Sciences

May 2024

Dedicated

To My beloved Parents,

Ashok Kumar Thamlagai, Yashoda Thamlagai

and Aunt

Radhika Thamlagai

In the garden of my ambitions, you've been the nurturing soil, the radiant sunshine, and the gentle rain. This thesis is a testament to the
strength of your support and the depth of your love.

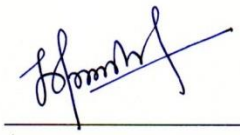
To My beloved wife,

Aruna

You have been the breeze beneath my wings, guiding me through skies uncharted and challenges untold. Your unwavering support and
sacrifice have lifted me higher than I ever imagined possible.

DECLARATION

I, Dependra Chamlagai, declare that the PhD thesis entitled "Study of Human Leukocyte Antigen gene complex and ABO system among the Indo-Aryan Kami population" submitted by me for the award of Doctor of Philosophy in Zoology, Sikkim University under the supervision of Dr. Bisu Singh, Associate Professor, Department of Zoology, Sikkim University is an original research work carried out by me at Department of Zoology, School of Life sciences, Sikkim University, Gangtok. This work has not been submitted for any other degree to this or any other University. All the relevant sources of previous research works have been properly cited.



Mr. Dependra Chamlagai
Department of Zoology
School Of Life Sciences
Sikkim University, Sikkim

Place: Gangtok
Date: 14/05/24

6 माइल, सामदुर, तादोंग - 737102
गंगटोक, सिक्किम, भारत
फोन-03592-251212, 251415, 251656
टेलीफैक्स - 251067
वेबसाइट - www.cus.ac.in

6th Mile, Samdur, Tadong-737102
Gangtok, Sikkim, India
Ph. 03592-251212, 251415, 251656
Telefax : 251067
Website : www.cus.ac.in

सिक्किम विश्वविद्यालय
SIKKIM UNIVERSITY

(भारत के संसद के अधिनियम द्वारा वर्ष 2007 में स्थापित और नैक (एनएएसी) द्वारा वर्ष 2015 में प्रत्यायित केंद्रीय विश्वविद्यालय)
(A central university established by an Act of Parliament of India in 2007 and accredited by NAAC in 2015)

SU/ZOO/2024/17 **10/05/24**

PLAGIARISM TEST CERTIFICATE

This is to certify that a plagiarism check has been carried out for the following Ph.D thesis with the help of DRILLBIT SOFTWARE, and the result is 8% tolerance rate, within the permissible limit (below 10%) as per the norm of Sikkim University.

"Study of Human Leukocyte Antigen gene complex and ABO system among the Indo-Aryan Kami population"

Submitted by **Mr. Dependra Chamlagai** under the supervision of **Dr. Bisu Singh,** Associate Professor, Department of Zoology, School of Life Sciences, Sikkim University.

Dependra Chamlagai
Roll no. - 17PDZ002
Signature of the candidate

Supervisor
Dr. Bisu Singh
Associate Professor
Department of Zoology
School Of Life Sciences
Sikkim University, Gangtok

Vetted by Librarian
Central Library
Sikkim University

पुस्तकालय
LIBRARIAN
केन्द्रीय पुस्तकालय Central Library
सिक्किम विश्वविद्यालय
SIKKIM UNIVERSITY

6 माइल, सामदुर, तादोंग - 737102
गंगटोक, सिक्किम, भारत
फोन-03592-251212, 251415, 251656
टेलीफैक्स - 251067
वेबसाइट - www.cus.ac.in

6th Mile, Samdur, Tadong-737102
Gangtok, Sikkim, India
Ph. 03592-251212, 251415, 251656
Telefax : 251067
Website : www.cus.ac.in

सिक्किम विश्वविद्यालय
SIKKIM UNIVERSITY

(भारत के संसद के अधिनियम द्वारा वर्ष 2007 में स्थापित और नैक (एनएएसी) द्वारा वर्ष 2015 में प्रत्यायित केंद्रीय विश्वविद्यालय)
(A central university established by an Act of Parliament of India in 2007 and accredited by NAAC in 2015)

CERTIFICATE

This is to certify that the PhD thesis entitled "**Study of Human Leukocyte Antigen gene complex and ABO system among the Indo-Aryan Kami population**" submitted to Sikkim University in partial fulfilment for the requirements of the degree of Doctor of Philosophy in Zoology embodies the research work carried out by **Mr. Dependra Chamlagai** at the Department of Zoology, School of Life Sciences, Sikkim University under our supervision. The results presented here are original and have not been submitted anywhere else for any other degree or diploma.

Bisu Singh 14.5.24
.....

Dr. Bisu Singh
(Supervisor)
Head of Department
Associate Professor
Department of Zoology
School Of Life Sciences
Sikkim University, Sikkim

Head
Department of Zoology
Sikkim University

ACKNOWLEDGEMENT

In embarking on my enriching Ph.D. journey in 2017, I have been profoundly fortunate to receive an outpouring of assistance and support from many caring and generous individuals. With sincere gratitude, I am delighted to extend my heartfelt thanks to all who have contributed directly or indirectly to realising my doctoral aspirations.

First and foremost, I extend my heartfelt gratitude to my Supervisor, Dr. Bisu Singh, Associate Professor and head of the Department of Zoology, whose unwavering support has been immense and immeasurable. His steadfast belief in my potential, dating back to my M.Sc. days, has significantly shaped my academic and personal trajectory. I am deeply thankful for the opportunity he provided me to join as a project fellow under his mentorship. His consistent motivation, invaluable guidance, and belief in my capabilities have been the cornerstone of my progress. It is noteworthy that, through his continuous inspiration, invaluable advice, selfless guidance, and astute suggestions, the completion of this research became achievable. I am profoundly grateful for working under such a dedicated supervisor who consistently prioritises his students and upholds academic standards of excellence, which have driven my commitment and diligence. Under his guidance, I have acquired priceless insights into diverse aspects of academia and garnered invaluable career lessons that will forever hold a special place in my heart.

I sincerely thank Prof. Dr. Bhoj Kumar Acharya, Department of Zoology, Sikkim University, for his valuable suggestions and encouragement throughout my academic journey. His belief in me and his confidence in my abilities have been a driving force, guiding me through challenges and uncertainties. I extend special thanks to all the faculties of the Department of Zoology, Sikkim University, Dr. Basundhara Chettri, Dr. Sudeep Ghatani, Dr. Arnab Banerjee, Dr. Suman Dahal (Guest faculty) and Dr. Kishore Sharma (Guest faculty) in the Department of Zoology for their invaluable guidance, suggestions, and encouragement throughout my study.

I extend my most profound appreciation to Dr. D. Purohit, the esteemed Principal of NBBGC, Tadong, Sikkim, whose constant support and cooperation have been a beacon of stability and encouragement. I want to express immense gratitude to Dr. Namrata Tamang, the esteemed Head of the Department of Zoology at NBBGC; I extend my heartfelt thanks for her invaluable support, cooperation, and guidance. Her dedication to nurturing and fostering academic growth has been a source of inspiration.

Furthermore, I am deeply grateful to Mr. Bedanta Saikia, Technical Assistant, Department of Zoology, Sikkim University, for his consistent support, timely assistance, and valued company throughout this journey. His immediate responses to my requests and willingness to help have been instrumental in overcoming numerous challenges.

Amidst the peaks and valleys of my doctoral journey, friends have enhanced my academic endeavours and offered solace and encouragement during challenging times. In the harmony of friendship, I am compelled to single out the unwavering support of Dr. Jiwan Gurung, a true rarity in the realm of friendship. His hands in the laboratory were always ready to help, supporting me through crucial experiments and sharing his expertise in data analysis, embodying the essence of true companionship. A noble soul with a heart of gold, Dr. Gurung stood steadfastly by my side through every challenge, infusing warmth and encouragement into every interaction. Equally deserving of recognition is Dr. Sailendra Dewan, whose constant support and friendship were essential to my doctoral journey. I am profoundly grateful to Dr. Prerna Dhakal for her unwavering support, timely advice, and cherished companionship throughout my Ph.D. journey. Additionally, I must express my deep appreciation to Mrs. Ananta Rai, whose invaluable guidance and steadfast support served as pillars of strength, guiding me through with her uplifting presence.

I extend my heartfelt appreciation to all the scholars of the Department of Zoology, including Mr. Bishal Thakuri, Dibyan Lama, Roshan Tamang, Dipsikha Saikia, Ishwar Chettri, Aita Hang Subba, Rabina Mothey, Siwangi Rai, Subodh Chettri and Anuj

Mukunda Sharma. Furthermore, Mr. Subodh Chettri's technical expertise proved invaluable in the analysis pertinent to my research. In the realm of gratitude, I reserve a special place for Ms. Tanushree Das, whose unwavering support and sisterly love guided me through the intricate process of structuring my thesis. Lastly, I profoundly appreciate Ms. Primila Neopaney's invaluable companionship and unwavering support. To all these remarkable individuals, I extend my deepest gratitude. Their presence has made this journey bearable and beautiful; I am eternally thankful.

As I express my gratitude to my family, it is a great privilege to express my sincere thanks and deep gratitude to my beloved father, mother, and aunt for their sacrifice, encouragement, inspiration, and unconditional love. Your sacrifices, seen and unseen, have shaped me into who I am today, and I am profoundly grateful for that. I also want to express my appreciation to my siblings, Dr. Manoj Chamlagai and Dinesh Chamlagai, whose steadfast support and camaraderie have brought strength and joy to my journey. Your encouragement and shared laughter have been comforting, even in the most challenging moments. Further, I extend my gratitude to Mr. Bhanu Chapagai for constant support and motivation.

I would also like to extend special gratitude to my wife, whose unwavering love has served as a steadfast anchor amidst the tumultuous seas of academia. Her enduring support and sacrifices, both seen and unseen, have been instrumental in navigating the challenges of this scholarly pursuit. Her dedication and selflessness have profoundly enriched my academic journey, from managing household responsibilities to providing emotional encouragement.

In expressing gratitude to my family, I reflect on my cherished companions who enriched my Ph.D. journey. I extend special mention to my dear brothers in crime, Dr. Tika Ram Pokhrel, Mr. Mayesh, and Bhawani, for the countless cherished memories and laughter-filled moments outside the confines of academia. I owe a debt of gratitude beyond measure to Dr. Tika Ram Pokhrel for his invaluable assistance in formatting my thesis. Amidst his commitments and responsibilities, he extended a helping hand with

unwavering generosity, ensuring that my work was presented with the utmost precision and professionalism.

Finally, I thank DST-SERB for their generous financial support of this study (Sanction No. ECR/2015/000296). My deepest appreciation goes to all the study participants whose selfless contribution through donating their blood samples has been invaluable to the advancement of this research. I am also grateful to Mrs. Sangeeta Darnal (Singh), Mr. Riben Ghatani, Mr. Mythless Baraily, Mr. Rohit Biswa, Mr. Suren Ghatani, Dr. Amir Darjee, Bharatiya Kami Unnayan Parishad (BKUP), and all the volunteers for their support and assistance throughout this journey. Above all, I am ever thankful to the divine, recognising that **"Everything has raison d'être."**

Dependra Chamlagai

Place: Gangtok, Sikkim

Contents

1 Introduction **1**

1.1 Introduction 1

1.2 Human Origins 3

1.2.1 Recent African Origin (RAO) 3

1.2.2 RAO and Hybridization (RAOH) 4

1.2.3 Assimilation Model (AM) 4

1.2.4 Multiregional Evolution (MRE) 4

1.2.5 Braided Stream model (BS) 5

1.3 The Evolution of Humans 5

1.4 Human Migration and Genetic Diversity 9

1.5 Determinants of Genetic Diversity 11

1.5.1 Natural selection 12

1.5.2 Genetic Drift 12

1.5.3 Mutations 13

1.5.4 Gene flow 13

1.6 Genetic Diversity and the Indian Subcontinent 13

1.7 Measuring Genetic Diversity 17

1.8 Use of Genetic markers to understand Genetic Diversity 19

1.8.1 mtDNA 19

1.8.2 Y-Chromosome 20

1.8.3 Autosomal STRs 20

1.8.4 Major Histocompatibility Complex 21

1.9 The rationale for the current investigation 22

1.10 Objectives of the Study 24

2 Literature Review **25**

2.1 Literature and Review 25
2.2 The Kami population 25
2.3 Human Leukocyte Antigen Gene Complex 26
2.4 Genetics of Major Histocompatibility Complex 27
2.5 HLA Nomenclature 30
2.6 Allelic polymorphism and Haplotypes at HLA loci in various populations 31
2.7 HLA diversity among the world population 32
2.8 HLA diversity among the Indian population 35
2.8.1 South Indian Studies 38
2.8.2 North Indian studies 40
2.8.3 West Indian Studies 42
2.8.4 East Indian Studies 58
2.9 HLA and Disease Association 59
2.10 ABO Blood group system 62
2.11 Biochemical Genetics of the ABO system 63
2.12 ABO system and population studies 64
2.12.1 International Status 64
2.12.2 National Status 66

3 Materials and Methods **68**

3.1 Ethical considerations 68
3.2 Population Sampling for HLA Study 68
3.3 Blood sample collection 69
3.4 Post-sampling stratification for disease association study 70
3.4.1 Selection Criteria for Disease Association Study 70
3.5 Extraction of Genomic DNA 71
3.6 Genotyping of HLA-A, B and DRB genes 72
3.7 Statistical analysis 74
3.7.1 Genetic Affinity Analysis 76
3.7.2 Statistical Analysis for Disease Association Study 76

3.8 Study area for ABO blood group study 76
3.9 ABO and Rh blood group determination 77

4 Results **79**
4.1 Results 79
4.2 Hardy Weinberg Equilibrium test 79
4.3 Frequency Distribution of genes at HLA-A locus in the Kami Population 80
4.4 Frequency Distribution genes at HLA-B locus in Kami Population . . . 80
4.5 Allele Frequency Distribution at HLA-DRB1 locus in Kami Population. 81
4.6 Haplotype Distribution in the Kami population. 82
4.7 Pairwise Linkage Disequilibrium 83
4.8 Genetic Relatedness of Kami Population with Other Populations. 84
4.9 Disease Association study 88
4.10 Results of ABO and Rh Blood Group Study in the Kami Population . . 89
4.10.1 Distribution of ABO and Rhesus blood group 89
4.10.2 Distribution of ABO and Rhesus blood group allele frequencies 91
4.10.3 Distribution of the ABO blood group in the subcastes 91

5 Discussion **94**
5.1 Discussion 94
5.2 Allelic Polymorphism of HLA-A, HLA-B, and HLA-DR Loci in the Kami Population 95
5.3 Comparative Haplotype Distribution in Kami population 98
5.4 Genetic affinity of the Kami population based on HLA 100
5.5 Disease association study 102
5.6 ABO blood group and Rh system in the Kami population 105

6 Conclusion **110**
6.1 Conclusion 110

References **112**

List of Tables

2.1 HLA-A, B, and DRB1 allelic polymorphisms among global populations. 36
2.2 HLA polymorphism among various Indian populations. 45
2.3 HLA alleles and their associations with diseases. 60
2.4 ABO blood group: genotypes, phenotypes, red blood cell surface antigens, and corresponding serum antibodies. 62
2.5 Distribution of ABO blood group in global populations. 65
2.6 ABO blood group distribution order in Indian populations. 67

3.1 Components and the volumes used to prepare the master mix. The stated total volume of the master mix contains a pipetting excess for seven extra wells. 73
3.2 PCR thermocycler program settings. 73

4.1 Results of hardy weinberg's equilibrium test for HLA-A, HLA-B, and HLA-DRB1 loci. 80
4.2 Allele frequency distribution at HLA-A locus in Kami polulation. . . . 81
4.3 Allele frequency distribution at HLA-B locus Kami polulation. 82
4.4 Allele frequency distribution at the HLA-DRB1 locus in Kami polulation. 83
4.5 Three locus HLA A B DRB1 haplotype among the Kami population. . . 84
4.6 Two locus HLA A∼B and HLA-B∼DRB1 haplotypes among the Kami population. 85
4.7 Pairwise linkage disequilibrium analysis results for HLA Loci in the Kami population. 85
4.8 Frequencies for HLA-A, -B, -DR alleles in type-2 diabetes mellitus individuals and controls belonging to Kami population. 89

4.9 Allele Frequency of HLA-A, -B, -DR in hypertensive individuals and normotensive controls belonging to Kami polulation. 90
4.10 Distribution of ABO and rhesus blood group percentage frequencies among the Kami population 90
4.11 Distribution of ABO and rhesus blood group allele frequencies among the Kami population 91
4.12 Distribution of the ABO blood group in the subcastes of the Kami population 92

5.1 Comparative analysis of ABO blood group and allele frequencies across diverse populations. 108

List of Figures

1.1 Migrations of early humans across the world. Photo courtesy of national geographic society, credit to international mapping. sources, Chris Stringer, 2010. 10

2.1 Schematic representation of classical Class I, Class II and Class III HLA genetic loci at chromosome 6 p21.1–21.3. 28
2.2 HLA nomenclature system. 31
2.3 Schematic representation of ABO blood group antigen determination pathways. 64

3.1 Geographic distribution of sampling sites for the HLA study. 69
3.2 Gel Electrophoresis image of genomic DNA extracted from blood samples, showing bands and molecular weight markers (ϕX174). 71
3.3 PCR-SSP HLA typing tray. 73
3.4 Gel electrophoresis image depicting distinct bands of HLA-A alleles: Lane M exhibits varying-sized DNA markers, Lane 1 serves as a Negative control, and Lanes 2 to 25 display bands corresponding to internal control gene amplicons positioned near the wells. Notably, Lanes 5, 14, 19, and 23 depict bands indicative of HLA-A alleles. 74
3.5 Gel electrophoresis image depicting distinct bands of HLA-B alleles: Lane M displays DNA markers of diverse sizes, while Lanes 1 to 19 exhibit bands corresponding to the positive control near the wells. Lanes 8, 16, 17, and 18 notably depict bands indicative of HLA-B alleles. . . . 75

3.6 Gel electrophoresis image displaying distinct bands of HLA-DRB1 alleles: Lane 0 presents DNA markers of various sizes, while Lanes 1 to 24 exhibit bands corresponding to the positive control near the wells. Bands indicative of HLA-DRB1 alleles are observed in Lanes 6, 7, 13, 18, and 22. 75
3.7 Map showing sampling sites for the ABO blood group distribution study. 77
3.8 ABO and Rh typing using the slide agglutination method. 78

4.1 Hierarchical cluster analysis based on HLA-A and -B allele frequency. Dendrogram depicting the cluster of 34 populations (including the population under investigation in the present study). 86
4.2 Principal component analysis based upon the frequency of HLA-A and -B alleles showing the genetic affinity of the Kami population. 87

Abbreviations

CI	: **Confidence Interval**
DNA	: **Deoxyribonucleic Acid**
EDTA	: **Ethylenediaminetetraacetic acid**
GalNAc	: **N-Acetylgalactosamine**
GTA	: **N-acetylgalactosaminyltransferase**
GTB	: **Galactosyltransferase**
HCA	: **Hierarchical Cluster Analysis**
HLA	: **Human Leukocyte Antigen**
HT	: **Hypertension**
HWE	: **Hardy-Weinberg Equilibrium**
LD	: **Linkage disequilibrium**
MHC	: **Major Histocompatibility Complex**
MLC	: **Microlymphocytotoxicity**
mtDNA	: **Mitochondrial Deoxyribonucleic acid**
NGS	: **Next Generation sequencing**
OR	: **Odds Ratio**
PCA	: **Principal Cluster Analysis**
PCR	: **Polymerase Chain Reaction**
RA	: **Rheumatoid arthritis**
SBT	: **Sequence-based typing**
SLE	: **Systemic lupus erythematosus**

SNPs : **Single nucleotide polymorphisms**

SSOP : **Sequence-Specific Oligonucleotide Probe**

SSP : **Sequence Specific Primer**

STRs : **Short tandem repeats**

T2DM : **Type 2 Diabetes Mellitus**

Publications

1. Chamlagai, D., Gurung, J., & Singh, B. (2024). Distribution of ABO and Rhesus (D) blood group antigens among the Indian Kami population from the Sub-Himalayan region of West Bengal and Sikkim, India. Asian Journal of Transfusion Science, 10-4103.

2. Singh, B., Chamlagai, D., & Gurung, J. (2021). HLA Profile of Kami Population Refutes the Earlier Proposition of Exclusive Closer Genetic Affinity of All the Gorkhas to Mongoloids. Human Heredity, 85(1), 45-50.

3. Singh, B., Chamlagai, D., & Gurung, J. (2020). HLA-A-B and-DRB1 distribution in Kami: A caste population of Gorkha community from the sub-Himalayan region of West Bengal, India. Human Immunology, 81(8), 395-396.

4. Chamlagai, D., Gurung, J., & Singh, B. (2024). HLA Association with Type 2 Diabetes Mellitus and Hypertension among the Kami Population from Sub-Himalayan West Bengal. International Journal of Human Genetics, 24(2), 199-204.

5. Chamlagai, D., & Singh, B. (2016). Study of in vitro anti-inflammatory activity of ethnomedicinal plants of Sikkim Viscum articulatum and Acorus calamus. Asian journal of pharmaceutical and clinical research, 119-122.

6. Gurung, J., Chamlagai, D., Bera, N. K., Chaudhuri, T. K.,& Singh, B. (2018). Elevated levels of C-reactive protein and IL-6 among the antipsychotic medicating

schizophrenia patients of Siliguri, West Bengal, India. Nordic journal of psychiatry, 72(4), 311-317.

Summary

The fascination with genetic variation has long captivated researchers, especially within diverse human populations, and this enduring intrigue serves as the impetus for this comprehensive investigation. This investigation was undertaken to explore the intricate genetic makeup of Kami Population belonging to Gorkha community based on Human Leukocyte Antigen (HLA) gene complex and the ABO and Rh blood group systems. The study may provide meaningful insights into the genetic diversity of the Kami population.

The study protocol was approved by IEC Sikkim University. After obtaining informed consent from participants, blood samples were collected from 158 unrelated individuals from various geographical regions of Sub-Himalayan West Bengal, including Darjeeling, Kalimpong, Kurseong, and Siliguri. Furthermore, for the ABO study a total cohort of 1420 Kami individuals, were recruited from regions namely Darjeeling, Kalimpong, Kurseong, and Siliguri of West Bengal and all six districts of Sikkim namely Gangtok, Pakyong, Soreng, Gyalshing, Namchi, and Mangan. In the laboratory of the Department of Zoology, Sikkim University, genomic DNA was isolated from the collected blood samples, and the HLA class I genes (A and B) and the HLA class II gene, DRB1, were genotyped utilising the HLA ABDRB1 PCR SSP typing kit. Further, blood samples collected from 158 individuals were used for ABO blood group serological typing, employing the slide agglutination technique. Additionally, ABO blood group data for 1262 individuals were collected through inquiries with individuals who had obtained their blood group information from diagnostic centres.

The present study identified sixteen distinct HLA-A genes, prominently featuring HLA-A*11 (31.01%), HLA-A*24 (21.84%), and HLA-A*33 (19.62%) as the most prevalent alleles. Additionally, twenty-one HLA-B genes were characterised, with HLA-B*15 (23.42%), HLA-B*40 (18.67%), and HLA-B*35 (13.92%) being the most frequent. Among the twelve identified HLA-DRB1 genes, HLA-DRB1*15 exhibited the highest frequency at 41.14%, followed by HLA-DRB1*14 (14.87%) and HLA-DRB1*07 (9.18%). Notably, the prevalence of HLA-DRB1*15 in this study is the highest reported among North Indian populations, signifying a notable finding. A total of one hundred and thirty-four putative three-locus haplotypes of HLA-A∼B∼DRB1 were delineated in our investigation. Among these, twenty-two haplotypes were observed with a frequency equal to or greater than 1%. The most prevalent haplotype identified in the Kami population was HLA-A*02∼B*40∼DRB1*15 (5.57%). It is a well-known Caucasian haplotype shared with other North and South Indian populations. Interestingly, the second most frequent haplotype, HLA-A*11∼B*15∼DRB1*15 (4.81%) observed among the Kami population is reported for the first time in an Indian population. Moreover, our analysis of pairwise linkage disequilibrium revealed non-random associations among alleles at the examined loci. These findings suggest potential genetic linkage or population-specific allele frequencies, adding depth to our understanding of the Indo-Aryan Kami Population's genetic diversity.

The dendrogram constructed based on HLA-A and HLA-B profiles exhibits a clear distinction between the Kami population and Mongoloid populations, confirming their genetic separation. As anticipated, the dendrogram illustrates a close clustering of the Kami population with Indo-Aryan-speaking populations, indicating genetic affinity or shared ancestral heritage with this linguistic group. Notably, previous genetic investigations have inadequately represented Indo-Aryan-speaking Gorkhas, primarily focusing on Tibeto-Burman-speaking Gorkhas with Mongoloid characteristics. This limited scope has led to generalisations suggesting a Mongoloid origin for the Gorkha community to which the Kami population belongs. Our results of hierarchical cluster analysis

in conjunction with principal component analysis refute the exclusive Mongoloid lineage attribution to the Gorkha population. The finding of the present study cautions against overly generalising the Gorkha community for genetic studies, which encompasses diverse populations, including Indo-Aryan Caucasoids and Tibeto-Burman Mongoloids, living harmoniously in close proximity. Furthermore, considering the historical migration of Kami ancestors from princely states in India to Nepal and, subsequently, to sub-Himalayan West Bengal, our findings strongly advocate for the Indian origin of the Kami population. This underscores the complex and multifaceted genetic dynamics within the Gorkha community, highlighting the need for nuanced and cautious approaches in cumulative studies concerning this diverse ethnic group.

The 158 randomly selected Kami individuals screened showed Hypertension (HT) and Type 2 Diabetes Mellitus (T2DM) as the predominant diseases within the cohort. Given the established associations between HLA genes and these diseases in previous studies, our investigation sought to elucidate the specific HLA gene associations in our study population. In patients with T2DM, our findings revealed a significantly elevated frequency of HLA-B*13 and HLA-B*15 alleles. Conversely, the frequency of HLA-A*33 was notably reduced among these patients. In contrast, patients with HT exhibited a significantly increased frequency of HLA-A*24 and HLA-B*13 alleles. This highlights the potential contribution of HLA genes to the pathogenesis and susceptibility of these common chronic diseases.

Our results showed that in the Kami population, the overall distribution of ABO blood groups is as follows: O (31.06%), B (29.15%), A (27.54%), and AB (12.25%). This pattern aligns with previous observations within the Kami community but diverges from findings in the broader Sub-Himalayan region, where the sequence typically follows B $>$ O $>$ A $>$ AB. Furthermore, a significant proportion of individuals were Rh (D) positive (97.89%), while the remaining were Rh (D) negative (2.11%). The allele frequencies of the ABO blood group, when ordered by prevalence, were identified as $I^O > I^B > I^A$. The prevalence of blood group O was notably higher across twelve subcastes (Lohar, Diyali, Dural, Gajmer, Ghimirey, Khati, Lohar Lohagun, Setisural, Sashankar,

Sinchury, and Tirwa) of Kami population, indicating its predominance within this population subset. Conversely, blood group A exhibited a higher incidence among seven other subcastes (Baraily, Gadaily, Ghatraj, Lakandri, Kalikote, Portel, and Lamgade). In comparison, blood group B predominated in seven additional subcastes (Rohpal, Ramudamu, Darnal, Ghatani, Rasaily, Shilal, and Singouray). Blood group AB, characterised by its unique dual antigen profile, was most prevalent among three specific subcastes (Poudel, Dutraj, and Rizal). These findings contribute to our understanding of genetic diversity within the Kami population and highlight intriguing contrasts with neighbouring populations. The observed variations highlight the complex interplay of genetic factors influencing blood group distributions, suggesting a dynamic relationship between genetic heritage, environmental influences, and the resulting phenotypic traits within the Kami population.

This research endeavour meticulously highlighted the genetic composition of Kami Population based on HLA, ABO, and Rh systems, offering profound insights into population genetics and genetic variation. The study unveils crucial findings regarding the genetic makeup and immunogenetic variations prevalent in this unique population cohort through rigorous genetic analysis and statistical methodologies. The present findings may not only significantly contribute to the field of population genetics and immunogenetics but also have far-reaching implications for healthcare practices, disease susceptibility, and the development of personalised medicine strategies tailored to the genetic intricacies of this population group. The present study warrants further research into the diverse genetic landscapes of Himalayan populations, which may propel advancements in precision medicine and population-based health interventions.

Chapter 1

Introduction

1.1 Introduction

In the explorations of human genetic variation, Lewontin's seminal book chapter, 'The Apportionment of Human Diversity,' published in 1972, remains a cornerstone as it delineated prevailing patterns of genetic variation among human populations. Since then, geneticists have made significant progress in understanding how the evolutionary history of our species has led to the current patterns of genetic variations in humans [1]. Studying human genetic diversity can provide insight into historical occurrences like migrations, colonialism, expansions, and selection. The genetic history of a group of populations is often examined by constructing a phylogenetic tree to elucidate their origins. However, the veracity of these reconstructions depends on the assumption that genetic divergence among populations primarily arises from population fissions, followed by independent evolutionary paths [2]. The intricate structure of genetic variability found in contemporary human populations results from complex interactions between genes and the environment through evolutionary processes. Gene-specific factors like mutation, recombination rates, selection pressure, and previous demographic history (effective population size, substructure, and migration) all have an impact on the

variation patterns seen in the modern human population. Genome variations serve as valuable genetic markers for various applications in diverse fields, including population genetics, drug response prediction, organ transplantation, forensic investigations, and routine clinical testing [3–6].

The ABO red cell blood group antigens, discovered through matching donors and recipients for blood transfusion, were the first polymorphic gene system identified in humans. Indeed, the very first documented human genetic markers were the ABO system antigens, which were also the first blood types to be discovered [7, 8]. This was a turning point for the advent of immunogenetic studies that focused on analysing human genetic variation. Genetic polymorphism across human populations is evident in the genes encoding essential components of the human immune system, including immunoglobulins, Human Leukocyte Antigen (HLA) molecules, and Killer-cell Immunoglobulin-like Receptors (KIR). They exhibit a high amount of polymorphism, displaying significant variations among human populations. The genesis of this polymorphism is likely due to natural selection and is majorly driven by host-pathogen interactions. HLA polymorphism is substantial in human population genetics and has been extensively examined through serological and molecular genetic techniques across various world populations [9–11]. Comprehensive analyses of the HLA data along with history, geography, archaeology, linguistics, and anthropology have yielded novel insights into the trajectories of human evolution and migration [12–14]. Studying these immune system markers provides an excellent opportunity to learn more about the evolution of modern humans and how they adapted to past environmental (e.g., pathogenic) changes. Current advancements in population genetic analysis and computer simulation improve our ability to discriminate among different stochastic or deterministic forces acting on the genetic evolution of human populations. As such, immunogenetic polymorphisms can be considered essential and complementary tools for anthropological studies, alongside markers such as Y-chromosomes, mitochondrial DNA, microsatellites, and single nucleotide polymorphisms [15].

1.2 Human Origins

The tale of modern human origin is still evolving and somewhat enigmatic, with the data lacking the clarity necessary to distinguish between several possibilities in many cases [16]. Nevertheless, with the advent of modern molecular tools in the late 1980s, genetic and morphological investigations have demonstrated that all humans are more closely related to one another than any late quaternary hominin group [17]. Current evidence suggests that most of the important events in the evolution of hominins occurred in East Africa [18]. Despite this uncertainty, current evidence points to East Africa as the focal point for many significant events in hominin evolution [18]. The discovery of the earliest anatomically modern human fossils in Ethiopia [19, 20] strongly indicates that the species originated in sub-Saharan Africa approximately 200,000 years ago (200 ka), a timeline that aligns with the genetic findings [21]. The four hypothesised models concerning modern humans' evolutionary history include the Recent African Origin (RAO), RAO with Hybridisation (RAOH), Assimilation (AM), and Multiregional Evolution (MRE) models. The ' braided stream ' model is a novel addition to these established models, garnering recent scholarly interest [22]. Aiello (1993) and Stringer (2002) provided conceptual frameworks that served as the foundation for the current investigation into the origins of our species[22, 23]. These models are described below.

1.2.1 Recent African Origin (RAO)

By the late 1980s, early explorations into human molecular diversity suggested that our species emerged from a confined African population [24–26]. Around 120–200 thousand years ago, researchers hypothesised that this population underwent global expansion, displacing earlier hominids. This replacement model was later termed the Recent African Origin or the "Out of Africa" Model. This model asserts that modern humans originated in Africa approximately 100,000 years ago and subsequently dispersed

worldwide. The migrating populations replaced premodern indigenous populations in various regions, with limited to no interbreeding between these groups.

1.2.2 RAO and Hybridization (RAOH)

The RAO and hybridisation model exhibit similarities to the RAO model but facilitate a more extensive degree of hybridisation between the migrating and indigenous premodern populations [22, 23].

1.2.3 Assimilation Model (AM)

The Assimilation Model (AM) represents a model of modern human evolution that accentuates an African origin for modern humans, coupled with a consistent pattern of relatively low, albeit noteworthy, contributions from Eurasian archaic populations to modern human gene pools during migrations across the Old World [27]. However, it deviates from prior models by rejecting replacement or population migration as a significant process in the emergence of modern humans. Instead, this model highlights the significance of gene flow, admixture, evolving selection pressures, and subsequent directional morphological changes [22, 23].

1.2.4 Multiregional Evolution (MRE)

Multiregionalism stands apart from the preceding three models by refuting a recent African origin for modern humans. It underscores the importance of genetic continuity over time and gene flow among contemporary populations, asserting that the rise of modern humans occurred not solely in Africa but also in Europe and Asia, stemming from their Middle Pleistocene ancestors [22, 23].

1.2.5 Braided Stream model (BS)

The Braided Stream model (BS) compares the evolution of humans across different parts of the globe during the Pleistocene to a stream with multiple larger or smaller channels. These channels split and combine to form a pattern that resembles a braid's threads. The smaller streams represent isolated populations, gene pools, or lineages during evolutionary events. Over time, these streams either become extinct or reunite into a shared gene pool, ultimately contributing to the global diversity of modern humans [22].

1.3 The Evolution of Humans

Exploring evolution aligns with the pursuit of knowledge, a universal motive that extends to comprehending the past, anticipating the future, and organising our understanding of the world. Fossils and molecular biology data have been the primary sources of information researchers use to understand the complexity of human evolution. Since fossils are the only material remains of past life, they are well-positioned to play an essential role in determining evolutionary relationships. Fossils provide the exclusive opportunity to observe evolutionary events from the remote past. Therefore, incorporating fossils into a phylogenetic framework is essential for precisely discerning their relationships and understanding evolutionary history [28]. The progression of hominin evolution, as observed in the fossil record, unfolds through four key stages: (i) the emergence of the earliest (proto) hominins, attributed to Sahelanthropus, Orrorin, and Ardipithecus genera, dating between 4 and 7 million years ago (Mya); (ii) the appearance of the Australopithecus genus around 4 Mya, accompanied by the robust Paranthropus genus around 2.7 Mya; (iii) the emergence of the Homo genus at the Plio-Pleistocene boundary, spanning 1.8 to 2.5 Mya; and (iv) the arrival of Homo heidelbergensis at 800 Kya and anatomically modern humans at around 200 Kya, as documented by Maslin et al. (2015) [18]. Taxa classified in genera other than Homo (Sahelanthropus, Ardipithecus,

Ororrin, Kenyanthropus, Australopithecus, and Paranthropus) that have a closer relationship to humans than to extant apes are known as early hominins. Frequently referred to as the 'bipedal apes' period, this phase of human evolution commenced between 7.0 and 2.8 million years ago, marked by the discovery of the first Homo fossil. The latter part, spanning from around 2.8 million years ago to the present, revolves around our genus and covers the evolution of modern humans. During this period, attention is directed toward encephalisation, alterations in life-history strategy, the advancement of technology, the broadening of diet and geographical range, and the emergence of cultural processes in evolution [29].

Although all fossils found to date have contributed to our understanding of human evolution, certain discoveries have provided deep insights into the morphological transformations and adaptive strategies that characterised this evolutionary process. Understanding the shift to bipedality is paramount in revealing the early stages of hominin evolution. Australopithecus afarensis, the most extensively studied early hominin species, is a fundamental reference point for interpreting locomotion in all early hominins. Australopithecus afarensis is widely acknowledged as exhibiting habitual bipedalism among researchers [30]. Similarly, Homo habilis is characterised by a skeleton that shows a combination of primitive and evolved features, suggesting an early hominid capable of bipedal walking, stone tool fabrication, and the retention of generalised hominoid abilities, including climbing trees [31]. Homo erectus achieved improved walking abilities in open environments, albeit at the cost of retaining adaptations for arboreal locomotion observed in the postcranial remains of australopithecines [32]. The migration of Homo erectus from Africa around 1.9 million years ago stands as one of the most critical, pivotal, and controversial events in the evolution of humans [33]. Additionally, studies have demonstrated that Neanderthals exhibited cognitive traits typically associated with contemporary human cognition, including language [34]. This furthers our understanding of our ancestors' evolutionary journey and the physical adaptations previously discussed.

Even though fossil specimens and morphological comparisons of living taxa have conventionally offered clues to understanding human origin and evolution, they are not devoid of limitations. There are biases in the preservation potential of fossils; organisms with durable skeletons possess a better chance of being preserved than those with softer bodies [35]. Furthermore, due to its incompleteness, the usefulness of the fossil record in providing a clear account of evolutionary history has been challenged [36]. Since some taxa are not preserved and cannot be included in area cladograms, the incompleteness of the fossil record can lead to artificial incongruences in phylogenetic paleobiogeographic investigations [37]. Taphonomic bias, which occurs when younger artefacts are overrepresented in the archaeological record relative to older ones because of damaging processes like weathering and erosion, is another drawback of fossils that may cause errors in our understanding of the biodiversity and distribution of ancient times [38].

Advancements in molecular biology have contributed to our comprehension of the evolutionary lineage of humans. Molecules such as DNA and proteins contain ancestral information, facilitating the reconstruction of previously unknown evolutionary events and confirming established views. In the current era, groundbreaking findings about human evolution are just as likely to emerge from a genetics lab as from the East African Rift Valley, indicating a paradigm shift in the techniques used to unravel our ancestry [39]. Once the ability to quantify genetic diversity in humans emerged, researchers began utilising this information to explore the relatedness among the populations and their ancestral roots. Significant demographic occurrences, such as population movements, bottlenecks, and expansions, leave lasting imprints on the human genome by modifying gene frequencies. As these imprints get inherited by future generations, the modern human genome is a permanent repository of our evolutionary journey [40]. The revelation that mutations accumulate consistently over time in the genes of all plant and animal lineages has provided a fresh understanding of evolution at both the molecular and organismal levels [41]. As protein and DNA sequences became accessible for various species, biologists began employing the molecular clock as an effective method

to estimate the times when species diverged [42]. This approach relies on the relatively consistent rates at which mutations occur and alter DNA across various lineages of organisms over extended durations. This allows investigators to establish a clocklike relationship between mutations and the passage of time. This molecular clock is calibrated using a select few fossils with precise dating, enabling the estimation of time elapsed since the divergence of specific groups of living species from their common ancestors and subsequently allowing scientists to estimate the divergence dates for numerous other species using molecular disparities [41]. Exploring connections between fossil and molecular evidence is essential for constructing a comprehensive story of human evolution.

The interdisciplinary approach integrating fossil evidence with molecular biology has significantly advanced our understanding of human evolution. Fossils, as tangible historical records, are the gateway to the past, allowing researchers to trace the evolutionary journey of early hominins and the emergence of our genus, Homo. Morphological analyses of critical species, such as *Australopithecus afarensis* and *Homo habilis*, throw light on essential adaptations like bipedality, tool use, and cognitive abilities. Meanwhile, molecular biology, with its molecular clock and genetic diversity analyses, has revolutionised our ability to explore the broader evolutionary landscape of Homo sapiens. The collaboration between these two methodologies has refined existing knowledge and opened new avenues, emphasising the dynamic interplay between physical and genetic aspects of our ancestral history. As we continue to understand human evolution, this synergistic method remains essential, providing a more comprehensive approach that connects the molecular imprints in our contemporary genome with the fossilised remnants of our past.

1.4 Human Migration and Genetic Diversity

Human migration significantly influences the intricate scheme of human history, culture, and social systems. The dissemination of ideas, the genesis of different groups, and advancing civilisations have all benefited from population movement across geographic borders. A multidisciplinary approach is necessary to fully comprehend the complexity of human migration, using knowledge from anthropology, archaeology, genetics, and history. Genetic variation, entwined with human population movements, is a molecular history that provides unique perspectives into our common ancestry and demographic past. The genetic variation patterns within and between populations offer evidence of the dynamic interactions underpinning migration, adaption, and our species' evolutionary history. Despite the consensus that modern humans evolved in Africa, there is still disagreement on when and how they migrated into Eurasia. The disagreement is because of few fossils and a scant amount of ancient DNA available [42]. Human migration history begins in prehistoric times when the ancestors of modern people left native Africa. According to the conventional view, Homo sapiens invaded the world by land, first migrating from Africa into the Middle East and then northward along the Levantine corridor. From there, they travelled to Europe and then over the middle of Asia to reach India and the Far East [22]. After that, they created the boats to go to Australia and adapted to live along the coast [22]. An alternative theory, that has gained attention lately, holds that early Homo sapiens may have migrated into the Levant through a land bridge exposed by a decline in sea level after leaving Africa using a southern coastal route along the Red Sea coast [43–45]. Multiple factors most likely drove this expansion. Assuming the primitive foraging abilities that Homo erectus groups may have had access to, it would not have taken long for their home regions to grow overcrowded to the extent where population pressure forced migration [46]. A second significant aspect would have been an extensive change in the climate. During the Pleistocene era, when Homo erectus existed, the globe underwent several minor changes, as well as four significant glacial advances that significantly altered habitats and climate regimes. A more

habitable desert stretched from the eastern Sahara to the Negev and northern Arabia because of the glaciers that covered most of north Europe and Asia, pushing climatic zones southward. During the warmer interstadial times, areas like southern Europe and the Caucasus would have been desirable locations. During the glacial maximum, sea levels dropped, creating the land bridges people used to travel from southeast Asia to the enormous landmass that would eventually become the Indonesian archipelago [46].

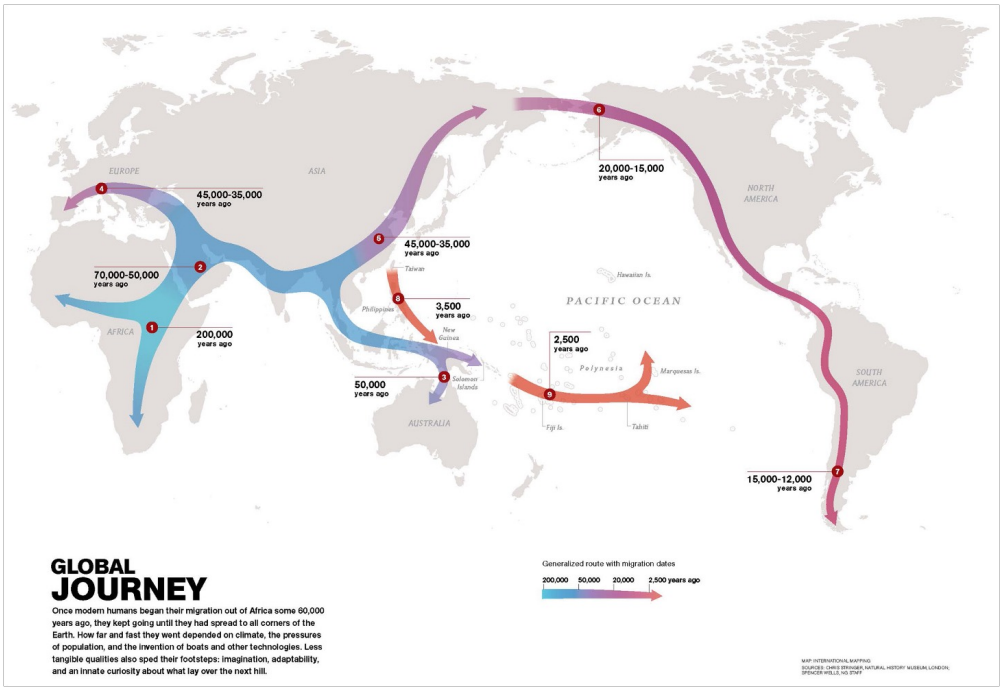


FIGURE 1.1: Migrations of early humans across the world. Photo courtesy of national geographic society, credit to international mapping. sources, Chris Stringer, 2010.

Despite no unanimous agreement on the precise migration routes, it is generally agreed that the migration of anatomically modern humans from Africa marks a significant turning point in human history. This migration event is considered unparalleled, significantly impacting the genetic diversity patterns found in contemporary human populations [47]. Human populations are dynamic, with rapid growth, migration, and subdivision that puts the classical mutation/genetic drift equilibrium model under scrutiny [48, 49]. This is demonstrated by the variations in allele frequencies affected by the adequate population size (Ne) and the magnitude of the migration rate (m), measured

by the index Nem. Smaller Nem values increase vulnerability to allele frequency variations, which frequently results in genetic divergence from ancestral populations. In contrast, significant Nem values act as a buffer against genetic drift, highlighting the impact of migration on genetic diversity [47]. The pattern of variation in the genome among individuals can provide information about population history, including migrations, colonisations, and expansions [50]. Investigations on genetic variation by analysing the genome of modern populations have helped reconstruct recent migration events, such as the colonisation of the New World, the Pacific, eastern Asia, central Asia, western Asia, and Europe [47]. Although genome analysis of modern humans offers the advantage of simple, straightforward sample collection, its disadvantage lies in the requirement of statistical methods of inference to determine the processes responsible for the observed patterns of diversity [51]. Another challenge in determining recent historical events from patterns of genetic diversity in the present day is that they are layered upon a series of ancient events [51]. These include the African origin of modern humans approximately 150,000 years ago, resulting in significant genetic diversity within Africa [52]. Further, successive founder events during the migration out of Africa and the subsequent settlement of the Old and New Worlds led to a gradient of decreasing genetic diversity as the distance from East Africa increased [53, 54]. Notwithstanding these complexities, research on genetic variety has influenced how we comprehend population dynamics and early human history.

1.5 Determinants of Genetic Diversity

Genetic diversity measures the genetic differences among individuals within a population and is a vital research parameter in several areas of biology [55]. Genetic diversity among humans provides insight into the long-term evolution of the human lineage and the demographic history of human populations [55]. The scientific discipline that studies genetic diversity within and between populations and the evolutionary processes that account for this variation is known as population genetics. This study area finds its base

in the Hardy-Weinberg law, which remains relevant in scenarios involving large populations, random mating, and negligible levels of mutation, selection, and migration[56]. Genetic variation must exist within and across populations for that population to evolve. In a population, allele frequencies are a measure of genetic diversity. The change in allele frequency over time can result from either genetic drift or the introduction of new mutations into the population. Genetic variations can arise through natural selection, genetic drift, migration, mutation, and other mechanisms.

1.5.1 Natural selection

Although Alfred Wallace and Charles Darwin contributed to the exploration of natural selection [57], Darwin is the leading proponent of the theory, and his most known work on the subject is 'On the Origin of Species' [58]. New forms naturally arise because of mutation, which may or may not increase an individual's fitness. An individual's alleles will be more common if their fitness results in a reproductive benefit. The individual's alleles are chosen in this manner. This is referred to as natural selection in a Darwinian context. When there is enough genetic variation, time, and population size to overcome genetic drift, natural selection produces traits that maximise fitness [59].

1.5.2 Genetic Drift

Since a small population size cannot accurately represent all alleles within a community, allele frequencies in small populations typically differ from those in larger populations. With wide-ranging implications ranging from speciation to species conservation initiatives, genetic drift is a fundamental evolutionary phenomenon that outlines random changes in allelic frequencies [60]. Genetic drift happens when a population is too small, leading to a random increase or decrease in the frequency of a particular allele, which may lead to deviation from Hardy-Weinberg equilibrium (HWE).

1.5.3 Mutations

Natural selection operates on heritable phenotypic variation, primarily derived from permanent modifications of the genetic material known as mutations (Tenaillon & Matic, 2020). Mutations can be broadly classified into three groups based on how they affect fitness: the ”good” or advantageous mutations that improve fitness, the ”bad” or deleterious mutations that decrease fitness, and the “indifferent” or neutral mutations that have too little effect to be affected by selection [61]. Mutations can be categorised as lethal, deleterious, neutral, or beneficial in each environment based on how they affect fitness. The relatively common detrimental and lethal mutations are useful for studying gene functions or the molecular causes of diseases, and rare beneficial mutations are vital to understanding adaptation. Generally, neutral mutations are employed as markers in phylogenies or genealogies [62].

1.5.4 Gene flow

Gene flow refers to all processes leading to the transfer of genes from one population to another. Sometimes, migration and gene flow are used interchangeably, which is accurate when the mechanism of gene flow involves migration between established populations. The movement of gametes, the extinction and recolonisation of entire populations, or the movement of extranuclear DNA segments like viruses, plasmids, and mitochondria are other causes of gene flow [63].

1.6 Genetic Diversity and the Indian Subcontinent

The genetic diversity within the Indian population is of great interest due to the diverse ancestral components, intricate social stratification, patterns of endogamy, and complex admixture [64]. These factors have led to distinct and exceptional genetic variations that

could be accountable for a range of prevalent diseases in India's indigenous and migratory populations [65]. The Indian subcontinent, a region of great cultural and biological diversity, is between 8 ° N to 37° N latitude and 68° E to 97° E longitude. It is now the country with the largest population in the world, with an increasing population of more than 1.43 billion people. The Indian subcontinent covers 3,200 km from north to south and comprises a diverse geography that includes the Himalayas, the Thar desert, plateaux, and rain forests. The country has a long history of migrations and invasions, a rigid caste system promoting endogamy, and almost 800 languages throughout the nation [66]. The influence of historic and prehistoric human migration in India is evident from its enormous social and cultural diversity [67]. Analysing biological findings within the framework of sociocultural history is critical, as cultural factors significantly impact the way gene flows among populations and, in turn, shape patterns of biological affinities [68]. The early stages of Indian historical discourse have been marked by conflicting theories regarding the origins of the foundational cultures of India and the peoples associated with them. These debates primarily revolve around the origins of the Aryans and, subsequently, the Dravidians [69]. One perspective posits that the Aryans originated from Central Asia, sharing cultural affinities with Western Eurasia, who migrated and settled in northwestern India before spreading throughout the subcontinent. Conversely, an alternative view suggests that the Aryans were indigenous to India and, as some argue, disseminated their cultural influence from India westward. Genetic analyses have been utilised to elucidate the identities of various population clusters and offer insights into these inquiries. However, the challenge persists as no such groups have endured as distinct entities from that era [70].

Human footprints in India date back to 400,000 to 200,000 B.C. [71], with humans expanding across the subcontinent during the middle Paleolithic era (50,000–20,000 B.C.). Additionally, evidence of several Neolithic settlements dating from 7000 to 5000 B.C. can be found in India [68]. The Indus Valley civilisations (ancient civilisations in present-day Pakistan and Northwest India) flourished from roughly 3000 B.C. to about

1500 B.C. They had strong trading networks extending to Mesopotamia and the Persian Gulf region. Although the exact causes of the collapse of the Mohenjo Daro and Harappan civilisations (prominent sites of the Indus Valley civilisation) are unknown, the earliest arrival of Indo-Aryans (2000–1400 B.C.), who migrated from central Asia via the Iranian plateau, coincides with this period [65]. The initial hypothesis attributing the decline of Harappan cities to Aryan invasions has been subject to scrutiny due to insufficient evidence, leading researchers to reconsider environmental factors as potential causes. Substantial archaeological investigations in the Indus Plain and adjacent regions uncovered numerous settlements representing diverse archaeological cultures contemporary with the Harappan cities. Some of these settlements persisted into the post-Harappan period. This comprehensive archaeological evidence has prompted a reassessment of the mechanisms underlying the spread of Indo-Aryan languages in northern India. It appears increasingly plausible that there were minor-scale migrations into the northwestern region, with settlements established close to pre-existing settlements, possibly integrating with them over time [70].

The Greeks (400–200 B.C.), Sakas (200 B.C.), Kushanas (100 A.D.), Huns (200–500 A.D.), and Arabs (800 A.D.) were among the prominent groups that arrived in India and progressively assimilated with the native populace [65]. In the medieval era, Turks and Afghans made significant invasions into the northern Indian subcontinent. Portuguese conquest of Goa in 1510 A.D. marked the start of European colonisation attempts in the sixteenth century. The British ascended to supremacy in India due to the differences between the Portuguese and Dutch, ultimately establishing the British Empire, encompassing a significant portion of India and a portion of Myanmar (Burma). Large Anglo-Indian and Indo-Portuguese communities were formed in Goa and Bombay due to the admixture of European genes with indigenous populations [72, 73]. During the extended period of cultural interaction and political uncertainty in northern India, the region of southern India predominantly experienced autonomous development. The main reason why inhabitants of south India are genetically different from those in north India is most likely because Dravidian-speaking Southern India was comparatively shielded

from Indo-Aryan influence. The Indo-Aryan speakers dispersed in all directions after passing through the western and northwest corridors. The prevalent, well-established Dravidian language in southern India was a barrier to the spread of Indo-Aryan influence further south. As a result, there was less genetic mixing between the Indo-Aryans and the indigenous populations in southern India compared to the northern regions where the influence of Indo-Aryans was more pronounced. Thus, the admixture of various genetic factors with indigenous genetic stocks in south India is reduced compared to northern India [68]. The multitude of migration events and subsequent settlements have substantially impacted the demographic composition of the Indian subcontinent. This historical process has contributed to the formation of diverse morphological, linguistic, religious, and socio-cultural groups within the Indian subcontinent, resulting in a rich genetic diversity and variation reservoir.

India's population can be broadly divided into four morphological groups: Australoid, Mongoloid, Negrito, and Caucasoid [74]. Recent research has identified four primary ancestral lineages in populations from mainland India: Ancestral South-Indian, Ancestral North-Indian, Ancestral Austro-Asiatic, and Ancestral Tibeto-Burman [64]. The population of western India is described to have Australoid or proto-Australoid elements with Indo-Aryan racial admixture [75]. While the genetic makeup of Eastern India is a mixture of Caucasoid and Mongoloid components [76]. The North Indian gene pool has a significant amount of genetic admixture from other racial groups, with profound oriental influence [77], while on its kind, the Southern part of India, being one of the oldest geophysical regions, is mainly occupied by Dravidian language-speaking people [78]. Based on language, the people of India are divided into four prominent language families: Austroasiatic, Dravidian, Tibeto-Burman, and Indo-European. Most people speak Indo-European languages, with Dravidian languages being most common in the country's southern parts. Whereas Tibeto-Burman speakers are concentrated in the northeastern states and the foothills of the Himalayas, Austroasiatic speakers are mainly found in central and eastern regions [65].

Geographical location and genetic structure are tightly correlated in other world regions. However, understanding the Indian population structure is made more difficult by the strict endogamy imposed by the caste system and the various languages spoken within the country [66]. Numerous religious groups shape the population's social structure in the Indian subcontinent. Most people in India practice Hinduism and the Hindu caste system. This system has four castes, ranked from high to low: the Brahmin (priestly class), Kshatriya (warrior class), Vysya (trader class), and Shudra (labour class). Other common religions practised in the country are Islam, Christianity, Sikh, Jainism, Buddhism, etc. Socio-cultural and biological characteristics determine many intrinsic ethnic groups within each language and religious community. These ethnic groups fall into broad categories - castes and scheduled castes, tribes and scheduled tribes, and other communities [72]. With such a diversified population, India holds tremendous genetic diversity among different regions. The abundance of an extensive genetic repertoire is an excellent resource for human diversity investigations and is highly informative for studying genetic susceptibility to various diseases [79, 80].

1.7 Measuring Genetic Diversity

One can deduce how selection and historical demographic events have affected genetic variation by looking at patterns of genetic polymorphisms. The development of blood grouping techniques and electrophoretic methods to identify polymorphisms at loci for red-cell enzymes and serum proteins provided researchers with additional opportunities that offered insight into the biological structure of Indian populations. Further, the advancement of molecular biology has tremendously contributed to understanding population structure. From a population geneticists' standpoint, the population can be defined as "a reproductive community of individuals who share in a common gene pool" [81]. The genes inherited by members of a population, who in turn inherited them from their parents within the same population, construct the genetic makeup of that population. Allele frequencies in a population are a measure of genetic diversity. Examining the

frequency of different alleles among the population allows for quantifying this genetic makeup, or gene pool. An allele's relative frequency at a genetic locus in a population is called its allele frequency. Estimating the incidence of alleles across various loci is a crucial step in comprehending the genetic structure of human populations. It is calculated by dividing the number of copies of all alleles at a particular genetic locus in a population by the number of times the allele under investigation is observed. In an individual, two alleles can be found for a specific gene located in the corresponding gene locus regions in a pair of chromosomes, one coming from the father and another from the mother. In nature, allele frequency is altered as individuals of a population reproduce, die away, and migrate between populations. Migration can alter allele frequency through a process known as gene flow (defined as the slow spreading of genes across a geographic or cultural barrier). When groups or populations (and thus their genomes) move from place to place and mate with the indigenous people, allele frequencies in both the migrant and the native populations may change.

Considering the population composition of the Indian subcontinent, gene flow due to migration might be the fastest and most significant cause of disbalance in allele frequency [72]. Genetic relationships between subpopulations increase through admixture or gene exchange. Cultural and linguistic constraints and geographic isolation are the main challenges to admixture. Therefore, it is expected that languages, cultures, and genes will all evolve simultaneously, especially for genes not subjected to different natural selection forces. This anticipated relationship suggests examining genetic diversities and affinities in tandem with cultural and linguistic histories. In general, the genetic distance between the two populations increases with the length of time that the two populations are separated. Thus, genetic distance is a valuable measure for dating evolutionary history [68]. Similarly, various statistical methods can be employed to analyse population differences, including heritability, population mean, effective population size, variance, gene frequency changes, genetic distance, and phylogenetic relationships. Additionally, determining the allele frequencies of candidate gene variants and single-nucleotide polymorphisms serves multiple purposes, including aiding in

data interpretation and research design, identifying genetic associations with diseases or health-related traits, estimating the prevalence of drug-resistant or disease-susceptible individuals within a population, and facilitating anthropological and evolutionary investigations.

1.8 Use of Genetic markers to understand Genetic Diversity

Genetic markers are independently segregating genetic elements used to classify populations based on their presence, absence, or variations in frequency between populations [82]. They can offer valuable information by detecting genetic variants in several areas, such as population structure, gene flow levels, phylogenetic relationships, historical biogeography patterns, and determining parentage and relatedness [83]. The following section describes several genetic markers frequently utilised to gain insights into the history of human evolution, trace ancient migration patterns, and aid in forensic investigations and human health studies.

1.8.1 mtDNA

Mitochondria contains its genetic material, the ancient remains of bacterial genomes co-evolved with proto-eukaryotic cells' nuclear DNA [84]. Human mitochondrial DNA (mtDNA) has characteristics such as high copy number, maternal inheritance, high mutation rate, and an absence of recombination, which makes it the preferred molecule for research on the evolution and history of the human population [85]. Most notably, as mtDNA haplogroups exhibit geographical distinctiveness, they are critical in population genetics. The mitochondrial haplogroups are collections of similar haplotypes defined by combinations of single nuclear polymorphisms (SNPs) found in mtDNA [86]. These haplogroups are created by gradually accumulating mutations along maternal lineages

[87]. Haplogroups correlate with the geographic origins of populations traced through the maternal lineages [88]. With the unique distribution of mtDNA sequences into haplogroups distinctive to different continents, it is possible to assign unique European, African, Asian/Native American, or other haplogroups to mitochondrial DNA (mtDNA) sequences [89]. The importance of mtDNA in population genetics is demonstrated by the fact that researchers used mtDNA to conclude that all contemporary human populations share an African genetic ancestor who lived approximately 200 thousand years ago[25].

1.8.2 Y-Chromosome

The Y chromosome inherited from the father harbours the largest non-recombining block of nucleotides (NRY) in the human genome, which was estimated to be 50 million base pairs (bp) in length and constituting 95% of the entire Y chromosome [90, 91]. The NRY contains various categories of markers, such as microsatellites (STRs), minisatellites, insertion/deletion polymorphisms (indels), and single nucleotide polymorphisms (SNPs) [92]. Due to their high heterozygosity, STRs are informative for forensic and paternity analysis and for determining affinities among closely related populations. Polymorphisms with lower probabilities of back and parallel mutation, such as SNPs and small indels with mutation rates of approximately $2-4\text{x}10-8$/site/generation, are ideal for studying ancient relationships among populations and reconstructing ancestral relationships among NRY lineages because of their determinable ancestral states and reduced homoplasy levels [93].

1.8.3 Autosomal STRs

Short tandem repeats (STRs) are regions of 2 to 7 bp length tandemly repeated DNA segments occurring throughout the genome varying in length due to insertion, deletion, or mutation with a core repeated DNA sequence [94]. The STRs are classified based

on the repeated pattern, dividing them into several categories; (1) simple repeats that possess repeat units of identical length and sequence, (2) compound repeats comprising two or more adjacent simple repeats, and (3) complex repeats that may contain several repeat blocks of variable unit length. STRs can be used to evaluate biological diversity at different levels of biological organisation and to reconstruct the history of a species' migration and evolution [95]. An analysis of the size difference between two STR alleles can provide important information; a more significant variation suggests more mutation events. Hence, newly formed mutants continue to have a relationship with the ancestral allele [96]. The length discrepancies between the alleles reveal evolutionary information about their relationship [97] that can be utilised in understanding human history.

1.8.4 Major Histocompatibility Complex

The major histocompatibility complex (MHC) refers to a cluster of genes found in both animals and humans, which encode diverse cell surface markers, antigen-presenting molecules, and other proteins crucial for immune functions. In humans, the term human leukocyte antigen (HLA) complex is used interchangeably with the human MHC. The HLA alleles vary among individuals and dictate the success of transplantation procedures. Hence, it is generally known as a major histocompatibility complex (MHC). Due to its highly polymorphic nature, the HLA system plays a significant role in genetic differentiation studies, disease association studies, and inferring phylogenetic relationships [98]. The generation of variability in this region results from evolutionary forces like mutation, recombination, and gene conversion [99]. The HLA system is also an incredible asset for population studies. Linkage disequilibrium estimates and haplotype data of the HLA system are paramount in population studies. Both measures are practical when comparing various populations, as each endogamous caste group, primary group, or even regional group has its characteristic haplotype profile and linkage disequilibrium estimates [100].

1.9 The rationale for the current investigation

The Indian subcontinent has served as a melting pot of diverse ethnic groups with distinct cultural practices and genetic heritage. People with diverse ancestral backgrounds originating from Eurasia and Southeast Asia constitute the Indian population. India's North and Northeastern regions stand out among this melting pot of ethnicities as an exciting area, home to populations with distinct socio-ethnic identities and genetic signatures. A significant and integral part of this vibrant region is the population that shares the Nepali lingua franca. As of 1978, three to three and a half million Nepalese (also known as Gorkhalis) were estimated to be inhabitants of northeastern India [101]. The Sub-Himalayan region of West Bengal, extending from $27^0$2'N latitude to $88^0$15'E longitude, showcases geographic continuity with segments of Northeast India, primarily due to the shared Himalayan foothills. This region notably hosts a substantial Gorkhali population, contributing significantly to its demographic and cultural landscape. The region shares borders with Nepal, Sikkim, and Bhutan, covering an area of 3149 km2 with a population of approximately 1.847 million inhabitants, predominantly consisting of the Gorkha (also known as Gurkha), Lepcha, and Bhutia communities [102]. The Gorkhalis or Gorkhas comprise diverse social and linguistic groups, including Rai, Magar, Limbu, Tamang, Newar, Bahun, Chettri, Kami, and so on [102]. Historically, the word 'Gorkha' is associated with the Khasas originating in northern India [103]. Subsequently, following the establishment of the Gorkha dynasty by Drabya Shah, the term 'Gorkha' became associated with the inhabitants of the principality under Gorkha rule [104]. However, in contemporary usage, 'Gorkha' denotes a specific community or distinguishes Indian citizens of Nepali ethnicity from Nepali citizens [104]. The two main racial groups that make up the Gorkha population are the Mongoloid group, which includes Limbu, Tamang, Rai, and Magar, and the Caucasoid group, which provides for Bahun, Chettri, Kami, Damai, and Sarki [105]. Linguistically, Caucasoid has no other languages except the Indo-Aryan Nepali language, while the mongoloids have many

Tibeto-Burman languages [105]. Based on the socio-ritual level, the Gorkha community encompasses three distinct ethnic groups: (i) The 'Kiratis,' comprising the Rai, Magar, Limbu, Lepcha, Tamang etc., (ii) The 'Newaris' or 'Newars,' (iii) The 'Tagadharis,' representing the Nepali equivalent of Indian Hindus, including Bahuns, Chettri, Kami, Damai, and Sarki [103].

Despite living close to one another, the Gorkha community shows a significant genetic difference, emphasising the need for research among individual ethnic Gorkha populations. Researchers have conducted a few genetic studies [106–108] to investigate the patterns of genetic diversity using different genetic markers in the Nepali/Gorkha population of India. However, most studies have sampled individuals from Tibeto-Burman mongoloids groups, and some studies have generalised that the Gorkha population has mongoloid descent, which may not be true given that the Gorkha population consists of populations belonging to both Caucasoid and Mongoloid groups. Moreover, limited studies have been conducted on Indo-Aryan Gorkha populations from the Gorkha community to understand their genetic diversity. Therefore, as one of the representative Caucasoid, Indo-Aryan population, the Kami population belonging to the Indian Gorkha community have been the subject of the current investigation.

This study explores the distribution of HLA and ABO blood groups among the Kami population. Investigating these genetic markers among the Kami population may help to uncover genetic variations and traits related to HLA and ABO gene loci that extend beyond the confines of a particular ethnic or linguistic group. This investigation may not only provide a deeper insight into the HLA profile and ABO blood group distribution among the Kami population but also help to understand the genetic affinity of Kamis with other populations. Additionally, as HLA and ABO systems are of paramount importance in the context of organ transplantation, blood transfusion, and disease association, the study holds significant potential to provide valuable insights for developing healthcare strategies specific to the needs of the studied population.

1.10 Objectives of the Study

1. To study the allelic polymorphism of HLA at A, B, and DR loci among the Indo-Aryan Kami population.

2. To study the two and three locus haplotypes of HLA at A, B, and DR loci among the Kami population.

3. To analyse the genetic affinity of the Kami population with the other Indian and World populations based on HLA.

4. To study the presence of proven disease-associated HLA alleles among the Kami population and correlate with the disease condition (if any).

5. To determine the allelic frequency and distribution of ABO and Rh system among the Kami population.

Chapter 2

Literature Review

2.1 Literature and Review

The Human Genome Project, which aimed to decipher the basic chemical building blocks that comprise the entire human genome, was completed in 2003 [109]. The Major Histocompatibility Complex (MHC), also known as the Human Leukocyte Antigen (HLA) Gene Complex in humans, is an intriguing region of the genome that stands out for having a wide range of genes and a notable clustering of genes involved in immune response processes [110]. This 4 Mb stretch of DNA, constituting 0.1% of the entire genome, located on the short arm of chromosome 6, is among the most extensively studied regions within the genome [111]. Hence, in this section, a review of HLA genetics and other relevant literature concerning the study of HLA gene complex and ABO system in the Kami population has been conducted under separate headings.

2.2 The Kami population

The Kami population are an Indo-Aryan-speaking Hindu socio-ethnic caste group within the Gorkha community, primarily concentrated in the sub-Himalayan region of West

Bengal and sporadically distributed in other Indian states. The precise origins of the Kami population remain obscure. However, historical references suggest their ancestors migrated from various regions of India, such as Punjab, Rajputana, Kashmir, and neighbouring areas, between the 10th and 15th centuries AD, settling in Nepal and subsequently to the Sub Himalayan region of West Bengal [112]. The traditional occupations of Kami centre around craftsmanship, particularly in metalwork, crafting the renowned 'Khukuri' knives used by the Gorkha army. Additionally, they are skilled in creating the traditional Nepali drum called 'Madal' and are practitioners of 'Maruni Nritya,' a folk dance of the Gorkha community. The Kamis predominantly practice monogamy, although historical instances of polygamy have been noted. They are organised into 54 exogamous clans, and inter-clan marriages are prevalent, while marriages outside the community are infrequent [112]. As per the 2011 census, the Kami population in West Bengal is approximately 52,178 and belongs to scheduled caste communities as per the Indian constitution.

2.3 Human Leukocyte Antigen Gene Complex

Human leukocyte antigens (HLA), responsible for presenting antigens to T lymphocytes, exhibit significant variability and are encoded by the human major histocompatibility complex (MHC) [113]. The assemblage of essential components of the MHC region can be traced back to the time before the origin of vertebrates, almost 525 million years ago [114]. It has been established that the closest extant invertebrate to vertebrates, Amphioxus, a member of the cephalochordate order, possesses prototypic MHC [115] . The modern Major Histocompatibility Complex (MHC) region is believed to have originated from ancient chromosomal duplications that occurred in the common ancestor of jawed vertebrates [116, 117]. The high degree of polymorphism in MHC proteins likely evolved as a response to counter the evasion tactics employed by pathogens [118].

The HLA system is an essential component of the human immune system and plays a fundamental role in regulating cell-mediated adaptive immune responses, innate immune responses, and the recognition of self and non-self [119, 120]. The HLA-encoded molecules capture self or foreign peptides, presenting them on cell surfaces for recognition by T cell receptors (TCRs) and NK cell receptors. These interactions guide T cell and NK cell development and activate memory/effector T cells, influencing immunity, cancer susceptibility, and autoimmunity [121]. The products of the human leukocyte gene complex substantially influence the success of transplantation, antigen presentation to CD4+ or CD8+ T cells, and susceptibility to or protection from certain diseases.

Since its discovery over 50 years ago, the human leukocyte antigen located on the petite arm of chromosome 6 has been a focal point in human genetic research [122]. Although their primary function involves presenting exogenous and endogenous peptides, they were initially termed 'antigens' because they engaged with T-cells during transplant rejection [123].

2.4 Genetics of Major Histocompatibility Complex

The MHC genomic region comprises genes, retrotransposons, transposons, regulatory elements, pseudogenes, and a few unidentified sequences, representing one of the most densely populated and well-defined gene regions [124]. The HLA gene complex, spanning around 4,000 kb, is on the short arm of chromosome 6p21.1-21.3. It originated through repetitive gene duplication and conversion processes throughout evolution. The HLA system encompasses nearly 27,000 alleles distributed across three distinct classes of genes: Class I, II, and III [125]. Spanning a length of 3.6 megabases, the three different segments (from centromere to telomere) are responsible for encoding class II antigens (1.1 megabases), class III antigens (0.7 megabases), and class I antigens (1.8 megabases) in sequential order [126]. The HLA region covers 0.12% of the human genome [127]. Contrary to expectations, within the approximately 3.6-megabase HLA

region, the segments responsible for encoding the class I and II genes exhibit significant polymorphism. In contrast, the approximately 1.7-megabase segments comprising at least 115 other genes remain significantly conserved [128]. The HLA super-locus contains twelve classical class I and class II genes responsible for coding antigen-presenting HLA proteins that engage with T-cell receptors to differentiate between self and non-self-peptides during the host's immune response [129].

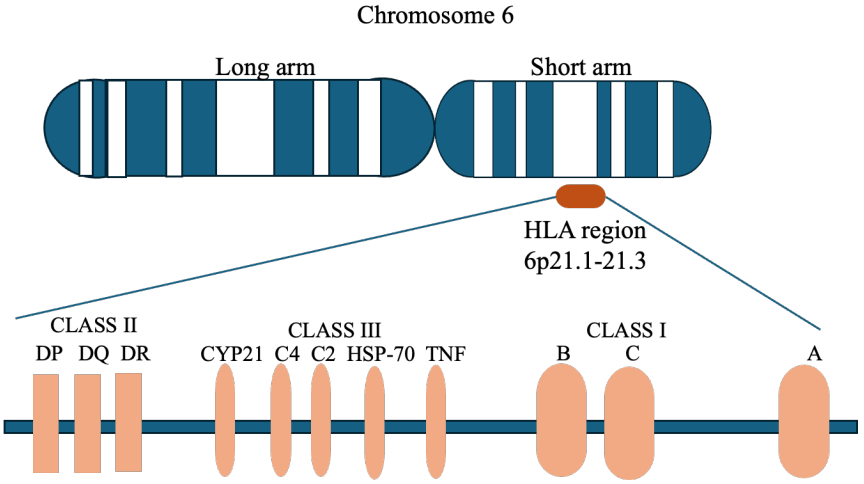


FIGURE 2.1: Schematic representation of classical Class I, Class II and Class III HLA genetic loci at chromosome 6 p21.1–21.3.

The classical HLA-Class I genes, namely HLA-A, HLA-B, and HLA-C, are integral components of the human leukocyte antigen class I (Class I HLA) gene group. Each of these genes displays polymorphism, characterised by many allelic variations. Based on data from April 2024, the IPD-IMGT/HLA database has carefully compiled a list of over 27,301 alleles for these well-known HLA-Class I genes. In contrast, non-classical genes within the HLA-Class I cluster, such as HLA-E, HLA-F, HLA-G, HFE, and 12 pseudogenes, exhibit comparatively lower levels of polymorphism [130]. The HLA Class II cluster encompasses classical class II genes, including HLA-DP, HLA-DQ, and HLA-DR, along with pseudogenes and non-classical class II genes such as HLA-DM and HLA-DO [130]. The IPD-IMGT/HLA Database has recorded over 11,674 alleles for these classical HLA-II genes as of April 2024. The proteins encoded by the Class I HLA genes are ubiquitously expressed and present peptides to CD8+ T cells,

while the Class II HLA products are heterodimers encoded by HLA-DRA/DRB1/3/4/5, -DQA1/DQB1, -DPA1/DPB1, and are primarily expressed on Antigen-presenting cells (APCs) and present peptides to CD4+ T cells [131]. While the HLA class I and class II genetic regions encode highly variable gene complexes, the class III region harbours a diverse array of non-HLA genes involved in functions such as stress response (HSPA1A, HSPA1B, and HSPA1L), complement cascade (C4A, C4B, C2, CFB), immune regulation (NFKBIL1, FXBPL, and DDX39B), inflammation (LTA, LTB, LST1, ABCF1, AIF1, NCR3, and TNF), leukocyte maturation (LY6G5B, LY6GSC, LY6G6D, LY6G6E, and LY6G6C), and the control of T cell development and differentiation (BTNL2) [132].

Since the HLA region harbours many immunologically essential genes, susceptibility to or protection from various diseases is associated with the HLA region in humans [133]. This tiny segment of the human genome has been reported to be associated with more than 100 different diseases, including common diseases such as diabetes, rheumatoid arthritis, psoriasis, asthma, and various other autoimmune disorders [134]. HLA allotypes are also implicated in unwanted immune reactions, including drug hypersensitivity syndrome, in which small therapeutic drugs interact with antigenic peptides to drive T-cell responses restricted by host HLA [135]. Polymorphisms in HLA class I genes and class II genes primarily result in variant amino acids within the antigen-binding cleft of expressed class I molecules and class II molecules, influencing the range of peptides that can be captured and presented to CD8+ T and CD4+ T cells, respectively [136]. The polymorphism in HLA is attributed to selection pressures from pathogens on HLA allomorphs, which exhibit varying antigen-binding capacities [137]. This extensive HLA polymorphism not only enhances the likelihood of heterozygosity in individuals, enabling them to carry two distinct HLA alleles for each HLA locus but also potentially doubles their antigen-presenting potential [138]. Consequently, this broadens the overall HLA repertoire to deal with diverse pathogens within a species.

The HLA system is also an incredible asset for population studies. Linkage disequilibrium estimates and haplotype data from the HLA system are paramount in population

studies. These measures are practical when comparing various populations, as each endogamous caste group, primary group, or even regional group has its characteristic haplotype profile [139] and linkage disequilibrium estimates. Owing to its highly polymorphic nature, the HLA system plays a significant role in genetic differentiation studies [140]. The generation of variability in this region results from evolutionary forces like mutation, recombination, and gene conversion [141]. Therefore, detailed studies of the HLA region are necessary to understand the evolution of disease-predisposing genes, epitopes, and complex multigene families [141]. This remarkable pattern of HLA polymorphism, characterised by a multitude of alleles at intermediate frequencies, suggests a form of balancing selection rather than selective neutrality, highlighting the complex interplay of factors shaping HLA diversity [142]. Thus, we can conclude that the HLA region is a crucial player in the ever-evolving battle between the immune system and pathogens, demonstrating a balance between diversity and specificity. The journey of understanding the HLA genomic landscape is ongoing, and the more we discover, the more we appreciate the complexity and adaptability of our immune system.

2.5 HLA Nomenclature

The WHO Nomenclature Committee for Factors of the HLA System oversees the formal naming of HLA alleles. It disseminates this information through the Immuno Polymorphism Database-International ImMunoGeneTics project/HLA (IPD-IMGT/HLA) database (https: //www.ebi.ac.uk/ ipd/imgt/hla/) and the HLA Nomenclature website (http://hla.alleles.org /nomenclature/naming.html) [143] [144]. The current HLA nomenclature system employs a unique numerical code consisting of up to four sets of digits separated by colons (Fig. 2.2). The HLA prefix indicates the human MHC gene complex, followed by the specific HLA genomic region. The first two digits (field 1) denote the allele group or family. The second field specifies the particular HLA allele (HLA

protein). The third field names alleles that differ solely by synonymous nucleotide substitutions within the coding region. The fourth field designates alleles that vary by sequence polymorphisms in introns, 3'-untranslated regions, and 5'-untranslated regions. The suffix, represented by a letter, indicates HLA protein expression level changes. For instance, "N" denotes null alleles with no HLA protein expression, while other letters like "L" signify low expression, "S" denotes secreted forms, and "Q" indicates questionable status [120]. This detailed nomenclature system aids in precisely identifying and categorising HLA alleles, facilitating research and clinical applications related to immune responses and transplantation compatibility.

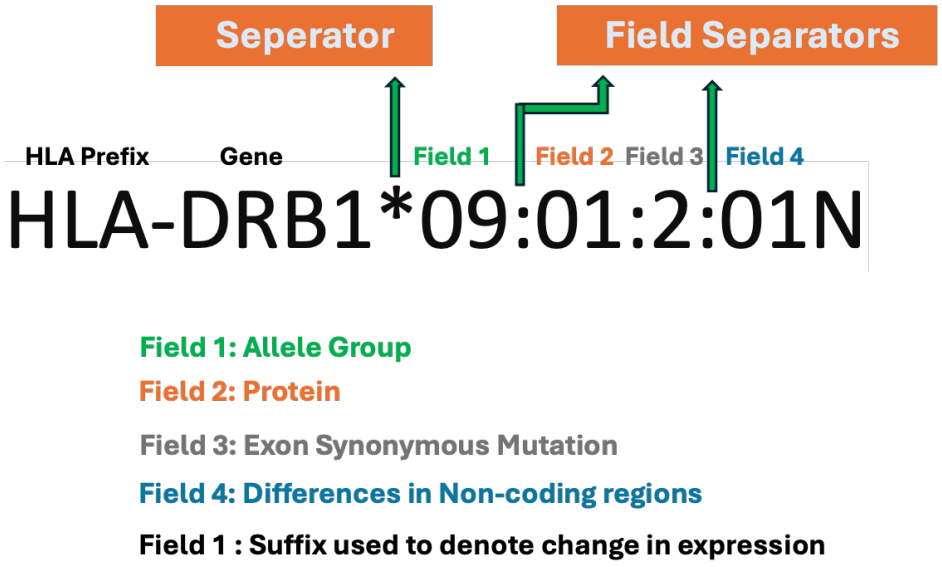


FIGURE 2.2: HLA nomenclature system.

2.6 Allelic polymorphism and Haplotypes at HLA loci in various populations

The remarkable levels of polymorphism, close linkage, a stochastic assortment of alleles, and the persistence of allelic lineages over generations render HLA genetic markers

indispensable in elucidating human history. The critical data about the quantity, structure, and geographical distribution of genetic diversity within HLA genetic markers across diverse populations facilitates the correlation between genetic profiles and historical migration patterns, contributing significantly to understanding population origins [145]. Several DNA-based HLA typing methods have been developed with the advent of polymerase chain reaction (PCR) technology [146]. DNA typing methods that are often used include sequence-specific primers (SSP), sequence-specific oligonucleotide probes (SSOP), sequence-based typing (SBT), and Next Generation sequencing (NGS) [147]. The significant variation observed in HLA frequencies in human populations since the emergence of our species and subsequent dispersal across Africa and various continents accentuates the importance of HLA data analysis at the level of specific populations or ethnic groups [148]. The distributions of HLA alleles and haplotypes and the linkage disequilibrium (LD) between HLA genes manifest distinct patterns of variation among diverse populations and regions. These patterns have served as valuable tools in exploring human evolutionary processes such as gene flow and migration [149]. The study of this genetic diversity is helpful in anthropological science as it helps to define evolution and the routes of the earliest worldwide migration of Homo sapiens [150]. Studies of polymorphism in the human genome have shed light on the relationships between various human groups [151] [152]. In addition, polymorphism analysis is a valuable tool for studying the genetic history of our communities [153].

2.7 HLA diversity among the world population

Various studies have been performed to understand the genetic diversity of the human leukocyte antigen system in diverse populations worldwide. The HLA studies conducted on global populations are summarised in Table 2.1. Williams et al. (2001) identified common HLA-B alleles in diverse populations from 5 continents. HLA-B*5101 (8.4%) and HLA-B*0801 (7.4%) were the most common alleles in the Brazilian population, while HLA-B*3501 (46%) and HLA-B*4005 (15.5%) were the most

common in the Mexicans Series. In Mexican Mestizos, HLA-B*3501 (7.3%) and HLA-B*1801 (6.1%) were the most common. HLA-B*0702 (7.8%) was the most common allele in Cuban Caucasoid, while HLA-B*5301 (10.7%) and HLA-B*3501 (9.5%) were the most common in Cuban Mulatto. HLA-B*4201 (12%) and HLA-B*4403 (10.5%) were the most common alleles in South African Zulu; HLA-B*5101 (17.5%) and HLA-B*0801 (11.5%) were the most common in Omani. Lastly, among Singapore Chinese, the most prevalent alleles were HLA-B*4001 (17.2%), HLA-B*4601 (13.2%), and HLA-B*4001 (17.2%) [154]. In another study, HLA frequencies were studied among the East African population. HLA-A*02:01:01 (10.90%), HLA-B*58:02 (8.79%), and HLA-C*06:02:01 (16.98%) were the most frequently found class I alleles. HLA-DPA1*01:03:01 (40.60%), HLA-DPB1*01:01:01 (23.45%), HLA-DQA1*01:02:01 (31.03%), HLA-DQB1*03:01:01 (21.79%), HLA-DRB1*11:01:02 (11.65%), HLA-DRB3* 02:02:01 (31.65%), HLA-DRB4*01:01:01 (10.50%), and HLA-DRB5*01:01:01 (10.50%) were the most prevalent class II alleles. This population's allele frequency distribution is remarkably similar to that of other sub-Saharan populations, except for the lower frequencies of DQA1*03 (4.79% versus 11.72%) and HLA-A*23 (5.55% versus 11.21%) and the higher frequencies of DPB1*30 (2.26% versus 0.37%) and DRB1*11 (21.51% versus 15.89%) [155]. HLA-A*02 (38.48%), HLA-A*01 (22.92%), HLA-B*51 (20%), HLA-B*35 (19.8%), HLA-CW*04 (23.4%), HLA-CW*07 (19.28%), HLA-DR2 (27.75%), HLA-DR3 (27.25%), HLA-DQ1 (23.25%), and HLA-DQ3 (23.25%) were the most common HLA alleles among Iraqi Arabs [156].

Numerous studies have also been conducted in South Asian countries to understand the diversity of HLA. The study of the frequencies of the HLA-A, -B, and -DRB1 alleles in the Bangladeshi population yielded important insights into the genetic diversity of this population. Among the HLA-A alleles, HLA-A*33 exhibited the highest frequency at 17.02%, followed closely by HLA-A*24 at 16.31% at 15.60%, and HLA-A*02 at 14.54%. With a frequency of 19.5%, HLA-B*15 was found to be the most prevalent allele at the B locus. HLA-B*35 and HLA-B*44 came in second and third, respectively, at 17.02% and 9.22%. Within the DRB1 locus, HLA-DRB1*15 was the

most prevalent allele, accounting for 29.07% of the observed alleles. Other notable alleles included HLA-DRB1*07 at 21.28%, HLA-DRB1*12 at 11.35%, and DRB1*04 at 10.28%. The most common two loci class I haplotypes were HLA-A*33∼B*44 (8.15%), whereas HLA-A*33∼B*44∼DRB1*07 (6.38%) was the most frequent three loci haplotype. The study revealed close associations between Bangladeshis and Indian non-tribal random Dravidians and north Indian Hindus and some relations between Mongolian and Pakistani populations [157]. According to a study, the Pakistani population exhibited the most frequent HLA alleles of HLA-A*02, HLA-B*35, and HLA-CW*07, with frequencies of 19.2%, 13.7%, and 20%. HLA-DRB1 alleles accounted for more than 60% of the subjects, with HLA-DRB1*03, HLA-DRB1*07, HLA-DRB1*11, and HLA-DRB1*15 being the most represented. The most common haplotype was DRB1*15∼DQB1*06. Interestingly, the study found that the HLA class I and II alleles of Pakistanis closely resembled those of European Caucasoids and Orientals [158] . A more specific study among the Gujjar population of Pakistan showed that the frequency of HLA-A*01, HLA-A*02, HLA-A*11, HLA-A*26, and HLA-A*31, HLA-B*08, HLA-B*51, HLA-C*07, and HLA-C*14 was found to be higher. At the DRB1 locus, DRB1*03, DRB1*13, and DRB1*15 were found to be higher, and at the DQB1 locus, the allelic groups DQB1*06 and DQB1*02 were higher. HLA-A*31∼B*51∼DRB1*13 (8.8%) was the most common haplotype observed in Gujjars. The study showed that Pakistani Gujjar is closely related to the Golla tribe from Andhra Pradesh in India [159]. The distribution of high-frequency HLA alleles is noteworthy among Nepalese donors and recipients of renal transplants. A*11 is the most common HLA-A allele, accounting for 34.5% of cases, followed by HLA-A*24 at 17% and HLA-A*33 at 13%. B15 has the highest frequency in the HLA-B group (27%), followed by HLA-B*35 (19%) and HLA-B*40 (10%). The HLA-DRB1 alleles with the highest frequency are DRB1*15 (33.5%), DRB1*12 (21.4%), and DRB1*04 (7.32%) R4. A thorough study of the allele frequencies at the HLA-A, -B, and -DRB1 loci in Filipinos living in the National Capital Region of the Philippines showed clear patterns. The

most common allele group, HLA-A*11, accounted for 33.3% of the genotyped population. HLA-B*15 had the highest frequency among the HLA-B alleles, at 29.41%. At a frequency of 49.02%, DRB1*15 accounted for nearly half of all detected alleles, making it the dominant allele in the HLA-DRB1 locus. Moreover, several other alleles were notable for their frequencies exceeding 10%. These included HLA-A*24 at 31.3%, HLA-A*34 at 16.7%, HLA-B*38 at 16.7%, and HLA-B*40 at 13.7% in the HLA-A and -B loci, respectively. Within the HLA-DRB1 locus, the frequency of HLA-DRB1*04 was 15.7%, whereas the frequency of HLA-DRB1*12 was 14.7% [160]. The HLA class I and II polymorphism in the healthy Iranian population of Yazd Province was examined through a thorough analysis of allele frequencies. The study revealed several notable findings regarding the prevalence of specific HLA alleles. At the HLA-A locus, the most frequent alleles identified were HLA-A*02:01, accounting for 18.889% of the population, followed by HLA-A*11:01 at 12.22% and HLA-A*26:01 at 9.44%. Moving to the B locus, prominent alleles included HLA-B*51:01 at 12.778%, HLA-B*35:01 at 10.556%, HLA-B*50:01, and HLA-B*38:01 each at 7.22%. At a frequency of 24.4%, HLA-DRB1*11 was found to be the most common allele within the DRB1 locus. HLA-DRB1*15 and HLA-DRB1*07 followed at 13.3% each [161].

2.8 HLA diversity among the Indian population

The Indian population is a conglomeration of various castes and races that differ in origin, migration, linguistics, socio-cultural aspects, and settlement. All these caste groups and races have distinctly structured the population of India. Therefore, different parts of the country have their unique gene pool. The country may be categorised into four regions to evaluate the distribution pattern of HLA alleles: the southern, northern, western, and eastern regions. The allele frequencies of the various populations in these regions are presented in Table 2.2

TABLE 2.1: HLA-A, B, and DRB1 allelic polymorphisms among global populations.

Place of Study and Author	Populations	HLA-A alleles	HLA-B alleles	HLA-DRB1 alleles
Five Continental Studies [154]	Brazillian		HLA-B*51:01 (8.4%), HLA-B*08:01 (7.4%),	
	Mexican Series		HLA-B*35:01 (46%), HLA-B*40:05 (15.5%)	
	Mexican Mestizos		HLA-B*35:01 (7.3%), HLA-B*18:01 (6.1%)	
	Cuban Caucasoid		HLA-B*07:02 (7.8%)	
	Cuban Mulatto		HLA-B*53:01 (10.7%), HLA-B*35:01 (9.5%)	
	South African Zulu		HLA-B*42:01 (12%), HLA-B*44:03 (10.5%)	
	Omani		HLA-B*51:01 (17.5%), HLA-B*08:01 (11.5%)	
	Singapore Chinese		HLA-B*40:01 (17.2%), HLA-B*46:01 (11.3%), HLA-B*40:01 (17.2%)	
East African Population [155]	East African Population	HLA-A*02:01:01 (10.90%)	HLA-B*58:02 \ (8.79%)	HLA-DRB1* 11:01:02 (11.65%)
Iraq [156]	Iraqi Arab Population	HLA-A*02 (38.48%), HLA-A*01 (22.92%)	HLA-B*51 (20%), HLA-B*35 (19.8%)	
Dhaka, Bangladesh [157]	Bangladesh Bengalee Population	HLA-A*33 (17.02%), HLA-A*24 (16.31%), HLA-A*110%), HLA-A*02 (14.54%),	HLA-B*15 (19.5%), HLA-B*35 (17.02%), HLA-B*44 (9.22%)	HLA-DRB1*15 (29.07%), HLA-DRB1*07 (21.28%), HLA-DRB1*12 (11.35%) & HLA-DRB1*04 (10.28%).

Place of Study and Author	Populations	HLA-A alleles	HLA-B alleles	HLA-DRB1 alleles
Pakistani Population [158]	Pakistani Population	HLA-A*02 (19.2%)	HLA-B*35 (13.7%)	HLA-DRB1*03 (16.8%), HLA-DRB1*07 (13%), HLA-DRB1*11 (12%), and HLA-DRB1*15 (20 %)
Pakistan [159]	Pakistani Gujjar Population	HLA-A*02 (19.6 %), HLA-A*11 (17.0 %), HLA-A*01 (14.4%), HLA-A*26 (12.9%), and HLA-A*31 (11.9%)	HLA-B*51 (25.7%), HLA-B*08 (13.9%),	HLA-DRB1*13 (28.0%), HLA-DRB1*15 (21.1%), and HLA-DRB1*03 (18.6%)
Nepal [162]	Renal transplant recipients & donors, Nepal	HLA-A*11 (34.5%), A*24 (17%), and A*33 (13%)	HLA-B*15 (27%), B*35 (19%), and B*40 (10%)	HLA-DRB1*15 (33.5%), HLA-DRB1*12 (21.4%), and HLA-DRB1*04 (7.32%)
National Capital Region of the Philippines [161]	Filipinos	HLA-A*11 (33.3%), LA-A*24 (31.3%), and HLA-A*34 (16.7%)	HLA-B*15 (29.41%), HLA-B*38 (16.7%), HLA-B*40 (13.7%)	HLA-DRB1*15 (49.02%), HLA-DRB1*04 (15.7%), HLA-DRB1*12 (14.7%)
Iran [161]	Iranian Healthy Population fromYazd Province	HLA-A*02:01 (18.889%), A* HLA-11:01 (12.22%), HLA-A* 26:01 (9.44%)	HLA-B*51:01 (12.778%), HLA-*35:01 (10.556%), HLA-B*50:01 (7.22%), HLA-B*38:01 (7.22%)	HLA-DRB1*11 (24.4%), HLA-DRB1*15 (13.3%), & HLA-DRB1*07 (13.3%)

2.8.1 South Indian Studies

South India is one of the oldest geophysical regions, majorly occupied by Dravidian language-speaking people, often considered the earliest inhabitants of India [163] [164] [165]. Dravidians of South India, with a unique social institution and culture, are subdivided into many gene pools, differing in origin, migration and settlement [166] [167]. Numerous investigations to understand the genetic relatedness of the South Indian population have been conducted based on HLA. The examination of HLA-A, B, and C antigen frequencies within the Koya tribal population of Andhra Pradesh revealed notable variations compared to the average Indian frequencies. Specifically, there was a decrease in the phenotypic frequency of A1, B5, B15, Bw37, and B40 antigens, while an increase was observed in A11, Aw23, B12, and B17 antigens [168]. Similarly, investigations into HLA-A and B antigen profiles were conducted among the Kotas and Badagas residing in the Nilgiri Hills of South India [169]. The Kota population exhibited elevated frequencies of A2 and B7 antigens and a notable presence of the HLA-A2∼B7 haplotype. Additionally, higher frequencies of Aw19, A28, and Bw22 antigens were observed in both the Kota and Badaga populations, indicating distinctive HLA antigen profiles within these groups. Subsequently, Selvakumar et al. (1988) investigated the distribution of HLA-A & B antigen profiles in Tamil-speaking South Indian Hindus residing in Madras. Tamil Hindus were found to lack HLA-A31, HLA-A32, HLA-Aw33, HLA-B16, HLA-B21, and HLA-Bw41. Except for minor differences (low occurrence of Awl9 antigen), the South Indian Hindus revealed similarity to North Indian and other Indian groups. Further, the study suggested that the haplotype HLA-A1∼B17 is shared among all Indians [170]. Subramanian and Damodaran (1995) studied HLA-A, B, C, DR, and DQ antigen profiles in the Sadhu Chetty community of Madras. There was a 28.2% rise in the frequency of HLA-B16 and the presence of haplotype HLA-A2∼B16 in the Sadhu Chetty community [171]. In another study, a primitive Dravidian caste, “Nadars” from Tamil Nadu, was analysed and compared with other populations of India and the world. The Nadars share

several traits with East Asian populations that are consistent with the demographic history of South India, according to high-resolution typing of HLA-A, B, and C alleles. However, they also have some unique traits, such as HLA-A*03011, HLA-A*31011, HLA-B*15011, HLA-B*3501, HLA-B*51011, and HLA-Cw*02022. Also, haplotypes like HLA-A*31011∼Cw*02022∼B*3501 HLA-A*03011∼Cw*04011∼B*4406, and HLA-A*2402101∼Cw*04011∼B*51011 were found to be very uncommon or not present at all in other Indian and global populations [163]. However, they were very common in Nadars and East Asian populations. Thomas et al. (2004) conducted a comparative analysis of the HLA-B and HLA-C alleles among two distinct groups in the Kerala region of India. The first group comprised seven tribal communities: Adiya, Kanikkar, Kattunaikka, Kuruma, Kurichiya, Malapandaram, and Paniya. The second group, the Random Non-Dravidian (RND) group, consisted of Malayalam-speaking individuals from various non-tribal castes. In this study, B*61 was found to be the most common allele. HLA-Cw*14 was one of the most frequent alleles. In contrast, HLA-Cw*17 was absent in all populations studied. It was also found that the most frequent HLA-B alleles in the RND population (B*07, B*61, B*40, B*44, B*51, B*35, and B*52) were similar to those in the North Indian population. Thus, the study suggested that the RND population of South India might have a crypto-Dravidian origin, while the smaller Dravidian tribal communities have a distinct Dravidian origin [165]. In another study, Thomas et al. (2006) compared six nontribal (Namboothiri, Nair, Zhava, Pulaya, Malabar Muslim, and Syrian Christian) communities based on the HLA-A, B, and C alleles. The study revealed that the HLA diversity of the Dravidian communities is very distinct from that of other world populations. Further findings of the study included the identification of strong Dravidian influence in Kerala's non-tribal communities and evidence of genetic admixture with populations from the Mediterranean, Western Europe, Central Asia, and East Asia. The genetic admixture of Dravidians with other populations has introduced many genetic elements, making the present-day South Indian population distinct from the past [172].

2.8.2 North Indian studies

The population of North India is believed to be descendants of Aryans [173]. The invasions during different periods from different directions (North-East and North-West) and mixing with the original inhabitants may have primarily influenced the population composition in these parts of the country. In the earliest study from North India, Mehra et al. (1986) investigated the HLA profile of 400 native North Indian Hindus of Aryan descent. The study reported the gene frequencies of a majority of class I and II antigens, showing similarity to the Caucasoid population. Apart from that, the study also reported the lack of the antigen HLA-B14 and the slightly elevated frequency of HLA-B16 and HLA-B41 [173]. In another study, Rani et al. (1998) examined the DR and DQ loci of the HLA class II system. The most common haplotype found in the population was HLA-DRB1*1501∼DRB5*0101∼DQA1*0103∼DQB1*0601. Also, DR2 haplotypes like HLA-DRB1*1502∼DRB5*00102∼DQA10103∼DQB1*0601 were recorded. These had been seen in Caucasians, Chinese, Gypsies, and Latin Americans before. The migration of Gypsies from India was proposed as an explanation for the relatively rare haplotype HLA-DRB1*1404∼DRB3*0202∼DQA1*01 01 ∼DQB1*0503 that was found in both Gypsies from the Czech Republic and Caucasians. The result also suggested a mixture of Caucasoid, black, and Chinese genes in the North Indian population [174]. Subsequent research revealed that 21 per cent of the samples that Rani et al. (1998) had previously identified as HLA-DRB1*1501 were HLA-DRB1*1506 due to an amino acid substitution from valine to alanine in HLA-DRB1*1506 caused by a nucleotide sequence change from GTG to GCG (at the 50th codon in the second exon) [175] (Rani et al., 1998). A non-radioactive method utilising biotinylated SSOPs was employed in a study by Mehra et al. (1991) to identify the DR-DQ genotypes associated with HLA-DR2, 4, DR51, and DR52 in Asian Indian families. Results showed that some of the HLA-DRB1*1601-associated Caucasoid haplotypes were absent, but the classic HLA-DRB1*1501 and HLA-DRB1*1502 alleles, which are typically associated with Caucasoid haplotypes, were prevalent. The DR52 group comprised the three most common alleles, including HLA-DRB1*0301, HLA-DRB1*1404, and HLA-DRB1*1101.

Additionally, DRB1*1301, DRB1*1302, DRB1*1401, and DRB1*1404 were found to be associated with DR6-alleles. These results indicate the presence of a complex range of HLA-DR2, DR4, and DR52-associated alleles among Asian Indians [176]. Further investigation by Mehra et al. (2001) sought to understand the diversity and distribution frequency of HLA-A*02 molecular subtypes in samples drawn at random from Punjab, Haryana, Uttar Pradesh, and Delhi. The investigation revealed a high occurrence of A*0211 (33.8%) in the population. The increased frequency of HLA-A*0211 among North Indians was mentioned to be a consequence of the founder effect, racial admixture, or selection pressure due to environmental factors in this population. Other common Oriental alleles with increased frequencies, A*0206 (7.5%) and A*0207 (32.5%) and also of HLA-A*0205 (15%) commonly observed in Negroid populations were also reported [177]. Jaini et al. (2002) performed a study to explore the diversity of the HLA-A*19 allele groups in North Indian and Japanese populations. It was found that the frequency of the HLA-A*19 groups of alleles is similar among North Indians and Japanese in comparison to Caucasians. North Indians were found to have all of the known HLA-A19 serological splits, including A33 (15.6%), A32 (8.6%), A31 (3.5%), A30 (3%), A29 (1.2%), and A74 (0.77%). High-resolution analysis revealed that every allele, except the HLA-A*33 group, was divided into two subtypes; A*3301 (4.3%) and A*3303 (43.7%). A new subtype of A*33, A*3306, was additionally identified from the North Indian samples [178]. Additionally, Jaini et al. (2002) looked at the genetic variability in the North Indian population at the HLA-DR4 locus and the DQA1-DQB1 haplotype combinations that are linked to it. A few DR4 subtypes were identified using PCR-SSP and PCR-SSOP procedures. The most common allele among these subtypes was DRB1-0403 (34.8%), followed by HLA-DRB1-0404 (27%), HLA-DRB1-0405 (11%), and HLA-DRB1-0405 (14.6%). Other subtypes observed in the study were HLA-DRB1*0410, HLA-DRB1*0406, HLA-DRB1*0418, HLA-DRB1*0407, HLA-DRB1*0408, and HLA-DRB1*0412, which occurred infrequently in a cohort of 85 HLA-DR4 positive samples studied. Most of these subtypes are combined with DQA1*03-DQB1*0302 (69.5%). It was also found that DRB1*0403

and DRB1*0404 exhibited maximum heterogeneity of DQB1 combinations [179]. In another study, Rani et al. (2007) studied the diversity of HLA-A, B and C alleles among individuals residing in Delhi who have their origins in the neighbouring states of Uttar Pradesh, Haryana, and Punjab. They reported HLA-A*0101, HLA-A*0206, HLA-A*0301, HLA-A*1101, HLA-A*6801, HLA-A*2401 and A*3101 as the most prevalent alleles at the A locus. At the B locus, HLA-B*58011 was found to have a high incidence of 15%, followed by HLA-B*5101, HLA-B*3503 and HLA-B*4006. Similarly, HLA-Cw*0602 and HLA-Cw*0401 were reported as the most frequent C-locus alleles. Other frequent C-locus alleles were HLA-Cw*0102, HLA-Cw*0302, HLA-Cw*0701, HLA-Cw*0702, HLA-Cw*1202, HLA-Cw*1203, HLA-Cw*1502 and HLA-Cw*1503. Besides the presence of Oriental alleles in North Indians like HLA-B*1301, HLA-B*1502 and HLA-B*4001, the alleles common to the African population like HLA-B*5801, HLA-A*68012, HLA-B*5301, HLA-B*44032, HLA-B*4006 and HLA-Cw*1701 were also reported [180]. In a study by Babita and Usha (2004), the HLA antigens were studied among the North Indian Sikh population from Punjab, which demonstrated an increase in frequencies of HLA-A10, B5, and B8 while the frequencies of HLA-A19, A28, B35 and B40 were found to be decreased when compared to other north Indian population. A unique Haplotype, HLA-A2∼B21 was found among the Sikhs [181]. Comprehending HLA studies done so far in North India, the collection of varied genetic elements from different parts of the world is evident in the North Indian population, similar to Southern India.

2.8.3 West Indian Studies

The populations of Western India are described to have Australoid or Proto-Australoid elements with Indo-Aryan racial admixture. Numerous studies have reported the diversity of HLA alleles in western India. HLA Class-I antigen distribution among Brahmins and related caste groups from Maharashtra were compared with other caste groups from

other parts of India [182]. It was revealed in the study that different alleles exhibited increased frequency among different caste groups such as HLA-A2, HLA-B7, HLA-B27, HLA-B40, HLA-B52 in Kunbis; HLA-B35 in Brahmins; HLA-A9, HLA-A19, HLA-A11, HLA-B37, HLA-B16 in Chandrasenia Kayastha Prabhu and HLA-A1, HLA-A2, HLA-A9, HLA-A19, HLA-B7, HLA-B35, HLA-B40, HLA-B53 in Mahars. Whereas HLA-A3, HLA-B13, HLA-B17, HLA-B37 in Kunbis, HLA-A10, HLA-B8, HLA-B13, HLA-B18, HLA-B21, HLA-B37, HLA-B52, HLA-B57, HLA-B60 in Brahmins, HLA-A10, HLA-B8, HLA-B14, HLA-B21, HLA-B55, HLA-B53 in Chandrasenia Kayastha Prabhu, HLA-A10, HLA-B8, HLA-B18, HLA-B22 in Mahars were observed to have significantly decreased frequency. Two locus haplotype analyses revealed that HLA-A10∼B8 was common in Brahmins and Mahars, and HLA-A19∼B12 was a haplotype common to Brahmins and Chandrasenia Kayastha Prabhu. However, HLA-A9∼B15 and HLA-A3∼B5 were the only haplotypes identified among the Brahmins [182]. In another study on Marathas, Shankarkumar et al. (2001) reported the increased gene frequencies of HLA-A1, HLA-A2, HLA-A9 (24), HLA-A11, HLA-A19 (33), HLA-B5, HLA-B7, HLA-B35, HLA-B40 (61), HLA-Cw3, HLA-Cw6, HLA-DR2, HLA-DR5, HLA-DR7, HLA-DQ1, and HLA-DQ2. On analysis of two-locus haplotypes, positive linkage disequilibrium was observed with A10∼B8, A1∼B17, A24∼B52, B5∼Cw9, B13∼Cw3, B15∼Cw2, B35∼Cw4, DR2∼DQ1, DR5∼DQ3, and DR1∼DQ9 haplotypes. Only the haplotype HLA-A2∼B12 was identified in negative linkage disequilibrium. The study suggests that the diversity at the genetic level in the Maratha population is probably a result of genetic drift. Considering the richness of allelic diversity and immense heterozygosity in haplotypes, it was suggested that the Hindu population is not a subject of a single panmictic population [183]. In a subsequent study, Shankarkumar et al. (2002) focused on HLA-A19 molecular subtypes and reported an increased HLA-A*3303 (56%) frequency in the Marathi-speaking Hindus and other common Oriental

alleles. Compared with other populations, it was observed that HLA-A*29 and HLA-A*74 occurred in low frequency among all A19 variants. The high occurrence of HLA-A*3303 (56%) was suggested to be the result of founder effect, racial admixture, or selection pressure due to environmental factors in the population [184]. In another study, Shankarkumar et al. (2003) investigated HLA-DRB1 and HLA-DQB1 locus antigen profiles in Marathas of Mumbai. The study reported frequencies of HLA-DRB1*02, HLA-DRB1*15, HLA-DRB1*0701, HLA-DQB1*06, and HLA-DQB1*0203 to be increased while that of HLA-DRB1*0301, HLA-DRB1*12, HLA-DRB1*09, and HLA-DQB1*04 were found to be decreased [185]. In a study conducted by Chhaya (2005), a sample of 1170 unrelated individuals from Mumbai, Maharashtra was examined. To verify the existence of HLA-B27, a total of 70 samples that tested positive for HLA-B27 were examined for their subtype using high-resolution typing. The analysis revealed the presence of the following subtypes: B*2702 (1.43%), B*2704 (14.29%), B*2705 (70%), B*2707 (12.86%), and B*2718 (1.43%) [186]. Similarly, HLA-A, B and DRB1 loci in the Sindhi population were investigated by Chhaya et al. (2010). The allelic and haplotypic frequency observed in the study was also compared with the Marathi, Gujarati, and North Indian populations in Mumbai. It was observed in the study that in all four groups, HLA-A*01, HLA-A*02, HLA-A*11 and HLA-A*24 at the A locus, HLA-B*35 and HLA-B*40 at the B locus and HLA-DRB1*07 and HLA-DRB1*15 were found to be the more frequent alleles. HLA-DRB1*03 was reported to be significantly high in the Sindhi community. The three locus haplotypes indicated Caucasian and Oriental influence. The analysis revealed that A*02∼B*40∼DRB1*15 and A*33∼B*44∼DRB1*07 were common haplotypes in all the groups. Overall, the literature review on HLA studies in Western India suggests that most investigations revolve around the Maratha population [187]. The studies on tribes and other populations are scanty and require consideration for a better overview of HLA diversity in Western India.

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(MLC) [168]	South India Koya tribe	A11 (29%) A2 (23%) A9 (19%)	B17 (26%) B7 (8%)	Cw3 (21%)	
	South India Kotas	A2 (45.63%)	B7 (45.63%)		
(MLC) [169]	Badagas	A3 (13.79%) Aw19 (120.07%) A28 (5.17%)	B8 (4.31%), B12 (11.21%), B13 (7.76%) B15 (8.62%), Bw22 (9.48%) Bw51 (2.59%)		

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(MLC) [170]	South India Tamil-speaking South Indian Hindus	Al (22.08%) A2 (8.33%) A9 (14.38%) A11 (14.17%)	B5 (14.38%) B17 (10.21%) B40 (10.21%)		
(MLC) [171]	South India Sadhu Shetty	A2 (34.8%) A3 (2.0%) A10(3.5%) All (4.6%)	B16 (28.2%)		DR5 (15.6%) DR7 (21.0%) DR10 (15.6%) DQ1 (44.6%) DQ2 (26.9%) DQ3 (37.1%)

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(PCR-RLS-SSOP) [163]	South India Nairs	A*02011 (10.66%) A*03011 (20.49%) A*3303 (6.56%) A*31012 (18.8%) A*2402101 (15.57%)	B*15011 (17.21%) B*2705 (4.10%) B*3501 (14.75%) B*51011 (15.57%) B*3503 (92.46%)	Cw*02022 (15.57%) Cw*04011 (21.31%) Cw*0702 (14.75%)	
	South India Overall study population		B*07 (13.4%) B*52 ((10.1%) B*51 (10.1%) B*35 (8.50%) B*58 (7.70%) B44 (0.049)	Cw*14 (28.9%) Cw*04 (18.3%) Cw*07 (16.4%), Cw*10 (9.8%) Cw*15 (8.7%) Cw*06 (5.2%)	

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
	(RND group) Random Non-Dravidian population		B*07 (1.92%) B*14 (0.60%) B*63 (0.60%) B*71 (0.60%) B*72 (0.60%) B*27 (1.9)	Cw*07 (25%)	
	Tribal Population Group		B*07 (35%)		
	Kurichiya		B*07 (35%)	Cw*07 (35%)	
	Adiya			Cw*14 ((38%)	
	Paniya			Cw*14 (50%)	

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(PCR-SSP Method) [165]	Kuruma			Cw*14 (40%)	
	Kanikkar			Cw*14 (38%)	
	Malapandaram			Cw*10 (35%)	
	South India Namboothiri	A24 (25%)	B*07 (28%)	Cw*04 (21.3%) Cw*07 (21.3%)	
	Nair	A*24 (23%)	B*07 (24%)	Cw*07 (31.85)	
	Ezhava	A*02 (25%)	B*40(61) (22%)	Cw*07 (22.9%) Cw*15 (27.1%)	
	Pulaya	A*11 (53%)	B*07 (28%)	Cw*07 (18.8%) Cw*01 (21.9%) Cw*14 (21.9%)	

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(PCR-SSP Method)	Malabar Muslim	A*02 (19%)	B*07 (255) B*54 (1.5%) B*15 (76) (1.5%)	Cw*04 (27.9%)	
[172]	Syrian Christian	A24(27%)	B*35 (17%) HLA B* 14 (65) (16%)	Cw*04 (0.226)	
(MLC) [173]	North India	A1 (14.5%) A9 (14.5%) A11 (13.8%) Aw19 (18.4%)	B5 (16%) B35 (14.5%) B40 (12.2%) B16(1.1%)	Cw4 (8.9%) Cw6 (3.7%)	DR2 (47%) DRw6 (9.4%)

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(PCR-oligotyping) [174]	North India				DRB1* 07 (29.8%) DRB3* 0202 (74.5%) DQA1* 0101 (42.6) DQB1* 0503 (34.0%)
(PCR-SSP) [177]	North India	A*0211 (33.8%) A*0206 (7.5%) A*0207 (32.5%) A*0205 (15%)			

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(SBT-Method) [178]	North India	A*33 (15.6%) A*32 (8.6%) A*3303 (43.7%)			
(PCR-SSOP) [180]	North India	A*0101 (10.5%) A*0206 (7.7%) A*0301 (7.2%) A*1101 (9.4%) A*6801 (9.4%) A*2401 (11%) A*3101 (6%)	B*5801 (15%) B*5101 (6%) B*3503 (6.6%) B*4006 (6.6%)	Cw*0102 (3%) Cw*0302 (8%) Cw*0701 (8%) Cw*0702 (9%) Cw*1202 (8%)) Cw*1203 (8%)	

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(Two-stage MLC assay) [182]	West India Kunbis	A2 (53.80)	B7 (23.10%) B27 (11.50%) B40 (38.50%)		
	Mahars	A1 (25.00%) A2 (40.60%) A9 (28.10%) A19 (53.10%)	B7 (28.1%) B35 (25.0%) B40 (43.8%)		
	Brahmins		B35 (36.40%)		
	Chandrasenia Kayastha Prabhu	A9 (64.0%), A19 (56.00%)			

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(Two-stage MLC assay) [183]	West India	A1 (28.37%) A2 (30.80%) A9 (25.26%) A11 (25%)	B7 (23.88%) B35 (26.30%) B40 (20.42%)	Cw3 (29.3%) Cw6 (20.7%)	DR2 (50.53%) DR5 (15.79%) DR7 (29.47%) DQ1 (74.47%) DQ2 (28.72%)
(PCR-SSP Method [185]	West India				DRB1*02 (14.40%) DRB1*15 (11.30%) DRB1* 0701 (14.30%) DQB1*06 (29.60%) DQB1* 0203 (10.80%)

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(PCR-SSP Method) [186]	West India		B*2704 (14.29%) B*2705 (70%) B*2707 (12.86%)		
(PCR-SSP Method) [187]	West India	HLA- A*01 (12.1%) A*02 (16.0%) A*11 (15.2%) A*24 (13.2%)	B*35(11.7%) B*40(9.1%)		DRB1*03 (18.7%) DRB1*07 (12.5%) DRB1*15 (16.4%)
(PCR-SSP Method) [188]	East India Toto Tribe	A*02 (40%)	B14 (32.5%)		
(PCR-SSP Method) [107]	East India Gurkha	A82 (14.00%) A*11(16.33%) A*33 (8.35%)	B*18 (7.3%) B*51 (9.45%)		

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
(PCR-SSP Method) [189]	East India Bengali	A*02 (10.55%) A*11 (12.82%) A*24 (9.17%) A*31 (9.72%) B*07 (8.89%)	B*37 (8.89) B*08 (11.96%) B*37 (8.89%) B*08 (11.96%)		
(PCR-SSP Method) [190]	Bengali Speaking Population	A*33 (20%) A*24 (19%) A*11 (16%) A*02 (14%) A*01 (11%),	B*15 (15%) B*35 (11%) B40 (9%) B*44 (6%)		DRB1*15 (29%) DRB1*07 (21%) DRB1*04 (11%) DRB1 *12 (9%)
	Kayasthas				DRB1*0701 (21.10%)
	Mathurs				DRB1*1313 (30%)

TABLE 2.2: HLA polymorphism among various Indian populations.

		Prevalent Alleles			
HLA Typing method	Place of Study & Ethnicity	HLA-A	HLA-B	HLA-C	HLA-DR, DQ, DP
SSOP Typing [191]	Sunnis				DRB1*1313 (27%) DQB1*0304 (53%)
	Shias				DRB1*0103 (18.4%) DQA1*0602 (5.3%) DQB1*0304 (26%)
	Vaish				DQA1*0603 (25%)
	Lachung				DRB1*0408 (1.70%)
	Kayasthas				DRB1*0701 (21.10%) DRB1*0406 (2.6%) DQA1*0402 (5.3%)

2.8.4 East Indian Studies

Very few studies have been undertaken to understand the diversity of the HLA in eastern India. In this section, East Indian Studies encompasses research conducted in key geographical regions, notably sub-Himalayan West Bengal, and Northeastern parts of India. An investigation was carried out among the very primitive and vanishing sub-Himalayan Indian tribe Toto of North Bengal to study the frequency of HLA-A and B alleles [188]. Alleles like HLA-A*02, HLA-A*23, HLA-A*11, HLA-A*26, HLA-B*07, HLA-B*13, HLA-B*14, and HLA-B*44 showed high frequency. One of the peculiar findings of the study was the presence of HLA-B14 (32.5%), the highest known frequency reported in any population in the world. Moreover, HLA-A*02 (40%) and HLA-B*14 (32.5%) were the highest among the Toto population. Subsequently, Debnath and Chaudhari (2006) studied HLA antigens among the 50 individuals of the Gurkha community from North Bengal. HLA-A and HLA-B loci were analysed and compared with other populations. HLA-A*02, HLA-A*11, HLA-A*33, HLA-B*18 and HLA-B*51 were observed to have increased frequency. Linkage disequilibrium estimates were significant for several haplotypes, including HLA-A3∼B35, HLA-A33∼B44, and HLA-A68∼B35. The study speculated the Gurkha population to be genetically closer to the population of Mongoloid origin. In addition, the study suggested that the origin of the Gurkha population most probably can be from Mongolian and/or Tibetan stocks [107]. Singh et al. (2009) studied HLA-A and B loci in one hundred individuals of the Bengali population. This study reported the increased frequency of HLA-A*02, A*11, A*24, A*31, B*07, B*08, and B*37 amongst the tested alleles. Along with the trend in other Indian populations, a higher incidence of HLA-B*37 and HLA-B*08 was also observed in the study. The two-locus haplotype analysis performed in this investigation revealed significant positive linkage disequilibrium for A*01∼B*37, A*01∼B*40, A*29∼B*40, A*30∼B*51, and A*31∼B*40 [189]. Another independent study examined the distribution of HLA alleles among the Bengali population. The analysis revealed the notable prevalence of certain HLA A alleles, particularly HLA-A*33 (20%), HLA-A*24 (19%), HLA-A*11 (16%), HLA-A*02 (14%), and HLA-A*01 (11%), compared to other HLA A alleles. Similarly, prevalent HLA B alleles included B*15 (15%), B*35, B*40, B*44 (11%), B*52 (9%), and B*57 (6%). Additionally, HLA-DRB1*15 (29%), HLA-DRB1*07 (21%), HLA-DRB1*04 (11%), and HLA-DRB1*12 (9%) were identified as prevalent alleles among Bengalis [190]. The genetic affinity between North Indians and the North-Eastern population was examined by Agarwal et

al. (2008) based on the HLA attributes of these populations. Four endogamous castes (Kayastha, Mathurs, Rastogies and Vaishyas), two inbreeding Muslim populations and three northeast Indian populations (Lachung, Mech and Rajbanshi) were considered for the study. It was found that the number and type of alleles varied among different loci and populations; 36 alleles were observed at the DRB1 locus in both Hindu castes and Muslims from the north, while 21 alleles were seen in North East Indians. At the DQA1 locus, the alleles ranged from 11 to 17 in the studied populations. In the DQB-1 locus, the total numbers of alleles were found to be 19, 12, and 20, respectively, in the studied castes, Muslims, and North-Eastern populations. The study suggested a significant differential admixture with North Indians among Sunnis and Shias. Additionally, a genetic affinity between North-Eastern samples and Mongoloids from South-East Asia was observed. The study reported that North Indians and North-Eastern populations were genetically unrelated [191]. Altogether, the studies conducted in Eastern India are very few compared with other regions. Since Eastern India comprises different ethnic communities with many caste and tribal groups and also due to its geographical location, the allelic polymorphism in this region is expected to be diverse. The episodic invasions also hugely influenced this region. Therefore, HLA-based genetic studies can be vital in discerning the descendants, genetic affinity and migration patterns of different populations in this region.

2.9 HLA and Disease Association

Over fifty years following the first description of HLA associations with disease, HLA molecules have emerged as pivotal components in physiological processes, safeguarding immunity against pathogens while also being implicated in autoimmune reactions that may lead to detrimental health conditions [192]. The specificity of HLA-peptide-T cell receptor tripartite interactions is fundamental in enabling the adaptive immune system to mount an efficient and appropriate response to counteract infection and malignancy while maintaining self-tolerance and preventing autoimmune disease. Understanding the molecular principles that underlie these interactions is crucial for gaining mechanistic insights into the dual role of HLA in initiating and defending against immunopathological phenomena. This pursuit remains a persistent challenge in biomedical research, yet it promises significant therapeutic potential [192].

TABLE 2.3: HLA alleles and their associations with diseases.

Population and Author	Disease	HLA Alleles, SNPs or haplotype	Type of Association	
Caucasian [193]	Ankylosing Spondylitis	HLA-B*27:02 & HLA-B*27:05	Susceptibility	
Caucasian[194]	Rheumatoid Arthritis	DRB1*04:01, DRB1*04:04, DRB1*04:05, DRB1*01:01	Susceptibility	
Iranian [195]	Tuberculosis	HLA-DRB1*07 & HLA-DQA1*01:01	Susceptibility	
Polish [196]	Tuberculosis	HLA-DRB1*16	Susceptibility	
		HLA-DRB1*13		Protection
Indian [197]	Tuberculosis	HLA-DRB1*15:01 & HLA-DQB1*06:01	Susceptibility	
Indian [198]	Leprosy	HLA-DQA1 rs1071630 & HLA-DRB1 rs9270650	Susceptibility	
Han Chinese[199]	Leprosy	HLA-DR-DQ rs602875	Susceptibility	
European [200]	COVID-19	HLA-DRB1*04:01		Protection against disease Severity
Japanese [201]	COVID-19	HLA-DRB1*09:01	Susceptibility	
Russian [202]	Hypertension	HLA-B*18	Susceptibility	
French [203]	Hypertension	HLA-B*18	Susceptibility	
Canadian [204]	Hypertension	HLA-B*18	Susceptibility	
Bahrain [205]	Type 2 Diabetes Mellitus	HLA-DRB1 *040101 &HLA-DRB1 *070101	Susceptibility	
Lebanon [206]	Type 2 Diabetes Mellitus	HLA-DRB1 *070101	Susceptibility	

The mechanism underlying the association between HLA and disease remains elusive, intriguing immunologists for decades and prompting the generation of numerous hypotheses, which can be broadly categorised into two groups. The first group suggests 'mistaken identity', where an HLA allele seems linked to the disease, yet the actual causative factor may reside in a different locus within the haplotype or be associated through linkage disequilibrium. The second group implicates immune reactivity towards self-antigens due to abnormal T cell repertoire selection [207, 208], immune cross-reactivity with external antigens [209] (Oldstone, 1998), or immune assault on 'altered self' antigens [210]. Over the past several decades, population studies have elucidated a broad spectrum of human diseases that exhibit significantly higher prevalence among individuals harbouring specific HLA alleles [194]. For instance, a notable observation is that over 90% of Caucasian individuals diagnosed with ankylosing spondylitis possess distinct HLA Class I alleles, such as HLA-B*27:02 and HLA-B*27:05 [193]. Rheumatoid arthritis (RA) serves as a prominent illustration of an HLA class II-associated disease. Around 90% of seropositive RA patients of Caucasian descent bear one or two HLA-DRB1 alleles, such as DRB1*04:01, DRB1*04:04, DRB1*04:05, DRB1*01:01 [194]. In addition to the diseases mentioned above, specific HLA alleles have been linked to susceptibility or resistance to infectious diseases such as Tuberculosis [195–197], Leprosy [198, 199], and COVID-19 [200, 201]. Furthermore, potential HLA associations have been proposed with cardiovascular disease, exemplified by HLA-B*18 associated susceptibility to hypertension [202–204], and with metabolic disorders like Type 2 Diabetes mellitus (T2DM). For instance, the HLA-DRB1*040101 and HLA-DRB1*070101 alleles have been associated with T2DM in patient cohorts from Bahrain [205], with the latter allele also showing an association with T2DM in patients from Lebanon [211]. Population studies have revealed a diverse array of human diseases associated with specific HLA alleles, ranging from autoimmune conditions like ankylosing spondylitis and rheumatoid arthritis to infectious diseases such as tuberculosis and leprosy. Furthermore, research suggests potential links between HLA alleles and other health conditions like cardiovascular disease and Type 2 Diabetes mellitus, highlighting the pivotal role of HLA genetics in understanding disease susceptibility and informing personalised medicine approaches. Some important associations of HLA alleles with diseases are represented in Table 2.3.

2.10 ABO Blood group system

Karl Landsteiner's seminal discovery of the ABO blood group antigens in 1901, followed by the identification of the Rhesus (Rh) blood group antigen in 1937 in collaboration with Alexander S. Weiner, marked a significant milestone in the field of blood transfusion research. In 1910, Hirszfield and von Dungem made an important discovery regarding the inheritance patterns of blood groups, establishing that these groups followed Mendelian inheritance principles [212]. The inheritance of ABO phenotypes was established as an inherited characteristic, with Bernstein proposing a one-gene locus-three allelic model in 1924 to elucidate the inheritance pattern of this genetic trait. Building upon the earlier work of Hirszfield and von Dungem in 1910, Bernstein's model articulated that individuals inherit a pair of allelic genes from the three possible options: A, B, and O. In cases of heterozygosity, where an individual possesses different alleles for a particular gene, both the A and B alleles manifest dominance over the O allele. However, it is noteworthy that within a heterozygous individual, the A and B alleles do not exhibit dominance over each other [213].

TABLE 2.4: ABO blood group: genotypes, phenotypes, red blood cell surface antigens, and corresponding serum antibodies.

Blood Group (Genotype)	Blood Group (Phenotype)	Antigen in the blood cells	Antibody present in the serum
OO	O	None	Anti-A and Anti-B
OA or AA	A	A	Anti- B
OB or BB	B	B	Anti-A
AB	AB	A and B	None

In a subsequent investigation conducted by von Dungern and Hirszfeld in 1911, it was revealed that blood group A comprises two distinct subgroups: Al and A2. This landmark discovery laid the foundation for the comprehensive classification of blood groups, with the International Society of Blood Transfusion recognising 36 blood group systems encompassing 360 antigens [214].

2.11 Biochemical Genetics of the ABO system

The ABO blood group antigens are encoded by one genetic locus, the ABO locus, which has three alternative (allelic) forms- A, B, and O. The ABO locus is situated on chromosome 9 at 9q34.1-q34.2 within the human genome. The human ABO genes are composed of at least seven exons, and the coding sequence across these seven coding exons spans approximately 18 kb of genomic DNA. These exons vary in size, ranging from 28 to 688 bp, with most of the coding sequence located within exon 7. Notably, exon six harbours a deletion commonly observed in most O alleles, leading to a loss of enzymatic activity [215, 216]. A precursor called the H antigen is essential for synthesising A/B antigens. This precursor is produced by an enzyme encoded by the H locus (FUT1) within red blood cells. The H locus is on chromosome 19 at 19q13.3 and consists of three exons spanning more than 5 kb of genomic DNA. It encodes a fucosyltransferase enzyme responsible for generating the H antigen found in red blood cells [217].

The concept of 'gene-specific glycosyltransferases' proposed by Watkins and Morgan in 1958 marked a revolutionary advancement in blood group genetics. According to this concept, rather than directly encoding for specific antigens, the ABO genes are responsible for producing glycosyltransferases A, B, and O. The A allele encodes a glycosyltransferase that makes the A antigen (N-acetylgalactosamine is its immunodominant sugar). The B allele encodes a glycosyltransferase that creates the B antigen (D-galactose is its immunodominant sugar). The O allele encodes an enzyme with no function, and therefore, neither A nor B antigen is produced, leaving the underlying precursor (the H antigen) unchanged [217]. N-acetylgalactosaminyltransferase (GTA) existing in individuals with A and AB phenotype utilises a UDP-GalNAc donor to catalyse the conversion of the H-antigen acceptor into the A antigen by transferring GalNAc into terminal β-Gal. Conversely, a galactosyltransferase (GTB) existing in individuals with B and AB phenotype employs a UDP-galactose donor to transform the H-antigen acceptor into the B antigen by transferring Gal into terminal β-Gal. In contrast, individuals lacking A and B enzymes present unglycosylated terminal Gal residues on H-substances [218] [219].

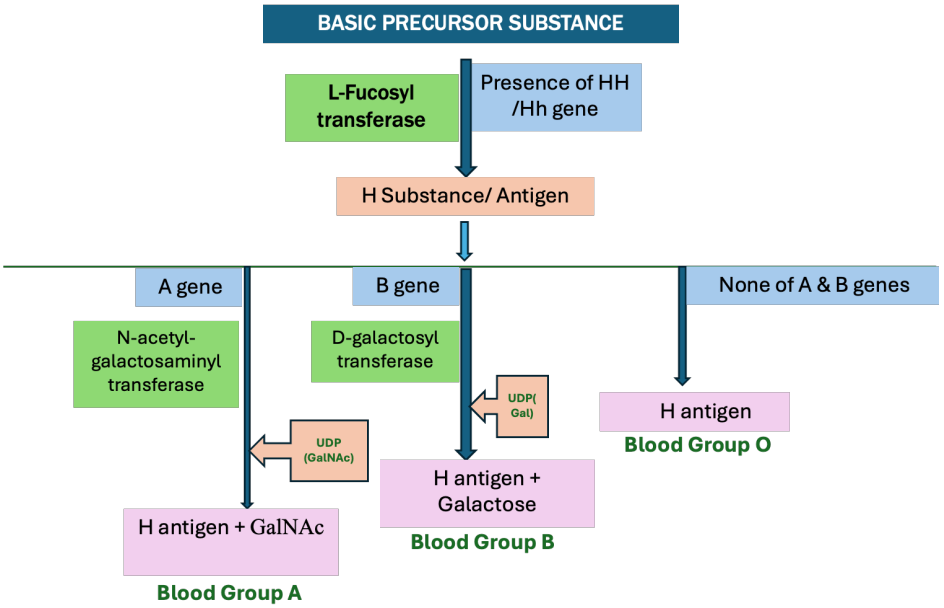


FIGURE 2.3: Schematic representation of ABO blood group antigen determination pathways.

2.12 ABO system and population studies

The ABO system, along with the Rh system, plays a crucial role in the context of blood transfusion and organ transplantation. Comprehensive investigations into the distribution and prevalence of ABO and Rh blood groups have been conducted globally, spanning diverse geographical areas and demographic cohorts. This literature review is centred on recent research endeavours concerning the frequency of ABO and Rh blood groups at both international and national levels.

2.12.1 International Status

Diversity in ABO and Rh antigens is notable among various ethnic groups, showing variation within and across geographical regions [220]. Studies have shown variations in ABO and Rh antigen prevalence among populations globally. In a retrospective study conducted in Gelephu, Bhutan, utilising data from 6045 blood donors over seven years, it was found that the distribution of blood groups followed the order: O (38.21%) $>$ A (29.31%) $>$ B (23.90%) $>$ AB (8.42%) [221]. Similar distribution patterns (O $>$ A $>$ B $>$ AB) have been observed in various other

TABLE 2.5: Distribution of ABO blood group in global populations.

Place of Study	Order of Blood group distribution
Gelephu, Bhutan [221]	O >A >B >AB
South West Ethiopia [215]	
Kurdish population, West Iran [216]	
Sabratha-Libya [222]	
Eastern Nepal [223]	
Ilorin, North-Central Nigeria [224]	
Morocco [225]	
Mexican population [226]	
China [227]	A >O >B >AB
Kayseri Province, Turkey [228]	
Switzerland [229]	
Greece [230]	
Pakistan [231]	B >O >A >AB
Southern Bangladesh [232]	

regions such as South West Ethiopia [215], the Kurdish population of West Iran [216], Sabratha-Libya [222], Eastern Nepal [223], Ilorin in North-Central Nigeria [224], Morocco [225], and the Mexican population [226]. Liu et al. (2017) conducted a large-scale population study in China involving 4,150,214 participants to enhance blood collection and management strategies. They found that blood group A was the most prevalent at 30.5%, followed closely by O at 30.4%, B at 29.4%, and AB at 9.7% [227]. This distribution pattern of ABO blood groups (A $>$ O $>$ B $>$ AB) aligns with findings from studies in Kayseri Province, Turkey [228], Switzerland [229], and Greece [230]. Interestingly, the highest frequency of the B blood group was noted in systematic reviews from Pakistan [231] and Southern Bangladesh [232], following the distribution order of B $>$ O $>$ A $>$ AB. Across these global studies, Rh-positive blood groups were consistently the most common, while Rh-negative blood groups were the least common.

Numerous studies have been carried out to determine ABO allelic frequencies. Dewan (2015) discovered that the allelic frequency order was O (0.5308) $>$ B (0.2585) $>$ A (0.2105) in southern Bangladesh [232]. Similar patterns were observed in Morocco, where the allelic frequency followed O (0.68) $>$ A (0.20) $>$ B (0.10) [225], aligning with the findings in the Mexican population [226]. However, Pakistan showed a different pattern with allelic frequencies of O (0.5756) $>$ A (0.2435) $>$ B (0.1809) [231].

2.12.2 National Status

Numerous studies have investigated the prevalence of ABO blood groups among Indian populations. In a study conducted at Bangalore, Periyavan et al. (2010) observed the order of O>B>A>AB [233], which aligns with findings from other areas such as Bastar district in Chhattisgarh [234], Dakshina Kannada in Karnataka [235], and Andhra Pradesh [236, 237]. Notably, the highest prevalence of blood group B has been consistently observed in many regions. Studies conducted in Gujrat [121, 238] [239], Maharashtra [240] [241], Shimla in Himachal Pradesh [242], Madhya Pradesh [243] [244], Jammu [245], the Kumaon region in Uttarakhand [246], western Rajasthan [247], Haryana [248], and Tirunelveli in Tamil Nadu [249] all reported a predominance of blood group B. This distribution pattern consistently follows B>O>A>AB across these diverse geographical regions. The pattern of blood group distribution, O>A>B>AB is noted among tribal students in Arunachal Pradesh[250], Manipur [251], Madhepura in Bihar [252], and Unnao district in Uttar Pradesh [253]. Conversely, the order A>O>B>AB is observed in Darjeeling [254], Sikkim [255], and Soliga in Tamil Nadu [256]. The prevalence of Rh-positive blood groups is notably higher among the Indian population, with Rh-negative blood groups being rare occurrences [254, 256]. In a recent systematic review covering 112 studies across 24 Indian states, it was noted that O is the most prevalent blood group (34.56%), followed by B (34.10%), A (23.16%), and AB (8.18%). Additionally, the Rh(D)-positive blood group was found to be more common at 94.13% compared to the Rh(D)-negative blood group at 5.87% among the Indian population [257]. The regional distribution of ABO blood groups indicates that B is the predominant blood group in the central (36.18%) and northern (35.62%) regions. In contrast, blood group O is more prevalent in the southern (38.95%), eastern (36.90%), and western (34.66%) regions of India. Additionally, the Rh(D)-positive blood group is more frequent in the Eastern region, with a frequency of 97.42% [257].

The study of allelic frequencies in the ABO blood group has been limited but reveals interesting patterns across different populations. In the Banjara population of Maharashtra [240], South Gujarat R75, Tirunelveli in Tamil Nadu [249] , and Tirupati in Andhra Pradesh [237], the allelic distribution follows the pattern $I^O > I^B > I^A$. However, in the Soliga Tribe of Tamil Nadu [256] and Darjeeling [254], the pattern is $I^O > I^A > I^B$.

TABLE 2.6: ABO blood group distribution order in Indian populations.

Place of Study	Order of Blood group distribution
Bangalore [233]	O >B >A >AB
Bastar district, Chhattisgarh [234]	
Burdwan in West Bengal [258]	
Guwahati in Assam [259, 260]	
Dakshina Kannada in Karnataka [235]	
Andhra Pradesh [236, 237]	
Gujrat [238, 239, 239]	B>O>A>AB
Maharashtra [240, 241]	
Shimla in Himachal Pradesh [242]	
Madhya Pradesh [243, 244]	
Jammu [245]	
Kumaon region in Uttarakhand [246]	
Western Rajasthan [247]	
Haryana [248]	
Tirunelveli in Tamil Nadu [248]	
Arunachal Pradesh [250]	O >A >B >AB
Manipur [251]	
Madhepura in Bihar [252]	
Unnao district in Uttar Pradesh [253]	
Darjeeling [254]	A >O >B >AB
Sikkim [255]	
Soliga in Tamil Nadu [256]	

In summary, studies across diverse geographical regions in India have consistently revealed varying patterns in the prevalence of ABO blood groups, with blood group B often showing the highest frequency. While regional differences exist, the overall distribution trends suggest a predominance of blood group O followed by B, A, and AB. Furthermore, the Rh-positive blood group appears more prevalent than the Rh-negative across the Indian population. Understanding these patterns not only sheds light on the genetic diversity of Indian populations but also holds implications for healthcare practices, such as blood transfusion and disease susceptibility screenings. Further research into allelic frequencies within specific populations promises to unveil additional insights into the complex genetic landscape of ABO blood groups in India.

Chapter 3

Materials and Methods

3.1 Ethical considerations

The study protocol for this investigation was approved by the Institutional Ethical Committee (IEC) of Sikkim University (SU/2012/REG-03/IEC/2323/95) (Annexure A), and the study was carried out as per the Declaration of Helsinki, 2013.

3.2 Population Sampling for HLA Study

The sub-Himalayan region of West Bengal, India, spans between the latitudinal coordinates of $27^0$2'N and longitudinal coordinates of $88^0$15'E. It shares borders with Nepal, Sikkim, and Bhutan, encompassing a land area of 3149 km2. The region sustains approximately 1.847 million individuals, predominantly comprising the Gorkha (or Gurkha), Lepcha, and Bhutia communities. As the substantial population of Kami can be found (52,178 as per the 2011 census) within the sub-Himalayan region of West Bengal, a simplified population sampling method was devised to ensure a comprehensive representation of the Kami population. This methodological strategy aimed to incorporate a diverse subcaste of the Kami population, facilitating a more accurate reflection of the population's demographic composition. Before the blood sample collection, the aim and objectives of the present study were described to prospective participants.

Participants were chosen from individuals whose parents were from the same ethnic group. They were made to fill out the informed consent form for their voluntary participation in the study. Additionally, to enhance the precision of this investigation, all participants were mandated to disclose their castes, subcastes, and other demographic details in the demographic detail form (Annexure B). During the recruitment process, two main criteria were considered for inclusion: firstly, individuals who belonged to the Kami population and were not related for at least three generations; and secondly, ensuring that the participants were also unrelated to each other for at least three generations, which was achieved by creating pedigree charts. The exclusion criteria included individuals with a family history of inter-caste marriages within the last three generations. The individuals were recruited from various sub-Himalayan regions of West Bengal, such as Darjeeling, Kurseong, Kalimpong, and some parts of Siliguri.

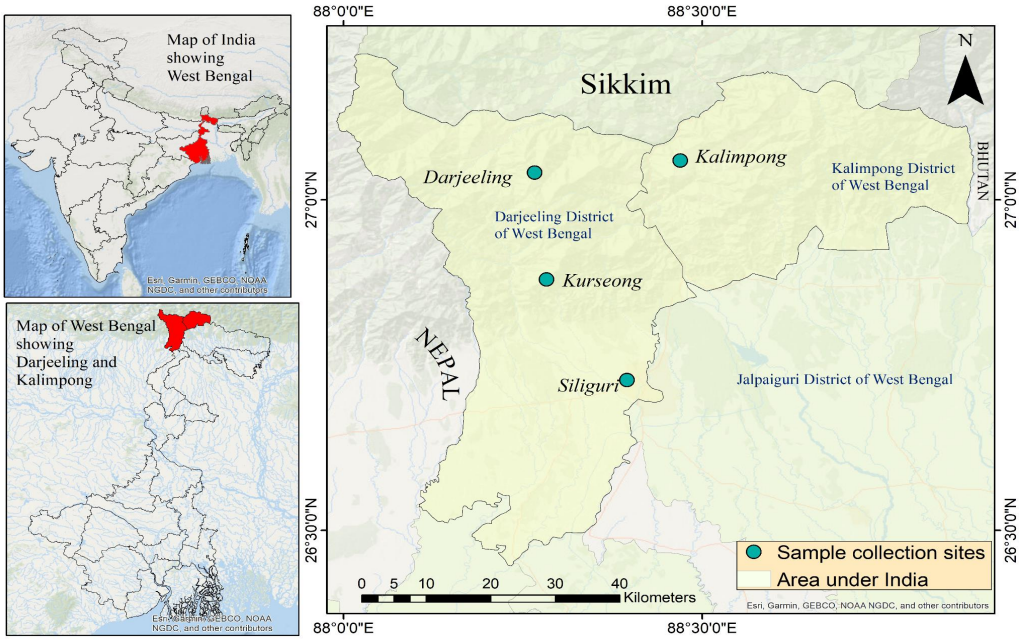


FIGURE 3.1: Geographic distribution of sampling sites for the HLA study.

3.3 Blood sample collection

The blood samples were collected from 158 randomly selected individuals from the Kami population. Approximately 2 millilitres of blood was obtained using the venipuncture technique and collected in vials coated with ethylenediaminetetraacetic acid (EDTA). The collected blood

samples were brought to the Department of Zoology at Sikkim University on the same day and stored in a freezer at -20 °C. All laboratory experiments were conducted at the Molecular Biology Laboratory, Department of Zoology, Sikkim University, Gangtok, India.

3.4 Post-sampling stratification for disease association study

All 158 recruited individuals were randomly screened for the diseases for the association study. The diseases reported by the subjects were recorded (Annexure B), and a comprehensive medical history review was undertaken for the individuals reporting any ailments. The subjects were inquired about any medical conditions and the medications they were currently using. Post-sampling stratification into cases and controls was carried out based on the following criteria for preliminary disease association study:

3.4.1 Selection Criteria for Disease Association Study

The inclusion criteria for the disease association study included Kami individuals with the disease diagnosed by a medical practitioner prior, and the exclusion criteria included Kami individuals without any medical history. The inclusion criteria for the controls included individuals belonging to the Kami population and the absence of a recent or prior medical condition (Annexure B).

Based on the established criteria, diseases with a patient count of 15 or more were designated for inclusion in the disease association study. Accordingly, two diseases meeting this criterion were identified: Hypertension (HT), comprising 20 cases, and Type-2 diabetes mellitus (T2DM), with 15 cases. The study enrolled all the cases after a comprehensive medical history review. Individuals who had any secondary cause for hypertension were excluded. For controls, 80 normotensive and 60 healthy Kami individuals were considered controls for HT and T2DM, respectively. The patient control ratio of 1:4 was maintained to improve the statistical power of the present study due to the small sample size [261].

3.5 Extraction of Genomic DNA

DNA extraction from the whole blood was conducted using a DNA extraction kit (Qiagen). The extracted DNA was then eluted in the TE buffer and ready for use. The extracted DNA was run in 0.8% agarose at 80 mV for 40 minutes. The presence of DNA was documented in the gel document system (Thermo Fisher Scientific). The uncontaminated DNA appeared as a single band near the wells. The purity of DNA was evaluated through the spectrophotometric method, wherein the optical density (O.D.) of the sample was measured at 280 nm to determine protein concentration and at 260 nm to ascertain DNA concentration using Nanodrop (Genetix). The ratio of OD260 to OD280 was then calculated, and DNA samples exhibiting a ratio of 1.7–2.0 were considered adequate for downstream processing. After the quality check, DNA was stored at -20°C for future use.

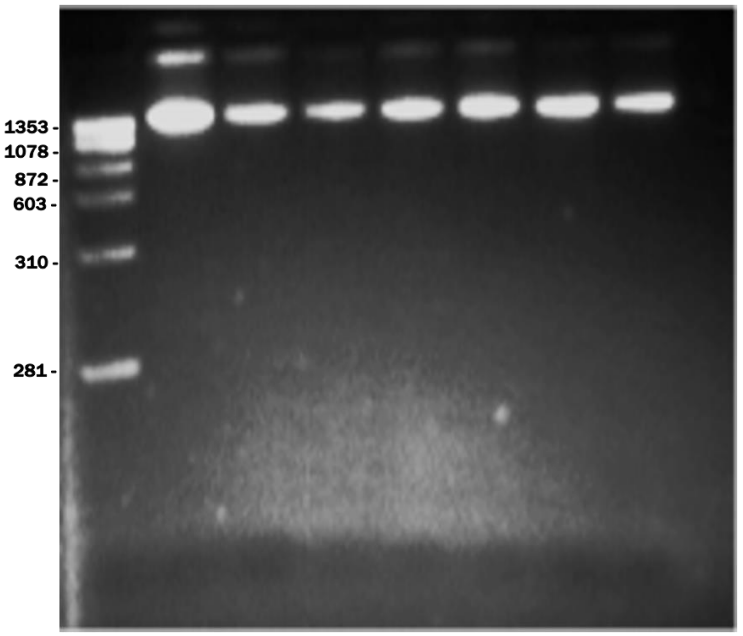


FIGURE 3.2: Gel Electrophoresis image of genomic DNA extracted from blood samples, showing bands and molecular weight markers (ϕX174).

3.6 Genotyping of HLA-A, B and DRB genes

The kit-based sequence-specific primer typing method was used in the present study to type the HLA-A, B, and DRB1 genes. The advantage of using a kit-based PCR-SSP (Polymerase Chain Reaction-Sequence-Specific Primer) over the traditional SSP method lies in the kit's convenience, consistency, and reliability. The HLA Ready Gene ABDRB1 PCR SSP typing kit (Inno-Train Diagnostik, Germany) identifies HLA genes through PCR using sequence-specific primers. The allele-specific primer pairs are designed to amplify target sequences specific to a particular allele or group of alleles. The basic principle of the PCR-SSP approach is that, in carefully controlled PCR conditions, only primers with sequences that precisely match the target sequences provide amplified products. The amplified DNA indicates the presence of the allele of interest in the sample. Apart from primers specific to a particular allele sequence, every PCR reaction mix contained an internal control primer pair. This pair amplifies a conserved area of the housekeeping gene, and the PCR result indicates the PCR reaction's success. The PCR amplification was carried out using the protocol mentioned in the kit. Each tray is designed to test the HLA genes existing in the genome of one individual. It consists of 96 PCR tubes containing pre-aliquoted dried primers (sequence-specific and internal control primers). Each pre-aliquoted well in the tray has the potential to detect 22 numbers of A genes, 35 numbers of B genes, and 13 numbers of DRB genes. The tray layout is shown below.

The PCR master mix was prepared as per the instructions given along with the kit (Table 3.1). After thoroughly mixing, 10μl of the sample was pipetted into the negative control tube, which was the first tube in the PCR typing tray. Subsequently, the DNA was added to the remaining master mix and thoroughly mixed. Following this, 10μl aliquots of the DNA-master mix were dispensed into the remaining wells of the PCR block. An adhesive foil was then used to cover the wells, and the tray was transferred to a thermal cycler (Proflex PCR System, Thermo Fisher Scientific), and PCR was initiated. The PCR program is presented in Table 3.2.

The products of the PCR reaction were assessed by agarose gel electrophoresis, which separates the amplified DNA fragments based on their size. The PCR product was subjected to electrophoresis on a 2% agarose gel, which was prepared in a 1X TBE buffer and contained ethidium bromide for DNA staining. The electrophoresis was conducted at a voltage of 70

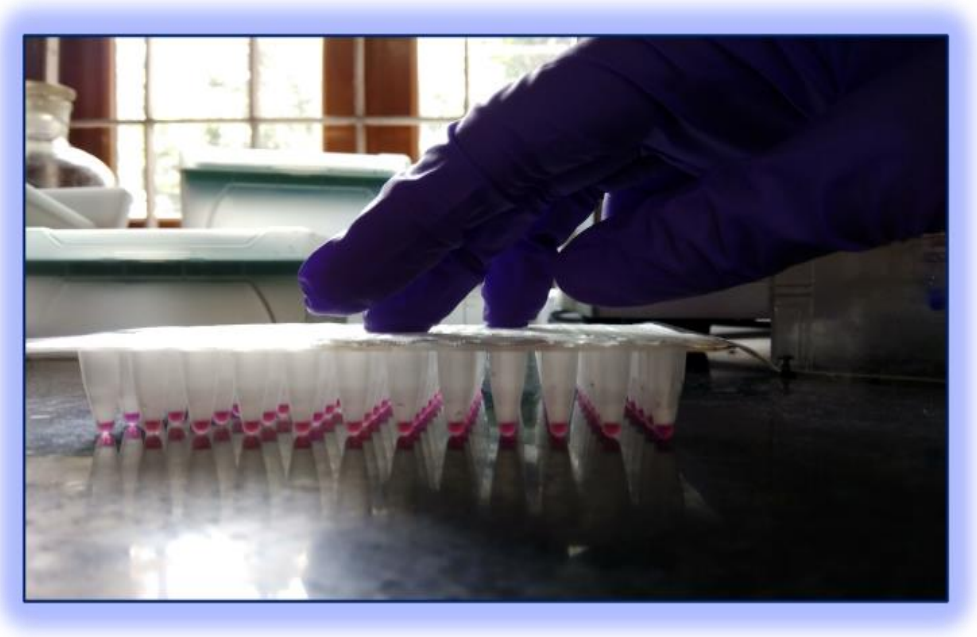


FIGURE 3.3: PCR-SSP HLA typing tray.

TABLE 3.1: Components and the volumes used to prepare the master mix. The stated total volume of the master mix contains a pipetting excess for seven extra wells.

	Volume Per Well (μl)	Volume Per Typing (96+ 7 wells) (μl)
Distilled water	6	(624 μl)
Ready PCR	3	312
Taq Pol	0.08	(8.3 μl)
End Volume Master Mix	9.08	944.3

TABLE 3.2: PCR thermocycler program settings.

1x Initial	10x cycles	20x cycles	Hold
96°C, 2 min.	96°C, 15 sec. 65°C, 60 sec.	96°C, 15 sec. 61°C, 50 sec. 72°C, 30 sec.	4°C endless

mV. The TBE buffer was used as the tank buffer. Bromophenol blue was used as the tracking dye. The amplified fragments were visualised as bands in the gel-doc system. The content of each well of the PCR tray was subjected to electrophoresis separately. The allele list from the IMGT/HLA sequence database release 3.22.0 (October 2015) was used for the first field of HLA Class I and Class II genotyping. The genotype assignment to the individuals was conducted using the HLA-ready gene ABDR gel interpretation worksheet (Annexure C), followed by the 'Ready Gene V.1.0.0.0' software for accuracy.

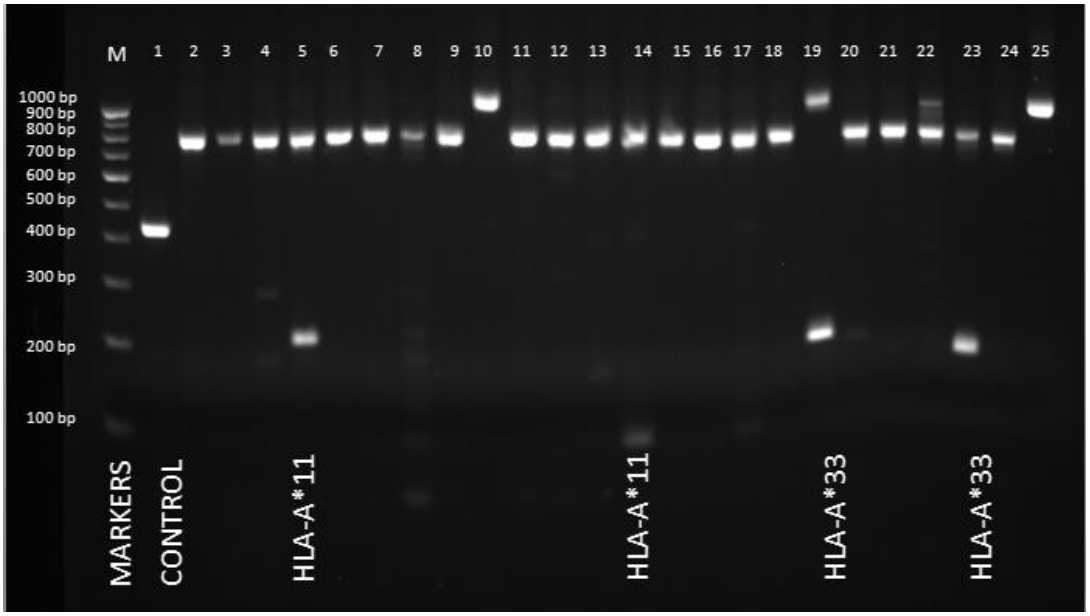


FIGURE 3.4: Gel electrophoresis image depicting distinct bands of HLA-A alleles: Lane M exhibits varying-sized DNA markers, Lane 1 serves as a Negative control, and Lanes 2 to 25 display bands corresponding to internal control gene amplicons positioned near the wells. Notably, Lanes 5, 14, 19, and 23 depict bands indicative of HLA-A alleles.

3.7 Statistical analysis

HLA gene frequency calculation and genotype assignment HLA gene frequency was determined through direct counting, and the loci where only one gene was detected were identified as homozygous. Allele frequency was computed utilising the formula: f=nA/(2N), where f denotes the allele frequency, nA represents the count of observed alleles, and N represents the total

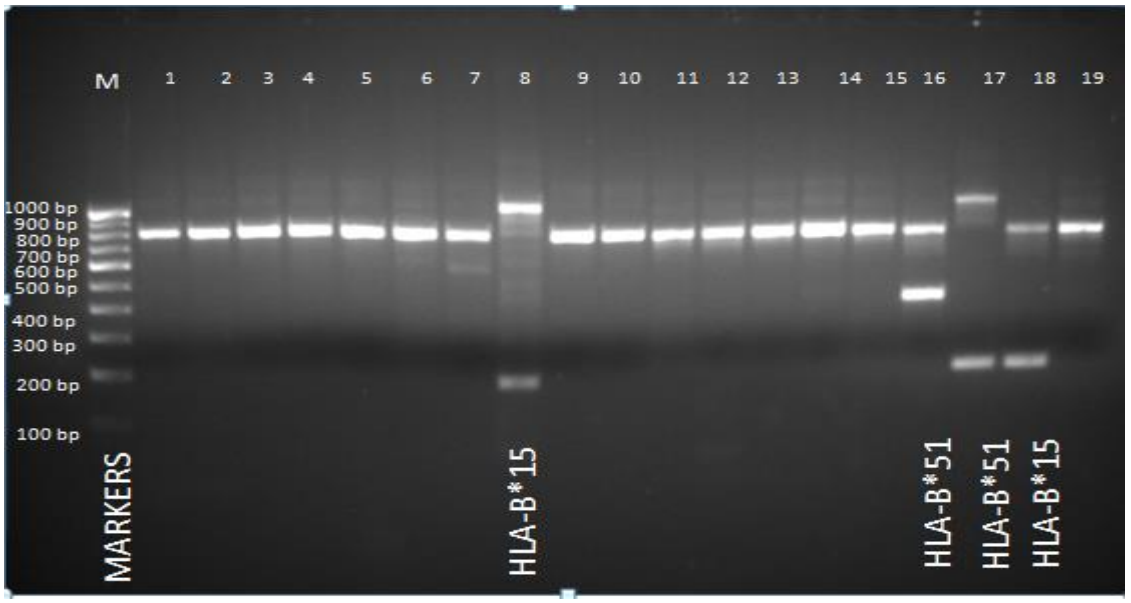


FIGURE 3.5: Gel electrophoresis image depicting distinct bands of HLA-B alleles: Lane M displays DNA markers of diverse sizes, while Lanes 1 to 19 exhibit bands corresponding to the positive control near the wells. Lanes 8, 16, 17, and 18 notably depict bands indicative of HLA-B alleles.

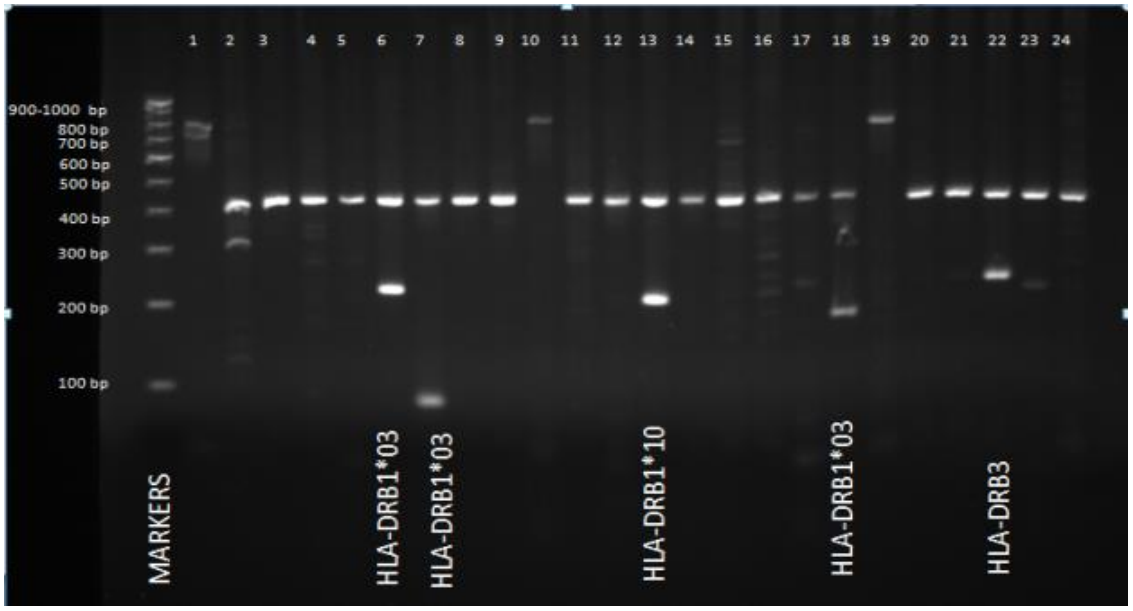


FIGURE 3.6: Gel electrophoresis image displaying distinct bands of HLA-DRB1 alleles: Lane 0 presents DNA markers of various sizes, while Lanes 1 to 24 exhibit bands corresponding to the positive control near the wells. Bands indicative of HLA-DRB1 alleles are observed in Lanes 6, 7, 13, 18, and 22.

number of individuals analysed. PyPopWin32-0.7.0 software estimated the Hardy-Weinberg equilibrium, genotype, and haplotype frequencies.

3.7.1 Genetic Affinity Analysis

The HLA genotype information of comparator populations was sourced from the http://www.allelefrequencies.net database, along with pertinent literature concerning Kannada, Malayalam, Tamil, Telugu, Tulu R2, Indian Bengalis [262], Iranian Kurds [263], Nepalese [162], Pakistan Gujjar [264], and Lebanese [265] populations. Following the computation of the Euclidean distance matrix with center-scaled variables, hierarchical cluster analysis (HCA) was performed to discern patterns of similarity among the dataset. The Ward (1963) method, known for its robustness in minimising variance within clusters, was selected as the clustering criterion for HCA. The ”hclust” algorithm, explicitly utilising the Ward.D2 method, available within the widely used ”stats” R package, facilitated this analysis. Additionally, to gain further insights into the underlying structure of the data, principal component analysis (PCA) was executed using the ade4 package. The resultant clusters and principal components were visually represented using the Factoextra package in R, enhancing interpretability and facilitating subsequent analyses.

3.7.2 Statistical Analysis for Disease Association Study

The odds ratio (OR) with a 95% confidence interval (CI) was obtained from a 2x2 contingency table. The Chi-square test was used to see if the differences between the allele frequencies found in people with HT and T2DM and those found in the control groups for both diseases were statistically significant. $P<0.05$ was considered statistically significant.

3.8 Study area for ABO blood group study

A total of 1420 individuals belonging to the Kami population aged between 18 and 65 years were included in this study from sub-Himalayan West Bengal (Kalimpong, Darjeeling, and Jalpaiguri

district) and Sikkim (Gangtok, Mangan, Pakyong, Soreng, Namchi, and Gyalshing district). The map showing the sample collection site is presented below (Fig.3.7).

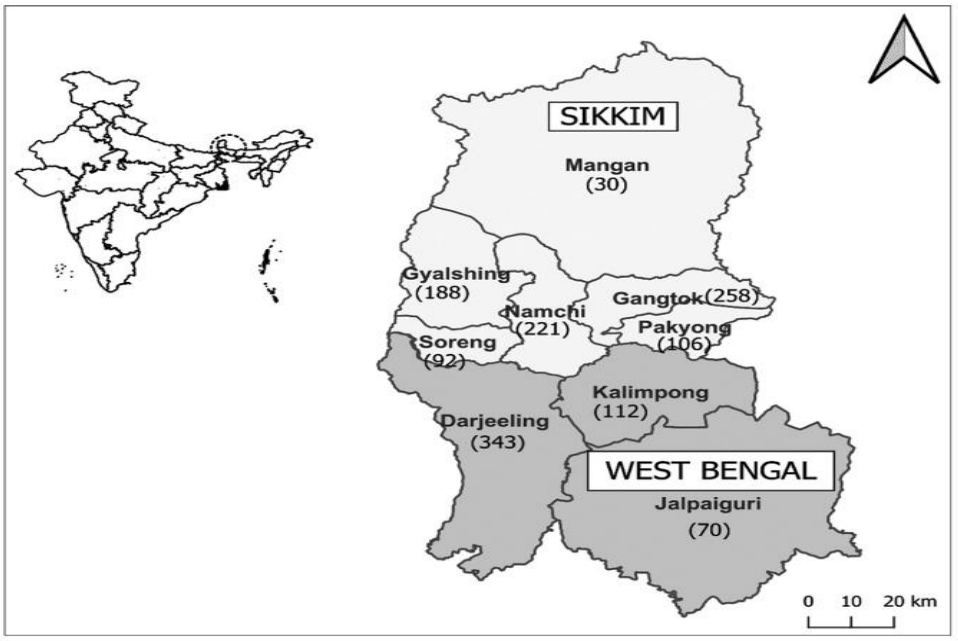


FIGURE 3.7: Map showing sampling sites for the ABO blood group distribution study.

3.9 ABO and Rh blood group determination

The ABO and Rh (D) blood groups of the 158 individuals were determined using the standard slide agglutination method. The procedure involved dividing a glass slide into three sections and mixing anti-A, anti-B, and anti-D separately with a drop of participant blood. The drop of blood was examined for agglutination, and clumping with anti-A and anti-B indicated the blood group for the participant as AB. Agglutination with anti-A or anti-B indicated blood group as A or B for the participant, respectively. The absence of agglutination with anti-A nor anti-B indicated the blood group as 'O' for the participants. The presence or absence of clumping with anti-D was considered Rh-positive or Rh-negative, respectively. In addition, ABO and Rh (D) blood frequency data were generated from 1262 individuals by documenting the blood group of individuals who had undergone routine blood group tests and were certified by diagnostic laboratories. Among the included individuals, 735 (51.76%) were females and 685 (48.24%)

were males. Out of 54 subcastes in the Kami population, the sample population comprised 34 subcastes.

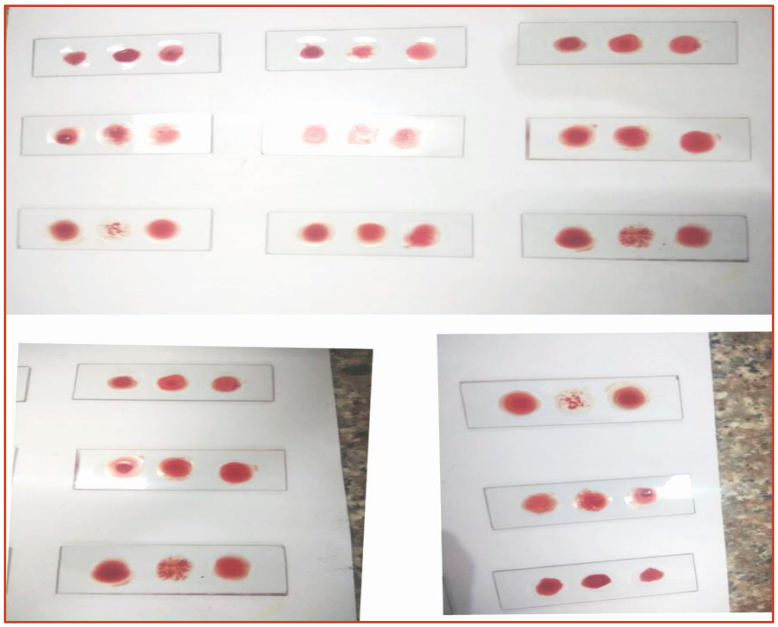


FIGURE 3.8: ABO and Rh typing using the slide agglutination method.

The statistical analysis was conducted using IBM SPSS software, version 20.0 (IBM Corp., Armonk, NY). The frequency of phenotypes for the ABO and Rh (D) blood groups was determined through direct counting and reported as a percentage. The ABO allele frequencies were computed based on the assumption of the Hardy-Weinberg principle, as outlined by Patidar and Dhiman (2021). The maximum likelihood ratio was calculated with the belief that a single ABO gene has three alleles: A, B, and O. The alleles A and B are codominant with each other. Both are dominant over O. Based on the observed allelic frequency values, the expected phenotypic frequencies of blood groups were calculated. A chi-square test was performed to determine the goodness of fit for the collected data. $P \geq 0.05$ was considered to be statistically significant.

Chapter 4

Results

4.1 Results

In the present study, low intermediate resolution PCR-SSP typing of two classical class I loci, HLA-A and HLA-B, was performed in the Kami population residing in the sub-Himalayan, West Bengal, India. Additionally, one HLA Class II locus, DRB1, was genotyped. The genotype and haplotype data of the Kami population have been submitted to the Allele Frequencies Net Database. They are publicly available at http://allelefrequencies.net under India, sub-Himalayan West Bengal, Kami. The genotype frequency distribution of HLA-A, -B, and -DRB1 in the Kami population is presented below.

4.2 Hardy Weinberg Equilibrium test

The frequencies of the various HLA-A, -B, and -DRB1 genes were assessed for Hardy-Weinberg equilibrium. The analysis revealed that these frequencies followed Hardy-Weinberg equilibrium ($P>0.05$). No deviation was observed at all the tested loci for Hardy Weinberg equilibrium.

TABLE 4.1: Results of hardy weinberg's equilibrium test for HLA-A, HLA-B, and HLA-DRB1 loci.

Hardy Weinberg Equilibrium						
Locus	Homozygous		Heterozygous		Fis	P
	Expected	Observed	Expected	Observed		
HLA-A	30.97	27	127.03	131	0.0046	0.6707
HLA-B	14.42	19	143.58	139	0.008	0.5311
HLA-DRB1	33.77	28	124.23	130	0.0059	0.6461

4.3 Frequency Distribution of genes at HLA-A locus in the Kami Population

The distribution of gene frequency at the HLA-A locus is presented in Table 4.2. Of the 22 HLA-A genes tested, 16 were found among the Kami population. The most common genotypes observed at the A locus were HLA-A*11 (31.01%), HLA-A*24 (21.84%), HLA-A*33 (19.62%), and HLA-A*02 (15.51%), which collectively accounted for 87.97% of the total HLA-A genes. Notably, only these four alleles exhibited frequencies exceeding 10%. Homozygous samples constituted 17.09% at the A Loci, with heterozygous samples comprising 82.91%. Rare alleles like HLA-A*23, HLA-A*26, and HLA-A*66 were observed at a frequency of 0.6%. However, other rare alleles, such as HLA-A*69, HLA-A*74, and HLA-A*80, were not found within the population. Similarly, HLA-A*34, HLA-A*36, and HLA-A*43 were absent among the Kami population.

4.4 Frequency Distribution genes at HLA-B locus in Kami Population

Gene frequency distribution at the HLA-B locus among the Kami population is presented in Table 4.2. The current investigation identified 21 out of the 35 tested HLA-B genes within the Kami population. The predominant HLA-B gene observed was HLA-B*15 (23.42%), followed by HLA-B*40 (18.67%) and HLA-B*35 (13.92%), collectively representing 55.99% of the total

TABLE 4.2: Allele frequency distribution at HLA-A locus in Kami polulation.

HLA-A Alleles	% of individuals that have the allele	Allele Frequency in % decimals	Sample size
A*01	5.1	0.0253	158
A*02	28.5	0.1551	158
A*03	1.9	0.0095	158
A*11	55.1	0.3101	158
A*23	0.6	0.0032	158
A*24	38.6	0.2184	158
A*25	1.9	0.0095	158
A*26	0.6	0.0032	158
A*29	3.2	0.0158	158
A*30	1.9	0.0095	158
A*31	1.9	0.0095	158
A*32	1.3	0.0063	158
A*33	36.7	0.1962	158
A*66	0.6	0.0032	158
A*68	5.1	0.0253	158

HLA-B genes analysed. None of the other alleles exhibited a frequency exceeding 10%. Homozygosity was observed in 13.93% of samples, with heterozygosity prevailing at 86.07% at the HLA-B loci. Rare alleles at the HLA-B locus, such as HLA-B*37, HLA-B*39, and HLA-B*78, were detected at a minimal frequency of 0.3%. Notably, the following alleles were not found in the Kami population; HLA-B*14, HLA-B*41, HLA-B*42, HLA-B*45, HLA-B*46, HLA-B*47, HLA-B*49, HLA-B*50, HLA-B*54, HLA-B*59, HLA-B*67, HLA-B*73, HLA-B*81, and HLA-B*82.

4.5 Allele Frequency Distribution at HLA-DRB1 locus in Kami Population.

The gene frequency distribution at the HLA-DRB1 locus is detailed in Table 4.4 Of the 13 HLA-DRB1 genes examined, 12 were identified within the Kami population. The most prevalent HLA-DRB1 genes in the Kami population were HLA-DRB1*15 (41.14%), followed by HLA-DRB1*14 (14.87%) and HLA-DRB1*07 (9.18%). Notably, only the first two most frequent

TABLE 4.3: Allele frequency distribution at HLA-B locus Kami polulation.

HLA-B Allele	% of individuals that have the allele	Allele Frequency in % decimals	Sample size
B*07	5.1	0.0253	158
B*08	1.9	0.0095	158
B*13	10.1	0.0506	158
B*15	43.7	0.2342	158
B*18	8.9	0.0443	158
B*27	5.1	0.0253	158
B*35	24.1	0.1392	158
B*37	0.6	0.0032	158
B*38	1.9	0.0095	158
B*39	0.6	0.0032	158
B*40	33.5	0.1867	158
B*44	13.3	0.0665	158
B*48	5.7	0.0285	158
B*51	12.7	0.0633	158
B*52	4.4	0.0222	158
B*53	1.9	0.0095	158
B*55	1.9	0.0095	158
B*56	2.5	0.0127	158
B*57	3.8	0.0190	158
B*58	7.0	0.0348	158
B*78	0.6	0.0032	158

alleles had a frequency exceeding 10%. Collectively, these three alleles represented 65.19% of the total HLA-DRB1 genes analysed. The gene DRB1*09 was observed with the lowest frequency of 0.6%. In 18.35% of participants, homozygosity was observed, while 81.65% displayed heterozygosity at the DRB1 loci. Importantly, HLA-DRB1*16 was not detected in the Kami population.

4.6 Haplotype Distribution in the Kami population.

Our results for three locus haplotypes A B DRB1 revealed the presence of a total of 134 haplotypes in the Kami population. The three-locus haplotype, whose frequency is equal to

TABLE 4.4: Allele frequency distribution at the HLA-DRB1 locus in Kami polulation.

HLA-DRB1 Allele	% of individuals that have the allele	Allele Frequency in % decimals	Sample size
DRB1*01	3.8	0.0190	158
DRB1*03	13.3	0.0665	158
DRB1*04	10.8	0.0538	158
DRB1*07	18.4	0.0918	158
DRB1*08	4.4	0.0222	158
DRB1*09	1.3	0.0063	158
DRB1*10	14.6	0.0728	158
DRB1*11	5.1	0.0253	158
DRB1*12	12.0	0.0601	158
DRB1*13	4.4	0.0222	158
DRB1*14	29.7	0.1487	158
DRB1*15	68.4	0.4114	158

or higher than 1%, is presented in Table 4.5 The observed most frequent haplotype HLA-A*02∼B*40∼DRB1*15(5.57%) among the Kami population is a Caucasian haplotype. However, HLA-A*11∼B*15∼DRB1*15 (4.81%), the second most frequent haplotype observed in our dataset, has been reported for the first time from an Indian population. The third most frequent haplotype observed among the Kami population is HLA-A*24∼B*40∼DRB1*15 (3.95%). The analysis of HLA-A∼B two-locus haplotypes in the Kami population showed that HLA-A11∼B15 (12%) was the most prevalent, followed by HLA-A11∼B35 (8%) and HLA-A11∼B51 (6.9%). Moreover, the B DRB1 haplotypes that were most frequently observed were B40∼DRB115 (7.2%), B35∼DRB115 (6.9%), and B15∼DRB114 (6.4%). The two-locus HLA-A∼B haplotypes and B∼DRB1 haplotypes, with frequencies equal to or exceeding 1%, are represented in Table 4.6.

4.7 Pairwise Linkage Disequilibrium

Pairwise LD analyses assessed the associations between alleles at different HLA loci within the studied population. The results of these analyses are presented in Table 4.7. The results indicate a significant association between alleles at the HLA-A and HLA-B loci, HLA-A and HLA-DRB1 Loci, and HLA-B and HLA-DRB1 Loci in the Kami population. The positive D value

TABLE 4.5: Three locus HLA A B DRB1 haplotype among the Kami population.

Haplotype	% frequency	Total Sample
A*02∼B*40∼DRB1*15	5.5700	158
A*11∼B*15∼DRB1*15	4.8100	158
A*24∼B*40∼DRB1*15	3.9500	158
A*11∼B*35∼DRB1*15	3.7100	158
A*11∼B*15∼DRB1*14	3.4300	158
A*11∼B*51∼DRB1*15	3.0600	158
A*33∼B*15∼DRB1*14	2.8400	158
A*24∼B*15∼DRB1*15	2.7300	158
A*33∼B*58∼DRB1*03	2.3900	158
A*33∼B*40∼DRB1*15	2.3800	158
A*11∼B*15∼DRB1*07	2.2200	158
A*33∼B*44∼DRB1*07	1.8900	158
A*02∼B*35∼DRB1*15	1.6600	158
A*24∼B*40∼DRB1*14	1.5300	158
A*11∼B*15∼DRB1*12	1.3700	158
A*11∼B*40∼DRB1*10	1.3600	158
A*24∼B*48∼DRB1*12	1.2700	158
A*33∼B*13∼DRB1*07	1.2700	158
A*02∼B*15∼DRB1*15	1.1900	158
A*02∼B*18∼DRB1*15	1.1100	158
A*24∼B*35∼DRB1*15	1.0900	158
A*11∼B*51∼DRB1*14	1.0500	158

and significant P-value suggest that specific alleles at these loci are inherited together more often than expected by chance, implying the presence of linkage disequilibrium between HLA-A and -B alleles. The results indicate non-random associations between alleles at the examined loci, suggesting potential genetic linkage or population-specific allele frequencies.

4.8 Genetic Relatedness of Kami Population with Other Populations.

Hierarchical cluster analysis (HCA) was performed on a comprehensive dataset of 34 distinct population groups, including the Kami population (Fig 4.1). The populations examined for genetic relatedness comprised various ethnic groups from the Indian subcontinent, such as the Nair (NAIR), Namboothiri (NAMB), Ezhava (EZHA), Pulaya (PULA), Punjab (PUNJ), Marathi

TABLE 4.6: Two locus HLA A∼B and HLA-B∼DRB1 haplotypes among the Kami population.

A∼B Haplotype	% Frequency	B∼DRB1 Haplotype	% Frequency	Total Sample
A*11∼B*15	12.06	B*40∼DRB1*15	10.99	158
A*11∼B*35	7.610	B*35∼DRB1*15	8.017	158
A*24∼B*40	6.256	B*15∼DRB1*14	7.493	158
A*11∼B*51	5.116	B*15∼DRB1*15	6.746	158
A*33∼B*44	4.613	B*51∼DRB1*15	4.244	158
A*24∼B*15	4.596	B*15∼DRB1*12	3.165	158
A*02∼B*40	4.358	B*40∼DRB1*10	2.854	158
A*33∼B*40	3.140	B*58∼DRB1*03	2.841	158
A*33∼B*15	3.124	B*13∼DRB1*15	2.816	158
A*33∼B*58	2.685	B*40∼DRB1*14	2.660	158
A*11∼B*40	2.678	B*18∼DRB1*15	2.468	158
A*02∼B*35	2.285	B*15∼DRB1*07	2.403	158
A*24∼B*48	1.899	B*35∼DRB1*14	2.239	158
A*33∼B*13	1.861	B*44∼DRB1*07	1.987	158
A*24∼B*18	1.785	B*13∼DRB1*07	1.529	158
A*33∼B*35	1.729	B*44∼DRB1*04	1.481	158
A*02∼B*15	1.559	B*18∼DRB1*12:	1.204	158
A*29∼B*07	1.266	B*07∼DRB1*15:	1.120	158
A*24∼B*13	1.266	B*27∼DRB1*15	1.089	158
A*68∼B*40	1.266	B*48∼DRB1*12	1.061	158

TABLE 4.7: Pairwise linkage disequilibrium analysis results for HLA Loci in the Kami population.

Locus Pair	D	D'	Wn	p-Value
A:B	0.01478	0.51885	0.35824	0.000
A:DRB1	0.01854	0.43766	0.32657	0.000
B:DRB1	0.01252	0.50048	0.39580	0.000

(MART), Gujrati (GUJR), Indian Sindhi (SIND), North India (NIND), Indian Bengali (IBEN), Gorkha (GORK), along with neighbouring countries such as Nepal (NEPL), Myanmar (MYAN), and Bangladesh, Dhaka Bangalee (DBEN). Additionally, populations classified racially as Mongoloids included Mongolia (MONG), Thailand (THAL), China (CHAN), Japan (JAPN), and Korea (KORN). Seven populations from neighbouring Pakistan were also included: Baloch (BALO), Brahui (BRAH), Kalash (KALS), Pakistani Gujjar (PGUJ), Pathan (PATN), Sindhi Pakistan (SPAK), and Bursho Pakistan (BUPK). Furthermore, one European population from Albania (ALBA) and five Middle Eastern populations, namely Arab (ARAB), Turkey (TURK), Lebanese (LEBN), Iran (IRAN), and Iraqi Kurds (IKUR), were considered in the analysis. The

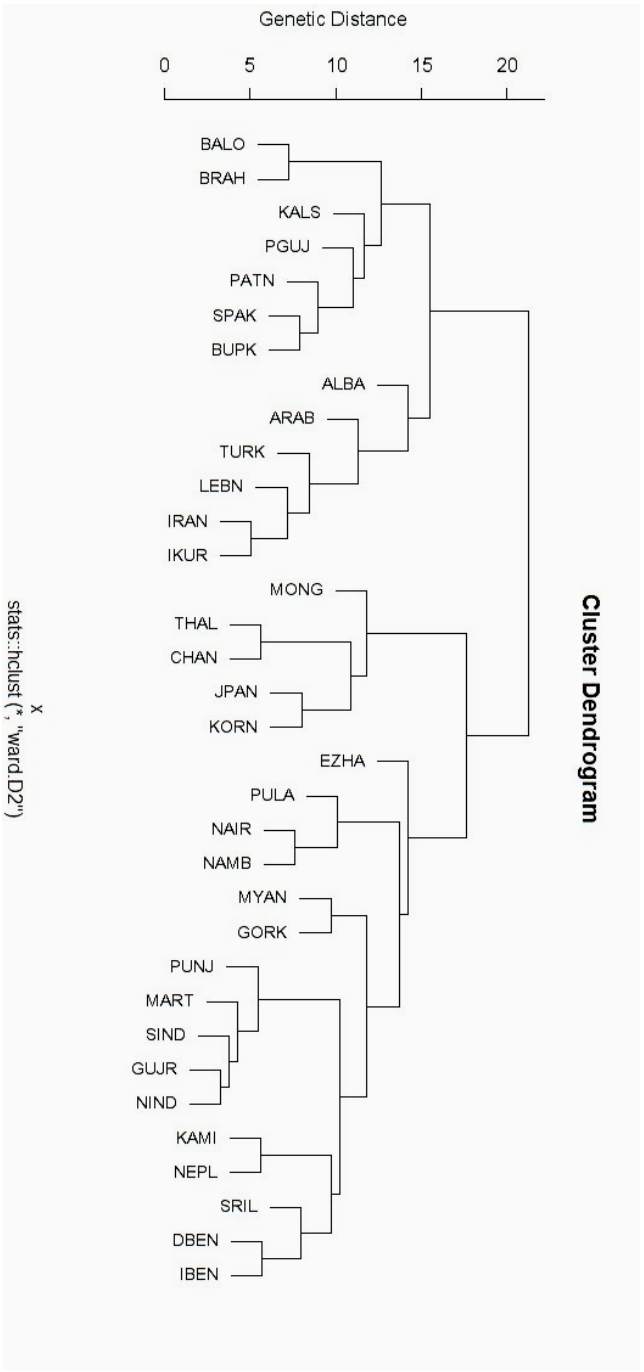


FIGURE 4.1: Hierarchical cluster analysis based on HLA-A and -B allele frequency. Dendrogram depicting the cluster of 34 populations (including the population under investigation in the present study).

Kami population was denoted as KAMI.

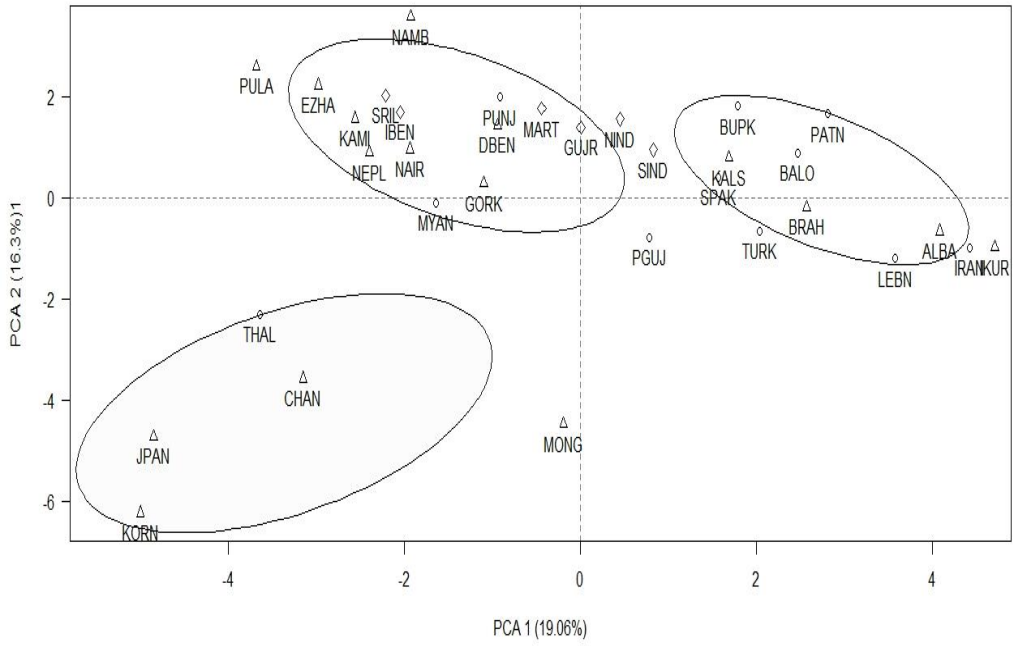


FIGURE 4.2: Principal component analysis based upon the frequency of HLA-A and -B alleles showing the genetic affinity of the Kami population.

The dendrogram presented in (Fig 4.1) delineates the hierarchical clustering of populations according to their genetic relatedness. Evident from the dendrogram are multiple clusters and subclusters, each representing varying degrees of genetic affinity among the populations under investigation. Four primary clusters emerged from the analysis, each with distinctive characteristics. Notably, populations from Pakistan coalesced into one major cluster, with Brahui and Baloch forming a distinct clade on a separate branch. Another major cluster encompassed Middle Eastern populations alongside the Albanian population, with populations from Iran and Iranian Kurds sharing a clade. Within the Mongoloid cluster, Mongolia appeared as a singular leaf, indicative of its genetic distinctiveness within this grouping. The fourth major cluster comprised Indian populations and populations from Nepal, Myanmar, and Bangladesh.

Further granularity within the Indian population cluster was observed, with distinct subclusters identified. These subclusters included Dravidian populations such as Nair (NAIR), Namboothiri (NAMB), Ezhava (EZHA), Pulaya (PULA), and North Indian populations including Punjab (PUNJ), Marathi (MART), Gujrati (GUJR), Indian Sindhi (SIND), and North India (NIND). Additionally, other subclusters comprised populations such as Myanmar and Gorkha, Sri Lankan, Indian Bengali, and Dhaka Bangalee, as well as populations from Nepal and Kami population.

The Sri Lankan population emerged as a singular leaf within its subcluster. Consistent with expectations, the hierarchical cluster analysis (HCA) based on Class I HLA-A and -B gene frequencies highlighted a close genetic association between the Kami population and populations from Nepal.

Principal Component Analysis (PCA) was conducted using the frequency data of 16 HLA-A and 21 HLA-B alleles from the Kami population and comparative data from 33 other populations. The outcomes of the PCA analysis are graphically depicted in Figure 4.2. The first principal component (PCA1) explained 19.06% of the variance, while the second (PCA2) explained 16.03%. Together, PCA1 and PCA2 accounted for 35.09% of the total variance in the dataset. PCA1 likely captures major patterns of genetic variation that distinguish the Kami population from other populations in the dataset. PCA2 captures additional variation, possibly reflecting differences within or between populations. The PCA results delineate three distinct clusters within the population: the Indian subcontinent, Pakistan and the Middle East, and the East Asian population. These findings align with the results obtained from the Hierarchical Cluster Analysis (HCA), indicating concordance between the two analytical approaches.

4.9 Disease Association study

The frequency distribution of HLA genes among the Type 2 Diabetes Mellitus (T2DM) patients and controls belonging to Kami polulation are summarised in table 4.7. Among the T2DM patients, an increased frequency was observed for HLA-B*15 and HLA-B*13, while there was a significantly decreased frequency of HLA-A*33 among the patients. The frequency distribution of HLA genes among the Hypertension patients and controls is summarised in Table 4.9. The hypertensive Kami individuals observed a significantly increased frequency of HLA-A*24, HLA-B*13, and HLA-B*27. At the same time, no association was observed for the genes at DRB1 loci.

TABLE 4.8: Frequencies for HLA-A, -B, -DR alleles in type-2 diabetes mellitus individuals and controls belonging to Kami population.

HLA	Type 2 diabetes mellitus patients (N=15) Gene frequency	Controls (N=60) Gene frequency	Odds ratio	95% CI	Chi- square	p-value
A*01	0.03	0.03	3.50	0.07-0.25	2.52	0.11
A*02	0.14	0.16	1.07	0.32-3.59	0.09	0.77
A*11	0.36	0.30	1.63	0.39-5.36	0.66	0.41
A*24	0.30	0.19	2.28	0.82-7.20	2.05	0.15
A*33	0.06	0.23	0.18*	0.03-0.90	5.07	0.02*
B*13	0.10	0.05	4.75	0.85-26.45	3.66	0.05*
B*15	0.32	0.20	3.45*	1.04-11.41	4.41	0.03*
B*18	0.13	0.05	3.27	0.78-13.56	2.88	0.08
B*27	0.03	0.01	2.07	0.17-24.49	0.34	0.55
B*35	0.03	0.16	0.16	0.020-1.36	3.45	0.06
B*40	0.10	0.20	0.40	0.12-1.57	1.78	0.18
B*44	0.06	0.04	1.38	0.25-7.66	0.14	0.70
B*51	0.10	0.10	0.61	0.17-3.10	0.35	0.55
B*57	0.03	0.02	0.19	0.18-19.72	0.06	0.79
DRB1*07	0.10	0.13	0.90	0.22-3.68	0.02	0.88
DRB1*15	0.50	0.34	2.66	0.67-10.45	2.08	0.14
DRB1*14	0.13	0.15	0.62	0.17-2.21	0.53	0.46
DRB1*13	0.03	0.01	4.21	0.24-71.58	1.15	0.28
DRB1*04	0.03	0.04	2.11	0.35-12.43	0.04	1.00
DRB1*12	0.10	0.06	1.62	0.37-7.05	0.42	0.51

4.10 Results of ABO and Rh Blood Group Study in the Kami Population

4.10.1 Distribution of ABO and Rhesus blood group

In the present investigation, the prevailing blood group within the Kami population was determined to be type O, with a prevalence of 31.06%, followed by B (29.15%), A (27.54%), and AB (12.25%) (Table 4.10). The hierarchy of blood group phenotypes, ordered by frequency, was delineated as O>B>A>AB. Additionally, a noteworthy % of individuals, totalling 97.89%, exhibited Rh (D) positivity, while 2.11% were Rh (D) negative. This distribution demonstrated statistical significance (χ^2 = 10; P<0.05).

TABLE 4.9: Allele Frequency of HLA-A, -B, -DR in hypertensive individuals and normotensive controls belonging to Kami polulation.

HLA	Hypertensive patients (N= 20) Gene frequency	Controls (N= 80) Gene frequency	Odd ratio	95% CI	Chi- square	p-value
A*01	0.02	0.31	0.00	0.00	0.276	0.63
A*11	0.22	0.32	0.70	0.26-1.88	0.49	0.48
A*24	0.37	0.18	4.60 *	1.58-13.25	8.69	0.03*
A*29	0.02	0.01	4.10	0.24-69.5	1.14	0.28
A*33	0.10	0.23	0.34	0.26-1.1	3.43	0.06
B*13	0.05	0.04	4.75*	1.07-21.01	4.89	0.02*
B*15	0.22	0.20	0.90	0.32-2.49	0.04	0.83
B*18	0.10	0.04	3.08	0.77-12.2	2.77	0.09
B*27	0.05	0.02	4.52	0.84-24.4	3.59	0.05*
B*35	0.12	0.15	0.62	0.18-2.05	0.62	0.43
B*40	0.10	0.20	0.94	0.33-2.64	0.01	0.91
B*44	0.05	0.08	1.10	0.27-4.41	0.02	0.88
B*51	0.02	0.09	0.29	0.84-2.44	1.41	0.23
B*57	0.02	0.00	1.05	0.11-9.96	0.60	0.60
DRB1*07	0.07	0.12	0.70	0.18-2.7	0.26	0.61
DRB1*04	0.01	0.02	2.11	0.35-12.43	0.70	0.40
DRB1*10	0.02	0.04	2.17	0.49-9.58	1.09	0.37
DRB1*12	0.07	0.06	1.00	0.00	0.31	0.69
DRB1*13	0.02	0.02	1.35	0.13-13.72	0.06	0.79
DRB1*14	0.10	0.20	0.62	0.20-1.88	0.72	0.39
DRB1*15	0.45	0.35	2.33	0.77-7.04	2.34	0.12

TABLE 4.10: Distribution of ABO and rhesus blood group percentage frequencies among the Kami population

Blood group	Phenotype	Total	Frequency (%)	χ^2	P
ABO	O	441	31.06%	8.172	0.0037
	B	414	29.15%		
	A	391	27.54%		
	AB	174	12.25%		
	Total	1420	100		
Rh D	Positive	1390	97.89%	10	0.02
	Negative	30	2.11%		
	Total	1420	100		

TABLE 4.11: Distribution of ABO and rhesus blood group allele frequencies among the Kami population

Blood group (phenotype)	Observed allele frequency	Genotype	Expected allele frequency	
A	0.28	AA	0.04	0.26
		AO	0.22	
	0.29	BB	0.06	0.33
		BO	0.27	
	0.31	OO	0.31	0.31
AB	0.12	AB	0.10	0.10
	1		1	1
	0.855	DD	0.73	0.855
		Dd	0.25	
Rh D negative (Id)	0.145	dd	0.02	0.145
Total	1		1	1

4.10.2 Distribution of ABO and Rhesus blood group allele frequencies

The allelic frequencies of I^A, I^B, and I^O alleles were computed and evaluated for conformity with the Hardy–Weinberg equilibrium using the maximum likelihood method. Within the Kami population, the allele frequencies of ABO were determined to be 0.200 for I^A (p), 0.243 for I^B (q), and 0.557 for I^O(r) in the sequence IO>IB>IA. The Chi-square test for goodness of fit between observed and expected phenotypes yielded non-significant results ($P \geq 0.05$). Similarly, for the Rh (D) blood group system, the calculated gene frequencies were 0.855 for Rh positive (ID) and 0.145 for Rh negative (Id). The observed frequencies of Rh phenotypes did not differ significantly from the expected frequencies derived from the maximum likelihood calculation ($P \geq 0.05$). The results are presented in Table 4.11.

4.10.3 Distribution of the ABO blood group in the subcastes

The Kami population encompasses 53 subcastes within it. However, only 34 different subcastes from 1367 individuals could be documented in the present study. The distribution of ABO among the various subcastes is presented in Table 4.12. The blood group O emerged as the predominant blood type among subcastes Lohar, Diyali, Dural, Gajmer, Ghimirey, Khati,

TABLE 4.12: Distribution of the ABO blood group in the subcastes of the Kami population

SUBCASTE	A, n (%)	B, n (%)	O, n (%)	AB, n (%)	χ^2	df	p-value
Baraily	47 (12.5)	36 (9.02)	41(10.9)	29(18.01)	4.50	3	0.20
Rasaily	46 (12.3)	50 (9.02)	36 (9.51)	11(6.83)	25.75	3	0.00
Sinchury	17 (4.52)	28 (7.01)	47 (10.20)	5 (3.10)	39.37	3	0.00
Lohar	32 (8.51)	36 (9.02)	44 (10.20)	24(14.90)	6.11	3	0.10
Ghatani	25 (6.64)	39 (9.77)	21 (4.87)	2 (1.24)	32.10	3	0.00
Sashankar	30 (7.97)	29 (7.26)	37 (8.58)	9 (5.59)	16.56	3	.001
Khati	25 (6.64)	16 (4.01)	27 (6.26)	6 (3.72)	14.90	3	0.00
Gajmer	17 (4.52)	10 (2.50)	21 (4.87)	15 (9.31)	3.98	3	0.26
Dural	5 (1.32)	4 (1.00)	16 (3.71)	-	10.64	2	0.00
Darnal	9 (2.39)	9 (2.25)	6 (1.39)	4 (2.48)	2.57	3	0.46
Ramudamu	5 (1.32)	26 (6.51)	4 (0.92)	2 (1.24)	40.94	3	0.00
Diyali	8 (2.12)	4 (1.00)	10 (2.32)	2 (1.24)	6.66	3	0.08
Dutraj	8 (2.12)	-	10 (2.32)	10 (6.21)	0.28	2	0.86
Gadaily	9 (2.39)	1 (0.25)	1 (0.23)	5 (3.10)	11.00	3	0.01
Ghatraj	14 (3.72)	7 (1.75)	8 (1.85)	-	2.96	2	0.22
Koirala	4 (1.06)	12 (3.00)	6 (1.39)	2 (1.24)	12.76	3	0.00
Lakandri	14 (3.72)	12 (3.00)	10 (2.32)	2 (1.24)	8.73	3	0.03
Lohagun	6 (1.59)	13 (3.25)	10 (2.32)	1 (0.62)	10.80	3	0.01
Lamichaney	-	1 (0.25)	-	-	-		-
Kalikote	8 (2.12)	6 (1.50)	7 (1.62)	2 (1.24)	3.60	3	0.30
Portel	3 (0.79)	2 (0.50)	2 (0.46)	-	0.28	2	0.86
Setisural	6 (1.59)	7 (1.75)	8 (1.85)	-	0.28	2	0.86
Pariyar	-	1 (0.25)	-	-	-		-
Ghimirey	3 (0.79)	-	9 (2.08)	4 (2.48)	3.87	2	0.14
Jaarkami	-	-	2 (0.46)	-	-	-	-
Rohpal	6 (1.59)	10 (2.50)	2 (0.46)	4 (2.48)	6.36	3	.009
Ruchal	-	3 (0.75)	-	3 (1.86)	0.00	1	1.00
Poudel	-	4 (1.00)	3 (0.69)	6 (3.72)	1.07	2	0.58
Rizal	-	-	2 (0.46)	4 (2.48)	0.66	1	0.41
Singhouray	-	4 (1.00) 2	2 (0.46)	2 (1.24)	1.00	2	0.60
Silal	2 (0.53)	12 (3.00)	6 (1.39)	-	7.60	2	0.02
Tirwa	2(0.53)	-	12 (2.78)	-	7.10	1	0.00
Sunar	-	-	4 (0.92)	1 (0.62)	1.80	1	0.18
Total	376	399	431	161	1367		1.00

Lohar Lohagun, Setisural, Sashankar, Sinchury, and Tirwa. On the other hand, blood group A was found to occur more frequently in subcastes Baraily, Gadaily, Ghatraj, Lakandri, Kalikote, Portel, and Lamgade. Blood group B exhibited a higher prevalence among subcastes Rohpal,

Ramudamu, Darnal, Ghatani, Rasaily, Shilal, and Singouray. The blood group AB had the highest prevalence in Poudel, Dutraj, and Rizal. Notably, the Chi-square test revealed significantly increased frequencies of blood group A among the Baraily and Lamghade subcastes, blood group O among the Sashankar, Dural, Diyali, and Dutraj subcastes, and blood group B among the Ramudamu, Koirala, and Silal subcastes.

Chapter 5

Discussion

5.1 Discussion

Recent advancements in molecular genetics have facilitated our understanding of historical demographic events by analysing contemporary gene pools within modern populations [266]. Various metrics are utilised to assess genetic variation, offering insights into the phylogenetic relationships among distinct groups. Genetic polytypism is a characteristic of the human species, reflecting variations in allele frequencies across populations at multiple genetic loci. This phenomenon is underpinned by polymorphism within breeding populations, where multiple alleles coexist at numerous loci. Such polymorphism often maintains Hardy-Weinberg equilibrium within populations, achieved through balancing forces such as mutation, genetic drift, or natural selection. Notably, only a tiny fraction of alleles are exclusive to specific regional populations. One of the crucial areas of genetic variation is the polymorphism of the HLA system, which plays a vital role in immune response and disease susceptibility. Polymorphism in the HLA gene complex represents a significant aspect of genetic diversity. This polymorphism refers to multiple alleles at HLA loci within populations, contributing to the extensive allelic variability observed across human populations. The investigation of allelic polymorphism and haplotype diversity in the HLA loci is crucial to the immunogenetics inquiry.

This section of the thesis encapsulates an elaborative discussion of allelic polymorphism and haplotype distribution within the HLA system, specifically focusing on HLA-A, HLA-B, and HLA-DRB1 loci within the Kami population. By analysing haplotype configurations and allelic diversity in these loci, we aimed to elucidate the genetic relationships that define the unique genetic identity of our study population. Additionally, our study delves into the prevalence and distribution patterns of proven disease-associated alleles within the HLA system. We particularly examine the correlation between these alleles and two prevalent medical conditions, hypertension and Type 2 Diabetes, widely acclaimed for their multifactorial origins. By extensively analysing the correlation between HLA variants and these medical conditions, we attempted to provide meaningful perspectives on the genetic underpinnings of these widespread diseases among the Kami population. Furthermore, this chapter extends its inquiry to encompass the distribution dynamics of the ABO and Rh blood group systems within the Kami population. By conducting a comprehensive analysis of ABO blood groups, the present study attempts to shed light on ABO distribution patterns within this population.

5.2 Allelic Polymorphism of HLA-A, HLA-B, and HLA-DR Loci in the Kami Population

Until now, no data on the frequency distribution of HLA alleles at any HLA loci in the Kami population is available. The Kami population belongs to the broader Gorkha ethnic group and is primarily confined to the Indian Sub-Himalayan region of West Bengal. The present study is the first to analyse the genetic diversity at the HLA gene complex among the Kami population. According to the present study's findings, there is a minor disparity in the HLA profile between the Kami population and other Indian populations. Nevertheless, few genes were found to have increased or decreased frequency among the studied sample. The predominant alleles at the A loci within the Kami population, namely HLA-A*02, HLA-A*11, HLA-A*24, and HLA-A*33, exhibit a noteworthy presence in numerous Indian populations [267, 268]. Notably, the frequency of HLA-A*11 in the Kami population was the highest among other Indian populations. Moreover, the other Asian populations share the HLA-A*11, HLA-A*24, and HLA-A*33 genes [162, 269, 270]. Previous high-resolution analysis revealed HLA-A*2402

and HLA-A*1101 as the most frequent alleles in Pacific/Asian populations [271]. Nevertheless, HLA-A*11 demonstrates an association with systemic lupus erythematosus (SLE), leprosy, and pulmonary tuberculosis in the previous studies [272–274]. On the other hand, HLA*24 is associated with Alzheimer's disease, and HLA*33 is protectively associated with Behchet disease [275, 276]. On the contrary, despite being Caucasoid racially, the Kami population did not demonstrate a substantial presence of HLA-A*01, while the HLA-A*0101 allele is highly prevalent among Caucasians and Jews [277]. Notably, Kami people exhibited a lower prevalence of the HLA-A*01 allele, which is common among North Indians and shows similar patterns with the Dravidian population of South India [278]. At the HLA-B loci, the observed frequency of HLA-B*15 among the Kamis is similar to the Pulaya population from Kerala [278]. HLA-B*15 is the most polymorphic allele, with 13 allelic variants among North Indians [279]. However, unlike the present study, HLA-B*15 was observed much more frequently in Oriental populations[280]. HLA-B*15 has been associated with conditions like spondyloarthritis [281]. HLA-B*40 is another common gene found in the Kami population that is also seen in the South Indian population. However, among the North Indian population, there was a notable decrease in frequency, a pattern consistent with the present study's findings [278, 282]. Moreover, HLA-B*40 is also a common gene in other Asian countries [280]. HLA-B*35, another gene commonly observed in Kamis', is shared by populations from the Mediterranean, Pakistan, Bangladesh, and Iraq [263]. On the other hand, in comparison with populations in North India and Western Europe, the Kami population has a low prevalence of HLA-B*07. As a result, the prevalence of HLA-B*35 genes suggests that the genetic makeup of the Kami ethnic group reflects significant influences from the Mediterranean gene pool. However, the notable absence of the HLA-B*07 gene among the Kamis, unlike the North Indian population, suggests minimal influence from the Indo-European gene pool, likely attributed to Kami's lower social hierarchy. Historically, the limited inter-clan marriages due to their lower social status restricted gene flow, contributing to this genetic distinction. Studies on mitochondrial DNA (mtDNA) have further observed a stratification of Indian genetic diversity based on social rank [283], corroborating the impact of social hierarchy on genetic patterns. Conversely, the high frequency of HLA-B*40 among the Kamis indicates the influence of the East Asian gene pool. This observation is consistent with studies on mtDNA haplogroups, affirming the East Asian genetic contribution in many Nepalese populations, including the Kami population's ancestral origins

[284]. Also, the higher-than-average increases in HLA-A*24, HLA-B*15, HLA-B*35, and HLA-B*40 frequencies seen in the Kami population have not been seen before in the Gorkha community, which includes the Kami population. An interesting observation is the presence of high-frequency alleles found within the Kami population among renal transplant recipients and donors in Nepal. These alleles include HLA-A*11 (34.5%), HLA-A*24 (17%), HLA-A*33 (13%), HLA-B*15 (27%), HLA-B*35 (19%), and HLA-B*40 (10%), as well as HLA-DRB1*15 (33.5%), HLA-DRB1*12 (21.4%), and HLA-DRB1*04 (7.32%) [162]. The high frequency of HLA-DRB1*03, HLA-DRB1*07, HLA-DRB1*14, and HLA-DRB1*15 at the HLA-DRB1 loci in the Kami population is similar to what has been found in other Indian caste groups R22. A noteworthy finding of this study was the high frequency of HLA-DRB1*15 (41.14%) among the Kamis, which is the highest reported frequency from North India to date. Previously, a high frequency of HLA-DRB1*15 was documented in South Indian populations such as the Kani tribe (45.91%) and Pallars (43%) [285]. Additionally, the Sinhalese population in Sri Lanka was observed to have a higher prevalence of HLA-DRB1*15 [286]. The increased frequency of HLA-DRB1*15 among the Kami population can be attributed to a combination of factors, including social exclusion and the tradition of assortative mating within their group due to their classification historically as untouchables in the Hindu caste system, which meant social exclusion and limited interaction with other caste groups. This social isolation may have led to genetic drift and the founder effect, where specific alleles, such as HLA-DRB1*15, became more frequent over generations. Additionally, cultural practices and selective pressures related to immune responses or disease resistance could have favoured the inheritance of HLA-DRB1*15 within the Kami population, contributing to its increased frequency. Furthermore, HLA-DRB1*15 has exhibited a positive correlation with multiple sclerosis, cervical cancer, Sjögren's disease, juvenile rheumatoid arthritis, and SLE [287–289]. In conclusion, the genetic analysis of the Kami population, a subset of the Gorkha ethnic group residing in the Indian Sub-Himalayan region of West Bengal, sheds light on their unique HLA allele distribution. Understanding the genetic diversity in the Kami population has broader implications for disease susceptibility and organ transplantation compatibility within this population. Further research elucidating the underlying mechanisms driving these genetic patterns is warranted for a comprehensive understanding of population genetics and its implications for healthcare and anthropology.

5.3 Comparative Haplotype Distribution in Kami population

The most prevalent haplotype, HLA-A*02∼B*40∼DRB1*15, observed within the Kami population with a frequency of 5.57%, represents the highest frequency documented globally. In South Asia, specifically among Sindhi and Pathan populations in Pakistan, this haplotype was observed as the second (2.60%) and third (2.50%) most frequent haplotype, respectively [269]. Furthermore, the Southeast Asian Myanmar Mon population exhibited the second-highest frequency worldwide for this haplotype at 3.90% [290]. Notably, this haplotype is a Caucasian haplotype commonly found among various Indian populations [282]. The second most frequent haplotype, HLA-A*11∼B815∼DRB1*15 (4.81%), observed within the Kami population, is documented for the first time in an Indian population. This haplotype occurs in the second highest frequency (2.0%) in Pakistan's Mixed Pathan Population in South Asia [269]. Furthermore, this haplotype has been shown to exist as an extended haplotype with four to six HLA Loci among the population from Tamil Nadu, India and Colombo, Sri Lanka, in South Asia [291, 292]. The highest reported frequency for HLA-A*11∼B*15∼DRB1*15 (11.90%) has been from the Philippines National Capital Region [290]. Similarly, the haplotype HLA-A*24∼B40∼DRB1*15 (3.95%), identified as the third most prevalent haplotype among the Kami population, has also been observed among the Indian Bengali population [267] and Mongolian population [280]. The results of this study provide compelling evidence for the genetic influence of Caucasian, Oriental, and Middle Eastern gene pools within the Kami population. In the present study, the most prevalent A∼B haplotype identified was HLA-A*11∼B*15, comprising 12.06% of the observed haplotypes within the Kami population. Notably, this haplotype is reported in the Maratha population (1.6%). However, it is worth mentioning that this haplotype exists as an extended haplotype, along with specific HLA-A alleles, within the broader Indian population. Another significant haplotype, HLA-A*11∼B*35 (7.610%), detected in the Kami population, exhibits prevalence among the Pakistan Burusho and Baloch populations. This observation suggests potential gene flow or genetic exchange between these populations, indicative of historical interactions or migratory patterns within the broader South Asian region. Among the HLA-B∼DRB1 haplotypes, HLA-B*40∼DRB1*15 (10.99%) emerged as the most prevalent haplotype within the Kami population. Intriguingly, the identification of

the HLA-B*35∼DRB1*15 (8.017%) variant as the second most frequent haplotype holds significance due to its prominence among the Dhaka Bangalee population from Bangladesh as an extended haplotype A*24∼B*35∼DRB1*15. However, it is noteworthy that the HLA-A*24∼B*35∼DRB1*15 haplotype is also observed in the Pakistan Burusho population, suggesting the intricate nature of genetic diversity and population structure within South Asia, reflecting unique historical and demographic factors shaping genetic variation among different populations.

The haplotype analysis conducted on our dataset unveiled a notable admixture in Kami, predominantly reflecting a Caucasoid component. Complementing the present findings, Y chromosome haplogroup diversity studies suggest extensive genetic admixture within the gene pools of the sub-Himalayan population[293]. Furthermore, this observation aligns with the assertions made by several European scholars, who posit that the presence of West Eurasian haplogroups in India stems from the influx of the Near East's West Eurasian gene pool. This interpretation primarily relies on the frequency distribution of haplogroups [294, 295]. Additionally, the prevalence of frequent South Indian and North Indian genes and haplotypes in the Kami population indicates past genetic admixture. Further supporting this notion, evidence from Y chromosome and mitochondrial DNA (mtDNA) studies suggests a significant impact on Nepal's gene pool due to recent immigration from neighbouring countries [296, 297]. Consequently, Y chromosome haplogroups such as H-M52*, H-M69*, H1-M82*, H1-M370*, R1a1-M198, and R2-M124, alongside mitochondrial DNA haplogroups M31, M33, M35, R6, and R30 [297, 298] commonly shared among South Asian populations, are frequently observed in Nepal's population. Notably, common haplotypes among the Kami, Sindhi, Marathi, and Gujarati populations indicated a shared genetic connection. These observations support the historical narrative of the Kamis' migration from Rajputana, encompassing present-day Gujrat, Madhya Pradesh, and Sindh, to Nepal and settling in Sub Himalayan West Bengal[112]. However, the variation in the allele frequency could be attributed to the founder effect, assortative mating, genetic drift, and selection pressures. Our findings further support the Indian origins of the Kami population since the Kami are said to have come to Nepal from princely states of India and then moved to different regions of north-east India.

5.4 Genetic affinity of the Kami population based on HLA

Most of the studies undertaken to explore patterns of genetic diversity utilising various genetic markers within the Gorkha population of India have primarily focused on individuals from Tibeto-Burman Mongoloid groups [106–108], leading to a common assumption that the Gorkha population possesses Mongoloid descent [107]. It is important to note, however, that the Gorkha population comprises both Caucasoid and Mongoloid communities, challenging the accuracy of such generalisations. In this study, we conducted a genetic affinity analysis to elucidate the genetic relatedness and phylogenetic relationships of the Indo-Aryan Caucasoid Kami population with various other populations. The genetic affinity analysis of the Kami population presented in this study is primarily based on genotype frequency data obtained from the HLA-A and -B loci. This approach is due to the limited availability of genotype frequency data for additional loci among most Indian population groups. The populations included in the Hierarchical cluster analysis (Fig 4.1) and Principal component analysis (Fig 4.2) for assessing the genetic affinity have been purposefully chosen to represent a diverse range of racial and linguistic groups. These populations encompass individuals from Mongoloid (Kalash, Pakistani Gujjar, Chinese, Japanese, Korean) and Caucasian (Baloch, Brahui Pakistan, Pathan, Sindhi Pakistan, Bursho Pakistan, Punjabi, Marathi, Indian Sindhi, Gujarati, North Indian, Nepali, Indian Bengali) racial backgrounds, as well as linguistic groups such as Indo-European (Albanian, Punjabi, Marathi, Indian Sindhi, Gujarati, North Indian, Nepali, Indian Bengali, Dhaka Bangalee), Dravidian (Ezhava, Pulaya, Nair, Namboothiri), and Sino-Tibetan (Kalash, Pakistani Gujjar, Pathan, Sindhi Pakistan, Bursho Pakistan, Chinese, Japanese, Korean, Sri Lankan) populations.

As evident from the phylogenetic analysis, the significant clustering of the Kami population alongside the Nepal population within the same clade points towards a shared genetic ancestry and geographical contiguity. This close clustering likely stems from historical migration patterns or shared evolutionary pressures due to geographic proximity. Furthermore, the substantial overlap in alleles across these populations emphasises their genetic affinity. It suggests a common genetic heritage, reinforcing a closely intertwined genetic relationship between the Kami and Nepal populations. An intriguing observation in the dendrogram is the clustering of

populations based on linguistic groups. Notably, populations belonging to the Indo-European linguistic group in and around the Indian subcontinent, such as Punjabi, Marathi, Indian Sindhi, Gujarati, the North Indian population, Nepali, Indian Bengali, Dhaka Bangalee, and the Sri Lankan population, tend to cluster together.

Similarly, populations from the Dravidian linguistic group, including Ezhava, Pulaya, Nair, and Namboothiri, form another distinct cluster. The close clustering of the Kami and Nepal populations with the Indian Bengali and Dhaka Bangalee clades in the dendrogram suggests a shared genetic heritage and possibly historical migration patterns or gene flow. Separation of the Sri Lankan population as a single leaf in the dendrogram may indicate distinct genetic characteristics or evolutionary divergence compared to the other clustered populations. Moreover, the Sri Lankan population compared in this study is Sri Lankan Sinhalese, who are Indo-Aryan language speakers. This observation opens avenues for discussing the complex interplay of genetic, historical, and cultural factors influencing population clustering and genetic diversity within the South Asian region.

Despite the distinctiveness of the Kami gene pool, both Hierarchical Cluster Analysis (HCA) (Fig. 4.1) and Principal Component Analysis (PCA) (Fig. 4.2) demonstrate the significant genetic closeness between the Kami population and other Indian populations, as well as those from neighbouring countries like Nepal. PCA particularly highlights a more substantial genetic relatedness of the Kami with populations from North India compared to South India. In our analysis, the first two principal components, PCA1 and PCA2, collectively explained 35.09% of the total variance. Indeed, while a greater variance is typically desirable in Principal Component Analysis (PCA) as it suggests a more comprehensive capture of the inherent data structure by the principal components, the extent of variance explained is subject to variability contingent upon several factors. These factors include but are not limited to, the complexity of the dataset, the dimensionality resulting from the number of variables, and the magnitude of variation intrinsic to the data. Therefore, although a higher percentage of variance explained is often sought after, it is imperative to contextualise this metric within the specific attributes and intricacies of the dataset under examination. While this percentage may seem relatively modest, it is important to consider that PCA is a dimensionality reduction technique aimed at capturing the most important features of the data. The interpretation and utility of the principal components should be

assessed not only based on the absolute percentage of variance explained but also in the context of the specific goals and characteristics of the dataset. Moreover, the HCA and PCA analyses delineate a distinct cluster encompassing Mongoloid, Middle Eastern, and Indian-subcontinental populations, within which the Kami population forms a subgroup closely aligned with the Indian subcontinent population. In our study, the observed concordance between PCA and Hierarchical Cluster Analysis (HCA) suggests that the identified principal components effectively captured meaningful patterns of genetic variation, despite explaining a moderate percentage of the total variance. This concordance reinforces the validity and robustness of our findings, highlighting the utility of PCA in elucidating the genetic structure of the studied populations. The results refute the earlier claim that all Gorkhas have a unique close genetic ancestry with Mongoloid populations [107]. Furthermore, the study indicates that despite being part of the Indian Gorkha community, the Kami people's genetic proximity aligns more closely with other Indian populations. In this study, HLA typing was conducted exclusively at the low-resolution level, specifically focusing on the first field/2 digits. Despite this limitation, the present study provides insights into the genetic affinities of the Kami population.

5.5 Disease association study

In the present study, we investigated the presence of proven disease-associated HLA genes among the Kami population and explored their potential correlation with two prevalent conditions, Type 2 Diabetes Mellitus (T2DM) and Hypertension (HT). The aetiology of T2DM and HT is known to be influenced by a complex interplay of genetic, ethnic, and environmental factors. Understanding the genetic components associated with the development of these diseases can provide valuable perspectives into disease management and shed light on their underlying pathophysiology. This study represents the first attempt to disclose the potential correlation between the HLA-A, HLA-B, and HLA-DRB1 genes with T2DM and HT among the Kami population. The observed association of HLA-B*15 with T2DM patients in this study align with previous reports in diabetic cohorts from the Papuan coast of Papua New Guinea [299]. The significant association of HLA-B*15 with susceptibility to T2DM observed in our study is

in concordance with the findings among the patients of North Indian descent [300], which suggests a substantial correlation between HLA-B*15 and risk for T2DM among the Indo-Aryan ethnic groups. In contrast to the current findings, although a higher prevalence of HLA-B*13 was observed among Papuan patients, this difference was not statistically significant [299]. Conversely, our study revealed a significantly higher frequency of HLA-A*33 in control individuals compared to patients, suggesting a potential protective role of HLA-A*33 against T2DM among the Kami population. It could be postulated that HLA-A*33 might contribute to enhanced self-tolerance, thereby mitigating autoimmunity mechanisms implicated in T2DM [301]. In contrast to our findings, it is noteworthy that HLA-A*33 has been linked with susceptibility to Type 2 Diabetes Mellitus (T2DM) in South Indian patients [302]. These discrepancies in findings highlight the multifactorial aetiology of T2DM, where factors such as serum uric acid levels, sleep patterns, smoking habits, dyslipidemia, hypertension, ethnicity, and obesity may collectively influence disease susceptibility [303] In our study however, a higher frequency of HLA-DRB1*15 was noted among T2DM patients, this variance did not attain statistical significance compared to the controls. The finding is consistent with the results observed among the South Indian T2DM patients [302], although they exhibited a lower frequency of HLA-DRB1*15 compared to the present study. Contrary to the present findings, previous research has documented susceptible associations between HLA-DRB1*040101 and HLA-DRB1*070101 alleles with T2DM in patient cohorts from Bahrain [205]. Similarly, in line with the findings of Motala et al. (2005), the HLA-DRB1*070101 allele was also associated with susceptibility to T2DM in patients from Lebanon [211]. The disparity observed between the outcomes of this investigation and the findings from conducted studies among patients in Bahrain, and Lebanon could be attributed to ethnic disparities among the studied populations. A recent Genome-Wide Association Study (GWAS) conducted within the Pashtun community of Khyber Pakhtunkhwa, Pakistan, revealed a positive correlation of T2DM with the single nucleotide polymorphisms (SNPs) rs2308655 within the HLA-B gene [304]. The heterogeneous range of HLA genetic markers implicated in T2DM research across various racial and ethnic cohorts indicates that variations within the HLA genes may substantially influence disease susceptibility within specific populations.

The association of HLA with HT has been well-documented across various populations [305,

306]. The results of this investigation suggest a potential linkage between HLA-A*24 and hypertension within the Kami population, which is consistent with observations among Greek populations [307]. Additionally, this study identified an association between HT and HLA-B*13 in the Kami population, aligning with previous investigations among Russian and Caucasian HT patients [308, 309]. Identifying HLA-A*24 and HLA-B*13 as potential risk factors for HT represents a significant outcome of the present study. Previous studies have investigated the possible effects of HLA genes on the degree of hypertension and damage to target organs. For instance, HLA-B*15 has been linked in White populations to diseases like grade III hypertension, malignant hypertension, and essential hypertension with stroke. Although HLA-B*18 was detected in 10% of Hypertensive Kami individuals; the association did not attain statistical significance (OR = 3.08, CI = 0.77-12.2, P = 0.09). On the contrary, previous studies have documented a susceptible association of HLA-B*18 in hypertensive patients from Russia, France, and Canada [202–204]. The absence of an observed association between hypertension and HLA-B*35 in the Kami cohort warrants further research despite its high frequency of 12%. Studies among hypertensive patients with non-insulin-dependent diabetes mellitus have reported associations of hypertension with HLA-B*35 in White Americans and HLA-DR3 in Black Americans [206]. This discrepancy highlights the complexity of genetic influences on hypertension susceptibility, which may differ among ethnic and racial groups. Similarly, specific HLA-DRB1 bearing haplotypes, such as HLA-DRB1*0406∼DQB1*0302, are linked with Idiopathic Pulmonary Arterial Hypertension (IPAH) in Korean patients and HLA-DRB1*0101/2 DQB1*0501/2 DQA1*0102 have been documented to be associated with an elevated risk of essential hypertension in Slovenians[305]. However, the current investigation did not ascertain any association between the HLA-DRB1 gene and hypertension. This disparity could be attributed to the relatively modest sample size, methodological variations, and the intricate interplay of genetic and environmental factors influencing these relationships. The present study is subject to several limitations that merit acknowledgement. Firstly, the study was constrained by its relatively small sample size, which may have limited the statistical power to detect associations. Secondly, HLA polymorphism was assessed only at the first field level, potentially overlooking more detailed associations. Thirdly, clinical parameters were not systematically investigated or correlated with HLA variants, which might have offered additional insights into

disease mechanisms. Despite these limitations, the current investigation offers preliminary evidence supporting the potential association of HLA genes with HT and T2DM. Further research with larger and more diverse cohorts, comprehensive HLA typing, and incorporation of clinical data is warranted to validate and expand upon these initial findings.

5.6 ABO blood group and Rh system in the Kami population

The present study represents the first comprehensive attempt to delineate the ABO blood group frequency exclusively within the Indian Kami population, working with an adequate number of samples derived from a broad geographic region. According to the findings, blood group O has the largest prevalence (31.06%), followed by B (29.54%), A (27.54%), and AB (12.25%). Notably, the ABO blood group distribution in the Kami community follows the same sequence: O>B>A>AB, which is similar in the general Indian population. These results are in line with our previous investigation of the Sikkim Kami population [255], which found a similar prevalence order (O>B>A>AB). Significantly, this study benefits from a larger sample size (1420 vs 220) collected from a more extensive area, improving its representativeness and statistical robustness. Consistent with the current investigation, previous research by Gurung (2019) on the Kami population of Darjeeling and Kurseong also observed blood group O as the most prevalent blood group [254]. However, the study reported blood group A as the second highest, which deviates from the current findings. This discrepancy may be attributed to the relatively small sample size of the previous study (n=354), which suggests the possible effect of sample size on the consistency of results. Contrary to our findings, Gurung (2019) reported that the inhabitants of Kurseong and Darjeeling in sub-Himalayan West Bengal (9030 Sample) exhibited an overall blood group frequency distribution of A>O>B>AB [254]. Furthermore, our findings do not corroborate with the findings among the cohort from southern parts of West Bengal, where the frequency distribution of the ABO blood group was observed to be B>O>A>AB[310]. Ethnic differences and heterogeneity within the populations included in previous studies may contribute to this lack of agreement in results. Blood group O is prevalent across Southern (38.95%), Eastern (36.90%), and Western (34.66%) regions of India [257]. Conversely, blood group B is more

common in Central (36.18%) and Northern (35.62%) areas [257]. Individuals with blood group O are more likely to have a heightened risk of severe cholera infection [311]. In contrast, blood group O protects against severe malaria due to the unsuitability of erythrocytes for rosette formation by Plasmodium falciparum [312]. Since the current demographic pattern of Gorkhas in the sub-Himalayan West Bengal is mainly due to migration across the Indo-Nepal border since 1700 CE, we compared the blood group frequency of Kami with the Nepalese population. The prevalence of blood groups among the Nepalese population reveals blood group O as the most common at 35.1%, followed by blood group A at 29.7%, with the overall distribution order being O>A>B>AB [313]. The observed discrepancy in these results compared to our findings may be attributed to the inclusion of a heterogeneous population in the Nepal study. Including diverse genetic backgrounds within the Nepalese population likely contributed to the observed variations in blood group frequencies between the studies. The blood group frequencies observed in the Nepalese population are comparable to those in the Bhutanese population, where O was more prevalent at 29.31%, followed by A at 29.31%, B at 23.90%, and AB at 8.42% [221]. The Pakistani population displayed a different pattern, with B as the most common blood type at 33.37%, O at 33.14%, A at 23.99%, and AB at 9.74% [231]. The results from this study are consistent with those from studies conducted in Afghanistan and Sri Lanka, where O was the most prevalent, followed by B, A, and AB [314, 315]. However, the ABO blood group distribution among Bangladeshis revealed that the B blood group was the most common, accounting for 34.15% of the population, followed by O at 29.67%, A at 26.57%, and AB at 9.61% [316]. These discrepancies may be attributed to ethnic differences within each population, influencing the observed blood group frequencies compared to our present findings. The current study's observation of Rh (D) predominance is consistent with previous reports from other populations. Compared to the Kami population, slightly higher proportions of the Rh (D) antigen were observed in populations from Bhutan (99.85%), Sri Lanka (98%), and Sikkim (99.47%) [221, 255, 315].

Despite reports indicating a high prevalence of blood group B in the Himalayan region, our investigation revealed the O blood group as most prevalent among the Kami population from sub-Himalayan West Bengal and Sikkim. This observation contrasts with the expected trend based on regional patterns. The discrepancy may be attributed to the historical migration patterns of the Kami population, whose ancestors are believed to have moved from many different places

of India, such as Punjab, Rajputana, Kashmir, and neighbouring areas between the 10th and 15th centuries AD, settling in Nepal before subsequently moving to sub-Himalayan West Bengal [102, 112]. These historical migration routes likely contributed to the observed pattern of blood group distribution among the Kami population, deviating from the typical regional prevalence in the Himalayan region.

Our study encompasses 34 subcastes of the Kami population, some with limited participant numbers, revealing notable patterns of blood group distribution in Kami subcastes (Table 4.11). The predominance of the O blood group across numerous subcastes within the Kami population, such as Lohar, Diyali, Dural, Gajmer, Ghimirey, Khati, Lohar, Lohagun, Setisural, Sashankar, Sinchury, and Tirwa, signifies a notable genetic inclination toward this blood type within sub-castes of the Kami population. This pattern suggests a significant genetic propensity towards blood group O within the Kami population, potentially influenced by shared ancestry, historical factors, or genetic drift. However, in contrast, the increased prevalence of blood group A among sub-castes such as Baraily, Gadaily, Ghatraj, Lakandri, Kalikote, Portel, and Lamgade, the predominance of blood group B in sub-castes like Darnal, Ghatani, Rasaily, Rohpal, Ramudamu, Shilal, and Singouray, and the higher occurrence of blood group AB in Dutraj, Poudel, and Rizal subcastes suggest distinctive genetic variations and potential historical influences shaping blood group distribution within these subgroups.

The ABO allele frequency distribution (O>B>A) among the Kami population is consistent with several regions in India, including West Bengal, Tamil Nadu, Uttarakhand, Andhra Pradesh, Gujrat, Punjab, and Chattisgarh (Table 5.1). However, discrepancies were noted with populations from Darjeeling, the general Indian population, Pakistan, Iraq, Oman, and Mexico (O>A>B). This variation indicates the influence of environmental factors, evolutionary changes, and population migration on ABO allele frequency distribution patterns [253]. Most samples included in this study were from Kalimpong, Darjeeling, and Jalpaiguri districts of sub-Himalayan West Bengal, along with Sikkim. However, no samples were acquired from the Cooch Bihar district of sub-Himalayan West Bengal, and a limited number of samples were collected from the Mangan district of Sikkim (n = 30). This limitation restricted the comprehensive analysis of ABO frequency across all subcastes of the Kami population. Despite these limitations, the study provides valuable data on the distribution of the Rh and ABO blood groups within the Kami population.

TABLE 5.1: Comparative analysis of ABO blood group and allele frequencies across diverse populations.

Place of study	Order of ABO Allele frequency Distribution	Allele frequency		
		I^A (p)	I^B (q)	I^O (r)
India [257]	O>A>B	0.2717	0.1772	0.551
Srinagar, Uttarakhand [317]	O>B>A	0.2403	0.2475	0.5122
Tamil Nadu [318]	O>B>A	0.1628	0.2177	0.6259
West Bengal [310]	O>B>A	0.189	0.244	0.244
Darjeeling and Kurseong [254]	O>A>B	0.2717	0.1772	0.5511
South Gujarat [238]	O>B>A	0.1844	0.2477	0.5679
Tirupati, Andhra Pradesh [237]	O>B>A	0.1398	0.2148	0.6454
Mexico [226]	O>A>B	0.1676	0.0562	0.7762
Punjab [319]	O>B>A	0.1710	0.2700	0.5590
Chattisgarh [320]	O>B>A	0.1716	0.2512	0.5792
Iraq [321]	O>A>B	0.20	0.14	0.64
Pakistan [322]	O>A>B	0.2330	0.1714	0.6062
Malaysia [323]	O>B>A	0.17	0.20	0.63
Oman [324]	O>A>B	0.15	0.14	0.71
Present study	O>B>A	0.200	0.243	0.557

Further, the patterns observed in blood group prevalence in Kami subcastes highlight the complexity of genetic diversity within the apparently homogeneous Kami population. Variations in blood group profiles can influence susceptibility to diseases, compatibility for blood transfusions, and immune responses to pathogens. Consequently, our study emphasises the importance of integrating sub-caste-specific genetic profiles into healthcare strategies, disease management protocols, and population-based genetic studies for the Kami population.

The ABO allele frequency distribution (O>B>A) among the Kami population is consistent with several regions in India, including West Bengal, Tamil Nadu, Uttarakhand, Andhra Pradesh, Gujrat, Punjab, and Chattisgarh, as demonstrated in Table 5.1. However, discrepancies were noted with populations from Darjeeling, the general Indian population, Pakistan, Iraq, Oman, and Mexico (O>A>B). This variation indicates the influence of environmental factors, evolutionary changes, and population migration on ABO allele frequency distribution patterns, as mentioned in the literature [253]. Most samples included in this study were from Kalimpong, Darjeeling, and Jalpaiguri districts of sub-Himalayan West Bengal, along with Sikkim. However, samples

were not obtained from the Cooch Bihar district of sub-Himalayan West Bengal, and a limited number of samples were collected from the Mangan district of Sikkim (n = 30). This limitation restricted the comprehensive analysis of ABO frequency across all subcastes of the Kami population. Despite these limitations, the study provides valuable data on the ABO and Rh blood group distribution within the Kami population. Further, the patterns observed in blood group prevalence in Kami subcastes highlight the complexity of genetic diversity within the apparently homogeneous Kami population. Variations in blood group profiles can influence susceptibility to diseases, compatibility for blood transfusions, and immune responses to pathogens. Consequently, our study emphasises the importance of integrating subcaste-specific genetic profiles into healthcare strategies, disease management protocols, and population-based genetic studies focused on the Kami community.

Chapter 6

Conclusion

6.1 Conclusion

This study is the first to examine HLA class I and II genes, ABO and Rh system in the Indo-Aryan Kami population, a subgroup of the broader Gorkha community. To the best of our knowledge, it also constitutes the pioneering investigation into the genetic composition of Nepali-speaking Indo-Aryan Gorkhas living in India. It is also the first attempt to analyse the constituent populations within the larger Gorkha community separately.

The genetic characterisation of the Kami population highlights significant occurrence rates of HLA alleles such as HLA-A*11, HLA-B*15, and HLA-DRB1*15, indicating unique genetic variability and evolutionary lineage. A noteworthy finding of the present study is the pronounced predominance of HLA-DRB1*15 in the Kami population, representing the highest recorded frequency observed among the North Indian population. The genetic clustering analysis conducted herein substantiates the genetic delineation between the Kami and Mongoloid populations, positioning the Kami closer to Indo-Aryan-speaking groups. The revelation challenges the prevailing assumption of an exclusive Mongoloid ancestry for the Gorkha community. It draws attention to how earlier research ignored Indo-Aryan Gorkhas, which resulted in false conclusions about their Mongoloid lineage. It is crucial to recognise that cumulative studies of Gorkhas may have inherent potential biases. This underscores the important need for careful

methodological design in research attempts pertaining to this group. Identification of significant HLA gene associations with common chronic diseases like Type 2 Diabetes Mellitus and Hypertension emphasises the role of genetic determinants in disease susceptibility within the Kami population. Future research should prioritise larger sample sizes and rigorous methodologies to validate findings and enhance precision medicine strategies adapted to the genetic predispositions of Kami individuals. The examination of ABO blood group distributions within the Kami population reveals discernible patterns of genetic inheritance and associated phenotypic traits. The observed variations highlight the complex interaction of genetic and environmental factors, adding to our understanding of blood group diversity and its implications for population health and disease susceptibility.

The current investigation has yielded genetic profile of Kami population that may help in unravelling the intricate web of genetic strands that make up genetic profile of Gorkha community. Additionally, the finding of the study may be indispensable for population genetics, immunogenetics, and clinical applications. By elucidating the complexities of human genetic diversity, evolutionary history, and disease susceptibility among the Kami population, this study lays the groundwork for future investigations into population dynamics, disease associations, and genetic diversity.

References

[1] Gopalan, S., Smith, S. P., Korunes, K., Hamid, I., Ramachandran, S., and Goldberg, A. (2022). Human genetic admixture through the lens of population genomics. *Philosophical Transactions of the Royal Society B*, 377(1852), 20200410.

[2] Cavalli-Sforza, L. L. (1997). Genes, peoples, and languages. *Proceedings of the National Academy of Sciences*, 94(15), 7719–7724.

[3] Tamaki, K. and Jeffreys, A. J. (2005). Human tandem repeat sequences in forensic dna typing. *Legal Medicine*, 7(4), 244–250.

[4] Petersdorf, E. W., Anasetti, C., Martin, P. J., Gooley, T., Radich, J., Malkki, M., Woolfrey, A., Smith, A., Mickelson, E., and Hansen, J. A. (2004). Limits of hla mismatching in unrelated hematopoietic cell transplantation. *Blood*, 104(9), 2976–2980.

[5] Karas-Kuzelicki, N. and Mlinaric-Rascan, I. (2009). Individualization of thiopurine therapy: thiopurine s-methyltransferase and beyond. *Pharmacogenomics*, 10(8), 1309–1322.

[6] Consortium, H. P.-A. S., Abdulla, M. A., Ahmed, I., Assawamakin, A., Bhak, J., Brahmachari, S. K., Calacal, G. C., Chaurasia, A., Chen, C.-H., Chen, J., et al. (2009). Mapping human genetic diversity in asia. *Science*, 326(5959), 1541–1545.

[7] Landsteiner, K. (1901). Agglutination phenomena in normal human blood. *Wien Klin Wochenschr*, 14, 1132–4.

[8] Storry, J. and Olsson, M. L. (2009). The abo blood group system revisited: a review and update. *Immunohematology*, 25(2), 48–59.

[9] Tiercy, J.-M., Sanchez-Mazas, A., Excoffier, L., Shi-Isaac, X., Jeannet, M., Mach, B., and Langaney, A. (1992). Hla-dr polymorphism in a senegalese mandenka population: Dna oligotyping and population genetics of drb1 specificities. *American journal of human genetics*, 51(3), 592.

[10] Serjeantson, S. and Gao, X. (1995). Homo sapiens is an evolving species: origins of the austronesians. *The Austronesians*, pages 165–176.

[11] Parham, P., Arnett, K., Adams, E., Little, A.-M., Tees, K., Barber, L., Marsh, S., Ohta, T., Markow, T., and Petzl-Erler, M. (1997). Episodic evolution and turnover of hla-b in the indigenous human populations of the americas. *Tissue Antigens*, 50(3), 219–232.

[12] Imanishi, K., Igarashi, H., and Uchiyama, T. (1992). Relative abilities of distinct isotypes of human major histocompatibility complex class ii molecules to bind streptococcal pyrogenic exotoxin types a and b. *Infection and immunity*, 60(12), 5025–5029.

[13] Serjeantson, S., Ryan, D., and Thompson, A. (1982). The colonization of the pacific: the story according to human leukocyte antigens. *American Journal of Human Genetics*, 34(6), 904.

[14] Riley, E. and Olerup, O. (1992). Hla polymorphisms and evolution. *Immunology today*, 13(9), 333–335.

[15] Sanchez-Mazas, A., Fernandez-Viña, M., Middleton, D., Hollenbach, J. A., Buhler, S., Di, D., Rajalingam, R., Dugoujon, J.-M., Mack, S. J., and Thorsby, E. (2011). Immunogenetics as a tool in anthropological studies. *Immunology*, 133(2), 143–164.

[16] Meneganzin, A., Pievani, T., and Manzi, G. (2022). Pan-africanism vs. single-origin of homo sapiens: Putting the debate in the light of evolutionary biology. *Evolutionary Anthropology: Issues, News, and Reviews*, 31(4), 199–212.

[17] Mirazón Lahr, M. (2016). The shaping of human diversity: filters, boundaries and transitions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1698), 20150241.

[18] Maslin, M. A., Shultz, S., and Trauth, M. H. (2015). A synthesis of the theories and

concepts of early human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1663), 20140064.

[19] McDougall, I., Brown, F. H., and Fleagle, J. G. (2005). Stratigraphic placement and age of modern humans from kibish, ethiopia. *nature*, 433(7027), 733–736.

[20] White, T. D., Asfaw, B., and Suwa, G. (2005). Pliocene hominid fossils from gamedah, middle awash, ethiopia: aspects of hominid evolution. *Transactions of the Royal Society of South Africa*, 60(2), 79–83.

[21] Fu, Q., Mittnik, A., Johnson, P. L., Bos, K., Lari, M., Bollongino, R., Sun, C., Giemsch, L., Schmitz, R., Burger, J., et al. (2013). A revised timescale for human evolution based on ancient mitochondrial genomes. *Current biology*, 23(7), 553–559.

[22] Stringer, C. (2022). The development of ideas about a recent african origin for homo sapiens. *Journal of Anthropological Sciences= Rivista di Antropologia: JASS*, 100.

[23] Aiello, L. C. (1993). The fossil evidence for modern human origins in africa: a revised view. *American Anthropologist*, 95(1), 73–96.

[24] Wainscoat, J. S., Hill, A. V., Boyce, A., Flint, J., Hernandez, M., Thein, S. L., Old, J. M., Lynch, J., Falusi, A. G., Weatherall, D. J., et al. (1986). Evolutionary relationships of human populations from an analysis of nuclear dna polymorphisms. *Nature*, 319(6053), 491–493.

[25] Cann, R. L., Stoneking, M., and Wilson, A. C. (1987). Mitochondrial dna and human evolution. *Nature*, 325(6099), 31–36.

[26] Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., and Wilson, A. C. (1991). African populations and the evolution of human mitochondrial dna. *Science*, 253(5027), 1503–1507.

[27] Smith, F. H., Ahern, J. C., Janković, I., and Karavanić, I. (2017). The assimilation model of modern human origins in light of current genetic and genomic knowledge. *Quaternary International*, 450, 126–136.

[28] Mongiardino Koch, N., Garwood, R. J., and Parry, L. A. (2021). Fossils improve phylogenetic analyses of morphological characters. *Proceedings of the Royal Society B*, 288(1950), 20210044.

[29] Foley, R. A., Martin, L., Mirazón Lahr, M., and Stringer, C. (2016). Major transitions in human evolution.

[30] Ward, C. V. (2002). Interpreting the posture and locomotion of australopithecus afarensis: where do we stand? *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 119(S35), 185–215.

[31] Susman, R. L. and Stern, J. T. (1982). Functional morphology of homo habilis. *Science*, 217(4563), 931–934.

[32] Bramble, D. M. and Lieberman, D. E. (2004). Endurance running and the evolution of homo. *nature*, 432(7015), 345–352.

[33] Carotenuto, F., Tsikaridze, N., Rook, L., Lordkipanidze, D., Longo, L., Condemi, S., and Raia, P. (2016). Venturing out safely: The biogeography of homo erectus dispersal out of africa. *Journal of Human Evolution*, 95, 1–12.

[34] Johansson, S. (2014). The thinking neanderthals: What do we know about neanderthal cognition? *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6), 613–620.

[35] Paul, C. R. (2009). The fidelity of the fossil record: the improbability of preservation. *Palaeontology*, 52(3), 485–489.

[36] Benton, M. J. and Storrs, G. W. (1994). Testing the quality of the fossil record: paleontological knowledge is improving. *Geology*, 22(2), 111–114.

[37] Lieberman, B. S. (2002). Phylogenetic biogeography with and without the fossil record: gauging the effects of extinction and paleontological incompleteness. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 178(1-2), 39–52.

[38] Surovell, T. A., Finley, J. B., Smith, G. M., Brantingham, P. J., and Kelly, R. (2009). Correcting temporal frequency distributions for taphonomic bias. *Journal of archaeological*

Science, 36(8), 1715–1724.

[39] Bradley, B. J. (2008). Reconstructing phylogenies and phenotypes: a molecular view of human evolution. *Journal of Anatomy*, 212(4), 337–353.

[40] Jorde, L. B., Bamshad, M., and Rogers, A. R. (1998). Using mitochondrial and nuclear dna markers to reconstruct human evolution. *Bioessays*, 20(2), 126–136.

[41] Wilson, A. C. (1985). The molecular basis of evolution. *Scientific American*, 253(4), 164–175.

[42] Dos Reis, M., Thawornwattana, Y., Angelis, K., Telford, M. J., Donoghue, P. C., and Yang, Z. (2015). Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. *Current biology*, 25(22), 2939–2950.

[43] Stringer, C. (2000). Coasting out of africa. *Nature*, 405(6782), 25–27.

[44] Walter, R. C., Buffler, R. T., Bruggemann, J. H., Guillaume, M. M., Berhe, S. M., Negassi, B., Libsekal, Y., Cheng, H., Edwards, R. L., Von Cosel, R., et al. (2000). Early human occupation of the red sea coast of eritrea during the last interglacial. *Nature*, 405(6782), 65–69.

[45] Field, J. S. and Lahr, M. M. (2005). Assessment of the southern dispersal: Gis-based analyses of potential routes at oxygen isotopic stage 4. *Journal of World Prehistory*, 19, 1–45.

[46] Kardulias, P. N. and Hall, T. D. (2007). A world-systems view of human migration past and present: providing a general model for understanding the movement of people. In *Forum on Public Policy*, pages 1–33.

[47] Tishkoff, S. A. and Verrelli, B. C. (2003). Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual review of genomics and human genetics*, 4(1), 293–340.

[48] Fay, J. C. and Wu, C.-I. (1999). A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear dna variation. *Molecular*

biology and evolution, 16(7), 1003–1005.

[49] Hey, J. and Harris, E. (1999). Population bottlenecks and patterns of human polymorphism. *Molecular Biology and Evolution*, 16(10), 1423–1426.

[50] Barbujani, G. and Colonna, V. (2010). Human genome diversity: frequently asked questions. *Trends in Genetics*, 26(7), 285–295.

[51] Jobling, M. A. (2012). The impact of recent events on human genetic diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1590), 793–799.

[52] Yu, N., Chen, F.-C., Ota, S., Jorde, L. B., Pamilo, P., Patthy, L., Ramsay, M., Jenkins, T., Shyue, S.-K., and Li, W.-H. (2002). Larger genetic differences within africans than between africans and eurasians. *Genetics*, 161(1), 269–274.

[53] Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., and Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in africa. *Proceedings of the National Academy of Sciences*, 102(44), 15942–15947.

[54] Prugnolle, F., Manica, A., and Balloux, F. (2005). Geography predicts neutral genetic diversity of human populations. *Current Biology*, 15(5), R159–R160.

[55] Osada, N. (2015). Genetic diversity in humans and non-human primates and its evolutionary consequences. *Genes & genetic systems*, 90(3), 133–145.

[56] Johnston, H. R., Keats, B. J., and Sherman, S. L. (2019). Population genetics. In *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics*, pages 359–373. Elsevier.

[57] Darwin, C. and Wallace, A. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the proceedings of the Linnean Society of London. Zoology*, 3(9), 45–62.

[58] Darwin, C. (1859). On the origin of species: facsimile of the first edition.

[59] Sheftel, H., Szekely, P., Mayo, A., Sella, G., and Alon, U. (2018). Evolutionary trade-offs

and the structure of polymorphisms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1747), 20170105.

[60] Lenart, P., Bienertov-Vaask, J., and Berec, L. (2022). Learning mitigates genetic drift. *Scientific Reports*, 12(1), 20403.

[61] Loewe, L. and Hill, W. G. (2010). The population genetics of mutations: good, bad and indifferent.

[62] Tenaillon, O. and Matic, I. (2020). The impact of neutral mutations on genome evolvability. *Current Biology*, 30(10), R527–R534.

[63] Slatkin, M. (1985). Gene flow in natural populations. *Annual review of ecology and systematics*, 16(1), 393–430.

[64] Sengupta, D., Choudhury, A., Basu, A., and Ramsay, M. (2016). Population stratification and underrepresentation of indian subcontinent genetic diversity in the 1000 genomes project dataset. *Genome biology and evolution*, 8(11), 3460–3470.

[65] Mastana, S. S. (2014). Unity in diversity: an overview of the genomic anthropology of india. *Annals of human biology*, 41(4), 287–299.

[66] Bose, A., Platt, D. E., Parida, L., Drineas, P., and Paschou, P. (2021). Integrating linguistics, social structure, and geography to model genetic diversity within india. *Molecular biology and evolution*, 38(5), 1809–1819.

[67] Thanseem, I., Thangaraj, K., Chaubey, G., Singh, V. K., Bhaskar, L. V., Reddy, B. M., Reddy, A. G., and Singh, L. (2006). Genetic affinities among the lower castes and tribal groups of india: inference from y chromosome and mitochondrial dna. *BMC genetics*, 7, 1–11.

[68] Majumder, P. P. (1998). People of india: Biological diversity and affinities. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 6(3), 100–110.

[69] Dolgin, E. (2009). The history of two distinct lineages led to most modern-day indians. *Nature*, 10.

[70] Thapar, R. (2014). Can genetics help us understand indian social history? *Cold Spring Harbor Perspectives in Biology*, 6(11), a008599.

[71] Mishra, S. (1992). The age of the acheulian in india: new evidence. *Current Anthropology*, 33(3), 325–328.

[72] Papiha, S. (1996). Genetic variation in india. *Human biology*, pages 607–628.

[73] Bhasin, M. K. and Walter, H. (2001). *Genetics of castes and tribes of India*. Kamla-Raj Enterprises.

[74] Malhotra, K. (1978). Morphological composition of the people of india. *Journal of Human Evolution*, 7(1), 45–53.

[75] Shankarkumar, U., Ghosh, K., and Mohanty, D. (2002). Defining the allelic variants of hla a19 in the western indian population. *Human immunology*, 63(9), 779–782.

[76] Kashyap, V., Ashma, R., Gaikwad, S., Sarkar, B., and Trivedi, R. (2004). Deciphering diversity in populations of various linguistic and ethnic affiliations of different geographical regions of india: analysis based on 15 microsatellite markers. *Journal of Genetics*, 83, 49–63.

[77] Jaini, R., Naruse, T., Kanga, U., Kikkawa, E., Kaur, G., Inoko, H., and Mehra, N. (2002). Molecular diversity of the hla-a* 19 group of alleles in north indians: Possible oriental influence. *Tissue Antigens*, 59(6), 487–491.

[78] Shankarkumar, U., Sridharan, B., and Pitchappan, R. (2003). Hla diversity among nadars, a primitive dravidian caste of south india. *Tissue antigens*, 62(6), 542–547.

[79] Srivastava, S. K., Antal, P., Gál, J., Hullám, G., Semsei, Á. F., Nagy, G., Falus, A., and Buzás, E. I. (2011). Lack of evidence for association of two functional snps of chi3l1 gene (hc-gp39) with rheumatoid arthritis. *Rheumatology international*, 31, 1003–1007.

[80] Pal, T., Akbari, M. R., Sun, P., Lee, J., Fulp, J., Thompson, Z., Coppola, D., Nicosia, S., Sellers, T. A., McLaughlin, J., et al. (2012). Frequency of mutations in mismatch repair genes in a population-based study of women with ovarian cancer. *British journal*

of cancer, 107(10), 1783–1790.

[81] Dobzhansky, T. (1950). Mendelian populations and their evolution. *The American Naturalist*, 84(819), 401–418.

[82] Crawford, M. H. (1973). The use of genetic markers of the blood in the study of the evolution of human populations. *Methods and Theories of Anthropological Genetics*, 19, 38.

[83] Féral, J.-P. (2002). How useful are the genetic markers in attempts to understand and manage marine biodiversity? *Journal of experimental marine biology and ecology*, 268(2), 121–145.

[84] Hernández, C. L. (2023). Mitochondrial dna in human diversity and health: From the golden age to the omics era. *Genes*, 14(8), 1534.

[85] Pakendorf, B. and Stoneking, M. (2005). Mitochondrial dna and human evolution. *Annu. Rev. Genomics Hum. Genet.*, 6, 165–183.

[86] Mitchell, S. L., Goodloe, R., Brown-Gentry, K., Pendergrass, S. A., Murdock, D. G., and Crawford, D. C. (2014). Characterization of mitochondrial haplogroups in a large population-based sample from the united states. *Human genetics*, 133, 861–868.

[87] Wallace, D. C. (2013). Bioenergetics in human evolution and disease: implications for the origins of biological complexity and the missing genetic variation of common diseases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1622), 20120267.

[88] Kenney, M. C., Hertzog, D., Chak, G., Atilano, S. R., Khatibi, N., Soe, K., Nobe, A., Yang, E., Chwa, M., Zhu, F., et al. (2013). Mitochondrial dna haplogroups confer differences in risk for age-related macular degeneration: a case control study. *BMC medical genetics*, 14, 1–9.

[89] Herrnstadt, C. and Howell, N. (2004). An evolutionary perspective on pathogenic mtdna mutations: haplogroup associations of clinical disorders. *Mitochondrion*, 4(5-6), 791–798.

[90] Tilford, C. A., Kuroda-Kawaguchi, T., Skaletsky, H., Rozen, S., Brown, L. G., Rosenberg, M., McPherson, J. D., Wylie, K., Sekhon, M., Kucaba, T. A., et al. (2001). A physical map of the human y chromosome. *Nature*, 409(6822), 943–945.

[91] Bachtrog, D. and Charlesworth, B. (2001). Towards a complete sequence of the human y chromosome. *Genome biology*, 2, 1–5.

[92] Hammer, M. F. and Zegura, S. L. (1996). The role of the y chromosome in human evolutionary studies. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 5(4), 116–134.

[93] Hammer, M. F. and Zegura, S. L. (2002). The human y chromosome haplogroup tree: nomenclature and phylogeography of its major divisions. *Annual Review of Anthropology*, 31(1), 303–321.

[94] Butler, J. M. and Hill, C. R. (2012). Biology and genetics of new autosomal str loci useful for forensic dna analysis. *Forensic Science Review*, 24(1), 15.

[95] Edwards, A., Hammond, H. A., Jin, L., Caskey, C. T., and Chakraborty, R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, 12(2), 241–253.

[96] Hardy, O. J., Charbonnel, N., Fréville, H., and Heuertz, M. (2003). Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, 163(4), 1467–1482.

[97] Goldstein, D. B., Ruiz Linares, A., Cavalli-Sforza, L. L., and Feldman, M. W. (1995). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences*, 92(15), 6723–6727.

[98] Agarwal, S. K., Tan, F. K., and Arnett, F. C. (2008). Genetics and genomic studies in scleroderma (systemic sclerosis). *Rheumatic Disease Clinics of North America*, 34(1), 17–40.

[99] Thomson, G. (1991). 3 hla population genetics. *Baillière's clinical endocrinology and metabolism*, 5(2), 247–260.

[100] Pitchappan, R., Kakkanaiah, V., Rajashekar, R., Arulraj, N., and Muthukkaruppan, V. (1984). Hla antigens in south india: I. major groups of tamil nadu. *Tissue antigens*, 24(3), 190–196.

[101] Sinha, N., Chatterjee, B., and Satsangi, P. (1982). Status of paleontological researches in the northeastern states of india. *Rec. Geol. Surv. India*, 112(4), 66–88.

[102] Singh, B., Chamlagai, D., and Gurung, J. (2020). Hla-a-b and-drb1 distribution in kami: A caste population of gorkha community from the sub-himalayan region of west bengal, india. *Human Immunology*, 81(8), 395–396.

[103] Gahatraj, D. (2016). Identity formation and identity crisis: Nepalis in india ". *SouthAsian Journal of Multidisciplinary Studies*, pages 2349–785.

[104] Subba, T. B. (2009). *Indian Nepalis: issues and perspectives*. concept publishing company.

[105] Sinha, A. C. and Subba, T. B. (2003). *The Nepalis in Northeast India: The Community in Search of Indian Identity*. New Delhi: Indus.

[106] Chaudhuri, T., Mandal, T., Sen, T., and Taneja, V. (1995). Distribution of hla antigens in indian gurkha population. *The Indian Journal of Medical Research*, 101, 170–172.

[107] Debnath, M. and Chaudhuri, T. K. (2006). Study of genetic relationships of indian gurkha population on the basis of hla-a and b loci antigens. *International Journal of Human Genetics*, 6(2), 159–162.

[108] Preet, K., Malhotra, S., Shrivastava, P., Jain, T., Rawat, S., Varte, L. R., Singh, S., Singh, I., and Sarkar, S. (2016). Genetic diversity in gorkhas: an autosomal str study. *Scientific Reports*, 6(1), 32494.

[109] Collins, F. S. and Fink, L. (1995). The human genome project. *Alcohol health and research world*, 19(3), 190.

[110] Rhodes, D. and Trowsdale, J. (1999). Genetics and molecular genetics of the mhc. *Reviews in immunogenetics*, 1(1), 21–31.

[111] Trowsdale, J. (1993). Genomic structure and function in the mhc. *Trends in Genetics*, 9(4), 117–122.

[112] Rasaily, M. (1996). Biswakarma ko thar ra gotrawali: Biswa brahman varnan.

[113] Mizuki, N. and Kimura, M. (1996). Gene structure of the human mhc region. *Nihon Rinsho. Japanese Journal of Clinical Medicine*, 54(6), 1705–1717.

[114] Smith, M. P., Sansom, I. J., and Cochrane, K. D. (2001). The cambrian origin of vertebrates. *SYSTEMATICS ASSOCIATION SPECIAL VOLUME*, 61, 67–84.

[115] Castro, L. F. C., Furlong, R. F., and Holland, P. W. (2004). An antecedent of the mhc-linked genomic region in amphioxus. *Immunogenetics*, 55, 782–784.

[116] Kasahara, M., Hayashi, M., Tanaka, K., Inoko, H., Sugaya, K., Ikemura, T., and Ishibashi, T. (1996). Chromosomal localization of the proteasome z subunit gene reveals an ancient chromosomal duplication involving the major histocompatibility complex. *Proceedings of the National Academy of Sciences*, 93(17), 9096–9101.

[117] Kasahara, M., Nakaya, J., Satta, Y., and Takahata, N. (1997). Chromosomal duplication and the emergence of the adaptive immune system. *Trends in Genetics*, 13(3), 90–92.

[118] Janeway, C., Travers, P., Walport, M., Shlomchik, M., et al. (2001). *Immunobiology: the immune system in health and disease*, volume 2. Garland Pub. New York.

[119] Douillard, V., Castelli, E. C., Mack, S. J., Hollenbach, J. A., Gourraud, P.-A., Vince, N., and Limou, S. (2021). Approaching genetics through the mhc lens: tools and methods for hla research. *Frontiers in Genetics*, 12, 774916.

[120] Medhasi, S., Chantratita, N., et al. (2022). Human leukocyte antigen (hla) system: genetics and association with bacterial and viral infections. *Journal of Immunology Research*, 2022.

[121] Jiang, Y., Lu, J., Wu, Y.-E., Zhao, X., and Li, L. (2019). Genetic variation in the hla-g 3 utr 14–bp insertion/deletion and the associated cancer risk: evidence from 25 case–control studies. *Bioscience Reports*, 39(5), BSR20181991.

[122] Vandiedonck, C. and Knight, J. C. (2009). The human major histocompatibility complex as a paradigm in genomics research. *Briefings in Functional Genomics and Proteomics*, 8(5), 379–394.

[123] Dausset, J. (1958). Iso-leuco-anticorps. *Acta haematologica*, 20(1-4), 156–166.

[124] Shiina, T., Hosomichi, K., Inoko, H., and Kulski, J. K. (2009). The hla genomic loci map: expression, interaction, diversity and disease. *Journal of human genetics*, 54(1), 15–39.

[125] Lorente, L., Martín, M., Franco, A., Barrios, Y., Cáceres, J., Solé-Violán, J., Perez, A., y Ramos, J. M., Ramos-Gómez, L., Ojeda, N., et al. (2021). Hla genetic polymorphisms and prognosis of patients with covid-19. *Medicina intensiva*, 45(2), 96–103.

[126] Tsujimura, A., Ota, M., Katsuyama, Y., Sada, M., Miura, H., Matsumiya, K., Gotoh, R., Nakatani, T., Okuyama, A., and Takahara, S. (2002). Susceptibility gene for non-obstructive azoospermia located near hla-dr and-dq loci in the hla class ii region. *Human genetics*, 110, 192–197.

[127] Beck, S. and Trowsdale, J. (2000). The human major histocompatibility complex: lessons from the dna sequence. *Annual review of genomics and human genetics*, 1(1), 117–137.

[128] Kumánovics, A., Takada, T., and Lindahl, K. F. (2003). Genomic organization of the mammalian mhc. *Annual review of immunology*, 21(1), 629–657.

[129] Kulski, J. K., Suzuki, S., and Shiina, T. (2022). Human leukocyte antigen super-locus: nexus of genomic supergenes, snps, indels, transcripts, and haplotypes. *Human Genome Variation*, 9(1), 49.

[130] Horton, R., Wilming, L., Rand, V., Lovering, R. C., Bruford, E. A., Khodiyar, V. K., Lush, M. J., Povey, S., Talbot Jr, C. C., Wright, M. W., et al. (2004). Gene map of the extended human mhc. *Nature Reviews Genetics*, 5(12), 889–899.

[131] Schindler, E., Dribus, M., Duffy, B. F., Hock, K., Farnsworth, C. W., Gragert, L., and Liu, C. (2021). Hla genetic polymorphism in patients with coronavirus disease 2019 in midwestern united states. *Hla*, 98(4), 370–379.

[132] Kulski, J. K., Pfaff, A. L., Marney, L. D., Fröhlich, A., Bubb, V. J., Quinn, J. P., and Koks, S. (2023). Regulation of expression quantitative trait loci by sva retrotransposons within the major histocompatibility complex. *Experimental Biology and Medicine*, 248(23), 2304–2318.

[133] Ricaño-Ponce, I. and Wijmenga, C. (2013). Mapping of immune-mediated disease genes. *Annual review of genomics and human genetics*, 14, 325–353.

[134] Shiina, T., Hosomichi, K., Inoko, H., and Kulski, J. K. (2009). The hla genomic loci map: expression, interaction, diversity and disease. *Journal of human genetics*, 54(1), 15–39.

[135] Bharadwaj, M., Illing, P., Theodossis, A., Purcell, A. W., Rossjohn, J., and McCluskey, J. (2012). Drug hypersensitivity and human leukocyte antigens of the major histocompatibility complex. *Annual review of pharmacology and toxicology*, 52, 401–431.

[136] Howell, W. (2014). Hla and disease: guilt by association. *International journal of immunogenetics*, 41(1), 1–12.

[137] Radwan, J., Kohi, C., Ejsmond, M., Paganini, J., and Pontarotti, P. (2023). Integration of the immune memory into the pathogen-driven mhc polymorphism hypothesis. *HLA*, 102(6), 653–659.

[138] Jin, P. and Wang, E. (2003). Polymorphism in clinical immunology–from hla typing to immunogenetic profiling. *Journal of translational medicine*, 1, 1–11.

[139] Pitchappan, R., Kakkanaiah, V., Rajashekar, R., Arulraj, N., and Muthukkaruppan, V. (1984). Hla antigens in south india: I. major groups of tamil nadu. *Tissue antigens*, 24(3), 190–196.

[140] Agrawal, S., Srivastava, S., Borkar, M., and Chaudhuri, T. (2008). Genetic affinities of north and northeastern populations of india: inference from hla-based study. *Tissue Antigens*, 72(2), 120–130.

[141] Thomson, G. (1991). 3 hla population genetics. *Baillière's clinical endocrinology and metabolism*, 5(2), 247–260.

[142] Hedrick, P. W. and Thomson, G. (1983). Evidence for balancing selection at hla. *Genetics*, 104(3), 449–456.

[143] Robinson, J., Barker, D. J., Georgiou, X., Cooper, M. A., Flicek, P., and Marsh, S. G. (2020). Ipd-imgthla database. *Nucleic acids research*, 48(D1), D948–D955.

[144] Marsh, S. G., Albert, E., Bodmer, W., Bontrop, R., Dupont, B., Erlich, H., Fernández-Viña, M., Geraghty, D., Holdsworth, R., Hurley, C., et al. (2010). Nomenclature for factors of the hla system, 2010. *Tissue antigens*, 75(4), 291.

[145] Panneerchelvam, S. and Norazmi, M. N. (2012). Hla polymorphism in anthropology. *INTECH Open Access Publisher*, pages p1–18.

[146] Middleton, D. and Williams, F. (1997). A history of dna typing for hla. *HLA*, pages 61–84.

[147] Middleton, D., Williams, F., Hamill, M., and Meenagh, A. (2000). Frequency of hla-b alleles in a caucasoid population determined by a two-stage pcr-ssop typing strategy. *Human immunology*, 61(12), 1285–1297.

[148] Maiers, M., Gragert, L., and Klitz, W. (2007). High-resolution hla alleles and haplotypes in the united states population. *Human immunology*, 68(9), 779–788.

[149] Meyer, D., Single, R. M., Mack, S. J., Erlich, H. A., and Thomson, G. (2006). Signatures of demographic history and natural selection in the human major histocompatibility complex loci. *Genetics*, 173(4), 2121–2142.

[150] Parham, P. (1993). Hla, anthropology, and transplantation. In *Transplantation proceedings*, volume 25, pages 159–161.

[151] Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., Cann, H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L., et al. (2008). Worldwide human relationships inferred from genome-wide patterns of variation. *science*, 319(5866), 1100–1104.

[152] Auton, A., Bryc, K., Boyko, A. R., Lohmueller, K. E., Novembre, J., Reynolds, A., Indap,

A., Wright, M. H., Degenhardt, J. D., Gutenkunst, R. N., et al. (2009). Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome research*, 19(5), 795–803.

[153] Arrieta-Bolaños, E., Hernández-Zaragoza, D. I., and Barquera, R. (2023). An hla map of the world: A comparison of hla frequencies in 200 worldwide populations reveals diverse patterns for class i and class ii. *Frontiers in Genetics*, 14, 866407.

[154] Williams, F., Meenagh, A., Darke, C., Acosta, A., Daar, A., Gorodezky, C., Hammond, M., Nascimento, E., and Middleton, D. (2001). Analysis of the distribution of hla-b alleles in populations from five continents. *Human immunology*, 62(6), 645–650.

[155] Peterson, T. A., Bielawny, T., Lacap, P., Hardie, R.-A., Daniuk, C., Mendoza, L., Thavaneswaran, S., Kariri, T., Kimani, J., Wachihi, C., et al. (2014). Diversity and frequencies of hla class i and class ii genes of an east african population. *Open Journal of Genetics*, 2014.

[156] Al-Hassan, A. A., Al-Naseri, S., Al-Ghurabi, B. H., Al-Faham, M., Al-Nnema, A. J., and Shereef, S. M. (2005). Distribution of hla–antigens class i and ii in iraqi arab population. *Iraqi J Gastroenterol*, 5, 2–9.

[157] Ali, M. E., Ahmed, M., Alam, S., and Rahman, M. H. (2008). Hla-a,-b and-drb1 allele frequencies in the bangladeshi population. *Tissue Antigens*, 72(2), 115–119.

[158] Moatter, T., Aban, M., Tabassum, S., Shaikh, U., and Pervez, S. (2010). Molecular analysis of human leukocyte antigen class i and class ii allele frequencies and haplotype distribution in pakistani population. *Indian journal of human genetics*, 16(3), 149.

[159] Raza, A., Firasat, S., Khaliq, S., Abid, A., Shah, S. S., Mehdi, S. Q., and Mohyuddin, A. (2013). Hla class i and ii polymorphisms in the gujjar population from pakistan. *Immunological investigations*, 42(8), 691–700.

[160] Villanueva III, E. Q., Li, C.-C., Lin, P.-Y., and Yang, K.-L. (2020). Hla-a,-b, and-drb1 genotyping and haplotype frequencies among filipinos living in the national capital region of the philippines. *Human Immunology*, 81(8), 397–398.

[161] Nikbin, B., Nicknam, M. H., Hadinedoushan, H., Ansaripour, B., Moradi, B., Yekaninejad, M., Aminikhah, M., Ranjbar, M. M., and Amirzargar, A. (2017). Human leukocyte antigen (hla) class i and ii polymorphism in iranian healthy population from yazd province. *Iranian Journal of Allergy, Asthma and Immunology*, pages 1–13.

[162] Tuladhar, A., Shrestha, S., Raut, P., Bhandari, P., and Shrestha, P. (2013). Hla antigen distribution in renal transplant recipients and donors. *Journal of Nepal Health Research Council*.

[163] Shankarkumar, U., Sridharan, B., and Pitchappan, R. (2003). Hla diversity among nadars, a primitive dravidian caste of south india. *Tissue antigens*, 62(6), 542–547.

[164] Shankarkumar, U., Ghosh, K., and Mohanty, D. (2002). Defining the allelic variants of hla a19 in the western indian population. *Human immunology*, 63(9), 779–782.

[165] Thomas, R., Nair, S., and Banerjee, M. (2004). Hla-b and hla-c alleles and haplotypes in the dravidian tribal populations of southern india. *Tissue Antigens*, 64(1), 58–65.

[166] Menon, T. M. (1996). *The encyclopaedia of Dravidian tribes*. International School of Dravidian Linguistics.

[167] Pitchappan, R. (2002). Castes, migration, immunogenetics and infectious diseases in south india. *Public Health Genomics*, 5(3), 157–161.

[168] Papiha, S., Wentzel, J., Roberts, D., Rao, Y. S., and Habibullah, C. (1983). Hla antigen frequency in the koya tribe of andhra pradesh, india. *American Journal of Physical Anthropology*, 62(2), 147–150.

[169] Selvakumar, A., Raj, B. M., Damodaran, C., and Sekharan, P. C. (1987). Distribution of hla antigens in kotas and badagas of the nilgiri hills, south india. *American Journal of Physical Anthropology*, 74(1), 125–129.

[170] Selvakumar, A., Damodaran, C., and Sekharan, P. C. (1988). Distribution of hla antigens in the native south indian tamil hindus. *Tissue antigens*, 31(3), 136–140.

[171] Subramanian, V. and Damodaran, C. (1995). Distribution of hla antigens in sadhu chetty of tamil nadu, south india. *Anthropologischer Anzeiger*, pages 221–230.

[172] Thomas, R., Nair, S., and Banerjee, M. (2006). A crypto-dravidian origin for the nontribal communities of south india based on human leukocyte antigen class i diversity. *Tissue antigens*, 68(3), 225–234.

[173] Mehra, N., Taneja, V., Kailash, S., Raizada, N., and Vaidya, M. (1986). Distribution of hla antigens in a sample of the north indian hindu population. *Tissue Antigens*, 27(2), 64–74.

[174] Rani, R., Fernandez-Vina, M., and Stastny, P. (1998). Associations between hla class ii alleles in a north indian population. *Tissue antigens*, 52(1), 37–43.

[175] Rani, R., Mukherjee, R., and Stastny, P. (1998). Diversity of hla-dr2 in north indians: the changed scenario after the discovery of drb1* 1506. *Tissue Antigens*, 52(2), 147–152.

[176] Mehra, N., Verduijn, W., Taneja, V., Drabbels, J., Singh, S., and Giphart, M. (1991). Analysis of hla-dr2-associated polymorphisms by oligonucleotide hybridization in an asian indian population. *Human immunology*, 32(4), 246–253.

[177] Mehra, N. K. and Kanga, U. (2001). Molecular diversity of the hla-b27 gene and its association with disease. *Modern rheumatology*, 11(4), 275–285.

[178] Jaini, R., Naruse, T., Kanga, U., Kikkawa, E., Kaur, G., Inoko, H., and Mehra, N. (2002). Molecular diversity of the hla-a* 19 group of alleles in north indians: Possible oriental influence. *Tissue Antigens*, 59(6), 487–491.

[179] Jaini, R., Kaur, G., and Mehra, N. K. (2002). Heterogeneity of hla-drb1 04 and its associated haplotypes in the north indian population. *Human immunology*, 63(1), 24–29.

[180] Rani, R., Marcos, C., Lazaro, A., Zhang, Y., and Stastny, P. (2007). Molecular diversity of hla-a,-b and-c alleles in a north indian population as determined by pcr-ssop. *International journal of immunogenetics*, 34(3), 201–208.

[181] Babita, K. and Usha, D. (2004). Hla antigen distribution in sikhs from punjab, india. *International Journal of Human Genetics*, 4(2), 111–113.

[182] Shankarkumar, U., Ghosh, K., and Mohanty, D. (2000). Hla class i antigen profile among brahmins and related caste groups from mumbai, maharashtra, india. *Indian J Hum Genet*.

[183] Shankarkumar, U., Ghosh, K., and Mohanty, D. (2001). Hla antigen distribution in maratha community from mumbai, maharastra, india. *International Journal of Human Genetics*, 1(3), 173–177.

[184] Shankarkumar, U., Ghosh, K., and Mohanty, D. (2002). Defining the allelic variants of hla a19 in the western indian population. *Human immunology*, 63(9), 779–782.

[185] Shankarkumar, U., Devaraj, J., Ghosh, K., and Mohanty, D. (2003). Hla drb1 and dqb1 gene diversity in maratha community from mumbai india. *International Journal of Human Genetics*, 3(1), 39–43.

[186] Chhaya, S. (2005). Hla-b27 polymorphism in mumbai, western india. *Tissue Antigens*, 66(1), 48–50.

[187] Chhaya, S., Desai, S., and Saranath, D. (2010). Hla polymorphisms in sindhi community in mumbai, india. *International Journal of Immunogenetics*, 37(5), 373–377.

[188] Debnath, M. and Chaudhuri, T. (2006). Hla-a and hla-b distribution in toto–a vanishing sub-himalayan tribe of india. *Tissue Antigens*, 67(1), 64–65.

[189] Singh, B., Mallick, G. C., Bandopadhyay, S., Nayak, C. R., and Chaudhuri, T. K. (2009). Study of selected hla-a and-b antigens by pcr-ssp method in bengali population of siliguri and adjoining areas of west bengal. *International Journal of Human Genetics*, 9(3-4), 245–249.

[190] Dedhia, L., Vijapukar-Valluri, M., Gadekar, S., Mehta, P., Parekh, S., and Ghosh, K. (2015). Distribution of hla genes and disease predisposition in bengali speaking people from india. *indian journal of transplantation*, 9(1), 2–6.

[191] Agrawal, S., Srivastava, S., Borkar, M., and Chaudhuri, T. (2008). Genetic affinities of north and northeastern populations of india: inference from hla-based study. *Tissue Antigens*, 72(2), 120–130.

[192] Dendrou, C. A., Petersen, J., Rossjohn, J., and Fugger, L. (2018). Hla variation and disease. *Nature Reviews Immunology*, 18(5), 325–339.

[193] Reveille, J. D. (2006). Major histocompatibility genes and ankylosing spondylitis. *Best Practice & Research Clinical Rheumatology*, 20(3), 601–609.

[194] Holoshitz, J. (2013). The quest for better understanding of hla-disease association: scenes from a road less travelled by. *Discovery medicine*, 16(87), 93.

[195] Amirzargar, A., Yalda, A., Hajabolbaghi, M., Khosravi, F., Jabbari, H., Rezaei, N., Niknam, M., Ansari, B., Moradi, B., and Nikbin, B. (2004). The association of hla-drb, dqa1, dqb1 alleles and haplotype frequency in iranian patients with pulmonary tuberculosis. *The International Journal of Tuberculosis and Lung Disease*, 8(8), 1017–1021.

[196] Dubaniewicz, A., Lewko, B., Moszkowska, G., Zamorska, B., and Stepinski, J. (2000). Molecular subtypes of the hla-dr antigens in pulmonary tuberculosis. *International Journal of Infectious Diseases*, 4(3), 129–133.

[197] Ravikumar, M., Dheenadhayalan, V., Rajaram, K., Lakshmi, S. S., Kumaran, P. P., Paramasivan, C., Balakrishnan, K., and Pitchappan, R. (1999). Associations of hla-drb1, dqb1 and dpb1 alleles with pulmonary tuberculosis in south india. *Tubercle and Lung Disease*, 79(5), 309–317.

[198] Wong, S. H., Gochhait, S., Malhotra, D., Pettersson, F. H., Teo, Y. Y., Khor, C. C., Rautanen, A., Chapman, S. J., Mills, T. C., Srivastava, A., et al. (2010). Leprosy and the adaptation of human toll-like receptor 1. *PLoS pathogens*, 6(7), e1000979.

[199] Zhang, F.-R., Huang, W., Chen, S.-M., Sun, L.-D., Liu, H., Li, Y., Cui, Y., Yan, X.-X., Yang, H.-T., Yang, R.-D., et al. (2009). Genomewide association study of leprosy. *New England Journal of Medicine*, 361(27), 2609–2618.

[200] Langton, D. J., Bourke, S. C., Lie, B. A., Reiff, G., Natu, S., Darlay, R., Burn, J., and

Echevarria, C. (2021). The influence of hla genotype on the severity of covid-19 infection. *Hla*, 98(1), 14–22.

[201] Anzurez, A., Naka, I., Miki, S., Nakayama-Hosoya, K., Isshiki, M., Watanabe, Y., Nakamura-Hoshi, M., Seki, S., Matsumura, T., Takano, T., et al. (2021). Association of hla-drb1* 09: 01 with severe covid-19. *Hla*, 98(1), 37–42.

[202] Titkov, Y. S., Ziskina, R. A., Temirov, A. A., Gybladze, D. V., and Bubnova, L. N. (1993). Hla antigens in borderline and essential hypertension. *American journal of hypertension*, 6(10), 885–887.

[203] Mercier, P. and Jouve, R. (1979). Antigenes hla chez les sujets ayant une hypertension artérielle essentielle.

[204] Sengar, D., Couture, R., Jindal, S., and Catching, J. (1985). Histocompatibility antigens in essential hypertension and myocardial infarction. *Tissue antigens*, 26(3), 168–171.

[205] Motala, A. A., Busson, M., Al-Harbi, E. M., Khuzam, M. A., Al-Omari, E. M., Arekat, M. R., and Almawi, W. Y. (2005). Susceptible and protective human leukocyte antigen class ii alleles and haplotypes in bahraini type 2 (non-insulin-dependent) diabetes mellitus patients. *Clinical and Vaccine Immunology*, 12(1), 213–217.

[206] Acton, R., Bell, D., Go, R., Roseman, J., Tseng, M., and Louv, W. (1993). Association of hla phenotypes with hypertension in african americans and caucasoid americans with type ii diabetes, a population at risk for renal disease. In *Transplantation proceedings*, volume 25, pages 2400–2403.

[207] Nepom, G. T. and Kwok, W. W. (1998). Molecular basis for hla-dq associations with iddm. *Diabetes*, 47(8), 1177–1184.

[208] Ridgway, W. M. and Fathman, C. G. (1998). The association of mhc with autoimmune diseases: understanding the pathogenesis of autoimmune diabetes. *Clinical immunology and immunopathology*, 86(1), 3–10.

[209] Oldstone, M. B. (1998). Molecular mimicry and immune-mediated diseases. *The FASEB Journal*, 12(13), 1255.

[210] Yin, L., Dai, S., Clayton, G., Gao, W., Wang, Y., Kappler, J., and Marrack, P. (2013). Recognition of self and altered self by t cells in autoimmunity and allergy. *Protein & cell*, 4, 8–16.

[211] Almawi, W. Y., Wakim-Ghorayeb, S. F., Arekat, M. R., Najm, P., Keleshian, S. H., Al-Sayed, N., Blanchon, B., Samaha, H. R., and Irani-Hakime, N. (2006). Association of selective hla class ii susceptibility-conferring and protective haplotypes with type 2 diabetes in patients from bahrain and lebanon. *Clinical and vaccine immunology*, 13(11), 1296–1298.

[212] Mourant, A. E. (1983). Blood relations: Blood groups and anthropology. *(No Title)*.

[213] Bernstein, F. (1924). Ergebnisse einer biostatistischen zusam-menfassenden betrachtung über die erblichen blutstrukturen des menschen. *Klinische Wochenschrift*, 3(33), 1495–1497.

[214] Storry, J. R., Clausen, F. B., Castilho, L., Chen, Q., Daniels, G., Denomme, G., Flegel, W. A., Gassner, C., de Haas, M., Hyland, C., et al. (2019). International society of blood transfusion working party on red cell immunogenetics and blood group terminology: report of the dubai, copenhagen and toronto meetings. *Vox sanguinis*, 114(1), 95–102.

[215] Zerihun, T. and Bekele, S. (2016). Pattern of abo and rhesus blood groups distribution of five years survey in jimma town blood bank, south west ethiopia. *J Health Edu Res Dev*, 4(177), 2.

[216] Ghobadian, Z., Sayemiri, K., Zeinali, M., and Sajjadi, S. M. (2014). Distribution of abo and rh blood groups in a major ethnic group of the west iran, the kurdish population. *Asian Journal of Medical Sciences*, 5(3), 26–29.

[217] Dean, L. and Dean, L. (2005). *Blood groups and red cell antigens*, volume 2. NCBI Bethesda.

[218] Yamamoto, F.-i., McNeill, P. D., and Hakomori, S.-i. (1995). Genomic organization of human histo-blood group abo genes. *Glycobiology*, 5(1), 51–58.

[219] Patenaude, S. I., Seto, N. O., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic,

M. M., and Evans, S. V. (2002). The structural basis for specificity in human abo (h) blood group biosynthesis. *Nature structural biology*, 9(9), 685–690.

[220] Jaff, M. S. (2010). Abo and rhesus blood group distribution in kurds. *Journal of blood medicine*, pages 143–146.

[221] Chhetri, V. (2018). Frequency distribution of abo and rh blood group among blood donors at centre regional referral hospital, gelephu: A retrospective study. *J Med Sci Clin Res*, 6(4), 939–42.

[222] Sakal, I. D., Emberesh, R. A., Alhamoudi, A. A., Habhab, S. T., and Waddan, M. A. (2019). Prevalence and distribution of abo and rh (d) factor among blood donors in sabratha-libya.

[223] Niroula, D., Jha, M., Limbu, P., Pokhre, I., Yadav, S., and Mukhopadhyay, S. (2018). Ethnic variations of blood groups in a medical college of eastern nepal. *Kathmandu Univ Med J*, 61(1), 18–22.

[224] Iyiola, O., Igunnugbemi, O., Raheem, U., and Anifowoshe, A. (2011). Gene frequencies of abo and rh (d) blood group alleles in ilorin, north-central nigeria. *World J Biol Res*, 4(1), 6–14.

[225] Benahadi, A., Alami, R., Boulahdid, S., Adouani, B., Laouina, A., Mokhtari, A., Soulaymani, A., Khdija, H., and Benajiba, M. (2013). Distribution of abo and rhesus d blood antigens in morocco. *The Internet Journal of Biological Anthropology*, 6(1), 1–6.

[226] Canizalez-Román, A., Campos-Romero, A., Castro-Sánchez, J. A., López-Martínez, M. A., Andrade-Muñoz, F. J., Cruz-Zamudio, C. K., Ortíz-Espinoza, T. G., León-Sicairos, N., Gaudrón Llanos, A. M., Velázquez-Román, J., et al. (2018). Blood groups distribution and gene diversity of the abo and rh (d) loci in the mexican population. *BioMed research international*, 2018.

[227] Liu, J., Zhang, S., Wang, Q., Shen, H., Zhang, Y., and Liu, M. (2017). Frequencies and ethnic distribution of abo and rhd blood groups in china: a population-based cross-sectional study. *BMJ open*, 7(12), e018476.

[228] Torun, Y. A., Kaynar, L. G., Karakükcü, Ç., Yay, M., Kurnaz, F., Mutlu, H., Çetin, M., and Eser, B. (2012). Abo and rh blood group distribution in kayseri province, turkey. *Turkish Journal of Hematology*, 29(1), 97.

[229] Volken, T., Crawford, R. J., Amar, S., Mosimann, E., Tschaggelar, A., and Mansouri Taleghani, B. (2017). Blood group distribution in switzerland-a historical comparison. *Transfusion medicine and hemotherapy*, 44(4), 210–216.

[230] Lialiaris, T., Mantadakis, E., Kareli, D., Mpountoukas, P., Tsalkidis, A., and Chatzimichail, A. (2010). Frequency of genetic diseases and health coverage of children requiring admission in a general pediatric clinic of northern greece. *Italian journal of pediatrics*, 36, 1–5.

[231] Rehman, G. U. and Shi, H. (2020). Abo and rh (d) blood groups distribution in pakistan: a systematic review. *Forensic Res Criminol Int J*, 8(6), 237–244.

[232] Dewan, G. (2015). Comparative frequency and allelic distribution of abo and rh (d) blood groups of major tribal communities of southern bangladesh with general population and their determinants. *Egyptian Journal of Medical Human Genetics*, 16(2), 141–147.

[233] Periyavan, S., Sangeetha, S., Marimuthu, P., Manjunath, B., and Seema, D. (2010). Distribution of abo and rhesus-d blood groups in and around bangalore. *Asian journal of transfusion science*, 4(1), 41.

[234] Badge, S. A., Ovhal, A. G., Azad, K., and Meshram, A. T. (2017). Distribution of blood groups in blood donors in the blood bank of jagdalpur, bastar district, chhattisgarh. *Medical Journal of Dr. DY Patil University*, 10(1), 32–35.

[235] Rao, C. and Shetty, J. (2014). Frequency of abo and rhesus (d) blood groups in dakshina kannada district of karnataka-a study from rural tertiary care teaching hospital in south india. *Journal of Health and Allied Sciences NU*, 4(03), 057–060.

[236] Prasad, B. S., Faheem, M. K., Krishna, N. R., Sekhar, B., Vijayalakshmi, S., Preethi Lakshmi, A., et al. (2018). Demographic distribution and prevalence of abo and rhesus blood groups in blood donors: study from a tertiary care centre in southern region of andhra

pradesh. *IOSR Journal of Dental and Medical Sciences*, 17(7), 01–05.

[237] Suresh, B., Babu, K. S., Mouli, P. C., Arun, R., and Jothibai, D. (2015). Distribution of abo and rhesus (d) blood group antigens among blood donors at a tertiary care teaching hospital blood bank in south india. *Journal of Clinical and Scientific Research*, 4(2), 129–135.

[238] Raja, K. A., Dobariya, G. H., Unagar, C. A., Pandya, A. N., Patel, J. N., Wadhwani, S. J., et al. (2016). Frequency and distribution of abo and rh blood groups among blood donors in tertiary care hospital of south gujarat, india. *Int J Res Med Sci*, 4(12), 5377–5381.

[239] Barot, T., Patel, D., and Shah, R. (2020). Distribution of abo and rhesus (rh) blood groups among voluntary blood donors in central gujarat, india. *Int j Contemp Med Res*, 7(7), 61–63.

[240] Chavhan, A., Pawar, S., and Baig, M. (2011). Allelic frequency of abo and rh d blood group among the banjara backward caste of yavatmal district, maharashtra, india. *Nature Precedings*, pages 1–1.

[241] Giri, P. A., Yadav, S., Parhar, G. S., and Phalke, D. B. (2011). Frequency of abo and rhesus blood groups: a study from a rural tertiary care teaching hospital in india. *Int J Biol Med Res*, 2(4), 988–90.

[242] Sachdeva, A., Kumar, M., Lal, P., Kumar, D., Kumar, S., and Salhotra, V. (2018). Frequency distribution of abo and rhesus blood groups among mbbs students in a hilly state of north india. *Int J Recent Sci Res*, 9(1), 23601–23605.

[243] Sharma, D. C., Rai, S., Iyenger, S., Jain, B., and Sao, S. (2013). Prevalence and distribution of abo and rh-d antigens along with its subgroups & rare types in greater gwalior region.

[244] Rawat, S., BhagoraR, H., Mane, A., Kapoor, P., and Patil, A. (2019). Frequency and distribution of abo and rhesus blood group of blood donors. *JMSCR*, 7(11), 148–152.

[245] Kotwal, Y., Yatoo, G., and Ahmed Jan, F. (2017). Morbidity and mortality among neonates admitted to a neonatal intensive care unit of a tertiary care teaching hospital

of jammu and kashmir (india). *Neonat Pediatr Med*, 3(2), 136.

[246] Garg, P., Upadhyay, S., Chufal, S. S., Hasan, Y., and Tayal, I. (2014). Prevalance of abo and rhesus blood groups in blood donors: a study from a tertiary care teaching hospital of kumaon region of uttarakhand. *Journal of clinical and diagnostic research: JCDR*, 8(12), FC16.

[247] Behra, R. and Joshi, Y. (2013). Distribution of abo blood group and rh (d) factor in western rajasthan. *National journal of medical research*, 3(01), 73–75.

[248] Singh, S., Mishra, S. K., Kalhan, S., Sharma, P., Satarkar, R. N., Singh, P., and Singh, N. K. (2018). Abo and rhesus (d) blood group distribution among blood donors in rural south haryana (mewat-region): A 5 years retrospective study. *Annals of International Medical and Dental Research*, 4(3), 30.

[249] Sivagurunathan, A., Subbulakshmi, A., Veluramesh, S., Nisha, J. A., Kalarani, S., and Kasilakshmi, G. (2020). Distribution of abo & rh blood group in relation to dermatoglyphics and bmi. *World Scientific News*, (144), 226–281.

[250] Nanda, P., Tunya, H., Sharma, H., and Tayeng, R. (2021). Distribution of abo and rhesus blood groups among tribal students of arunachal pradesh, india. *Antrocom: Online Journal of Anthropology*, 17(1).

[251] Meitei, S. Y., Asghar, M., Achoubi, N. D., Murry, B., Saraswathy, K. N., and Sachdeva, M. P. (2010). Distribution of abo and rh (d) blood groups among four populations in manipur, north east india. *Anthropological notebooks*, 16(2).

[252] Ranjan, R. P., Singh, D., and Kumar, A. Heterogeneity and distribution patterns of abo and rh phenotypes in the madhepura, bihar.

[253] Sodhi, B. S. (2020). Frequency and distribution of abo and rh blood groups among blood donors in a tertiary care hospital in unnao district, uttar pradesh. *Indian Journal of Basic & Applied Medical Research*, 9(3).

[254] Gurung, Y. (2019). Distribution of abo blood groups and allele frequency in major ethnic groups of darjeeling and kurseong hills, west bengal, india. *Int J Recent Sci Res*, 10,

33533–7.

[255] Singh, B. (2017). Distribution of abo blood groups and rhesus factor percentage frequencies among the population of sikkim, india.

[256] Nandakumar, R., Raghu, P., Deva, A., and Prabhakaran, K. (2015). Abo and rh (d) blood group polymorphms in soliga tribal population of tamil nadu, india. *Journal of Life Sciences*, 7(1-2), 18–18.

[257] Patidar, G. K. and Dhiman, Y. (2021). Distribution of abo and rh (d) blood groups in india: A systematic review. *ISBT science series*, 16(1), 37–48.

[258] Nag, I. and Das, S. S. (2012). Abo and rhesus blood groups in potential blood donors at durgapur steel city of the district of burdwan, west bengal. *Asian journal of transfusion science*, 6(1), 54–55.

[259] Talukdar, L. and Sarma, U. (2014). Frequency of major blood group antigens among blood donors at a tertiary level hospital in north-east india. *IJSR*, 3(9), 2130–2.

[260] Kumbhakar, D. (2016). Distribution of abo and rh-d blood groups in the cachar district of barak valley of assam: implication for regional blood transfusion service. *JOURNAL OF EVOLUTION OF MEDICAL AND DENTAL SCIENCES-JEMDS*, 5(58), 4017–4021.

[261] Coggon, D., Barker, D., and Rose, G. (2009). *Epidemiology for the Uninitiated*. John Wiley & Sons.

[262] Dedhia, L., Vijapukar-Valluri, M., Gadekar, S., Mehta, P., Parekh, S., and Ghosh, K. (2015). Distribution of hla genes and disease predisposition in bengali speaking people from india. *indian journal of transplantation*, 9(1), 2–6.

[263] Arnaiz-Villena, A., Palacio-Grüber, J., Muniz, E., Campos, C., Alonso-Rubio, J., Gomez-Casado, E., Salih, S. F., Martin-Villa, M., and Al-Qadi, R. (2017). Genetic hla study of kurds in iraq, iran and tbilisi (caucasus, georgia): relatedness and medical implications. *PLoS One*, 12(1), e0169929.

[264] Raza, A., Firasat, S., Khaliq, S., Abid, A., Shah, S. S., Mehdi, S. Q., and Mohyuddin,

A. (2013). Hla class i and ii polymorphisms in the gujjar population from pakistan. *Immunological investigations*, 42(8), 691–700.

[265] Mansour, I., Klayme, S., Naman, R., Loiselet, J., Hallé, L., and Kaplan, C. (1996). Hla phenotype polymorphism in the lebanese population. *Transfusion clinique et biologique*, 3(5), 289–295.

[266] Benjeddou, M., Leat, N., and Davison, S. (2006). Use of molecular genetics and historical records to reconstruct the history of local communities. *African Journal of Biotechnology*, 5(25).

[267] Dedhia, L., Gadekar, S., Mehta, P., and Parekh, S. (2015). Hla haplotype diversity in the south indian population and its relevance. *Indian Journal of Transplantation*, 9(4), 138–143.

[268] Singh, B., Mallick, G. C., Bandopadhyay, S., Nayak, C. R., and Chaudhuri, T. K. (2009). Study of selected hla-a and-b antigens by pcr-ssp method in bengali population of siliguri and adjoining areas of west bengal. *International Journal of Human Genetics*, 9(3-4), 245–249.

[269] Mohyuddin, A., Ayub, Q., Khaliq, S., Mansoor, A., Mazhar, K., Rehman, S., and Mehdi, S. Q. (2002). Hla polymorphism in six ethnic groups from pakistan. *Tissue antigens*, 59(6), 492–501.

[270] Machulla, H. K., Batnasan, D., Steinborn, F., Uyar, F. A., Saruhan-Direskeneli, G., Oguz, F. S., Carin, M. N., and Dorak, M. T. (2003). Genetic affinities among mongol ethnic groups and their relationship to turks. *Tissue Antigens*, 61(4), 292–299.

[271] Bugawan, T., Mack, S., Stoneking, M., Saha, M., Beck, H., and Erlich, H. (1999). Hla class i allele distributions in six pacific/asian populations: evidence of selection at the hla-a locus. *Tissue Antigens*, 53(4), 311–319.

[272] Xingpei, X., Senbin, L., Chaoying, W., and Quanhui, L. (1986). Study on the association of hla with pulmonary tuberculosis. *Immunological Investigations*, 15(4), 327–332.

[273] Kim, S., Choi, I., Dahlberg, S., Nisperos, B., Kim, J., and Hansen, J. (1987). Hla and leprosy in koreans. *Tissue Antigens*, 29(3), 146–153.

[274] Mohd-Yusuf, Y., Phipps, M. E., Chow, S. K., and Yeap, S. S. (2011). Hla-a* 11 and novel associations in malays and chinese with systemic lupus erythematosus. *Immunology letters*, 139(1-2), 68–72.

[275] Ma, S. L., Tang, N. L. S., Tam, C. W. C., Lui, V. W. C., Suen, E. W. C., Chiu, H. F. K., and Lam, L. C. W. (2008). Association between hla-a alleles and alzheimer's disease in a southern chinese community. *Dementia and geriatric cognitive disorders*, 26(5), 391–397.

[276] Kang, E. H., Kim, J. Y., Takeuchi, F., Kim, J. W., Shin, K., Lee, E. Y., Lee, Y. J., Lee, E. B., Park, M. H., and Song, Y. W. (2011). Associations between the hla-a polymorphism and the clinical manifestations of behcet's disease. *Arthritis research & therapy*, 13, 1–9.

[277] Clayton, J., Lonjou, C., and Whittle, D. (1997). Allele and haplotype frequencies for hla loci in various ethnic groups. *Genetic diversity of HLA. Functional and medical implications*, 1, 665–820.

[278] Thomas, R., Nair, S., and Banerjee, M. (2006). A crypto-dravidian origin for the nontribal communities of south india based on human leukocyte antigen class i diversity. *Tissue antigens*, 68(3), 225–234.

[279] Mishra, V. C., Chandra, D., Raina, V., and Sharma, G. (2021). Analysis of hla-b allele polymorphism in north indian population: Experience at tertiary care centre. *Gene Reports*, 22, 100996.

[280] Li, X.-F., Zhang, X., Chen, Y., Zhang, K.-L., Liu, X.-J., and Li, J.-P. (2014). An analysis of hla-a,-b, and-drb1 allele and haplotype frequencies of 21,918 residents living in liaoning, china. *PloS one*, 9(4), e93082.

[281] Londono, J., Santos, A. M., Peña, P., Calvo, E., Espinosa, L. R., Reveille, J. D., Vargas-Alarcon, G., Jaramillo, C. A., Valle-Oñate, R., Avila, M., et al. (2015). Analysis of hla-b15 and hla-b27 in spondyloarthritis with peripheral and axial clinical patterns. *BMJ*

open, 5(11), e009092.

[282] Chhaya, S., Desai, S., and Saranath, D. (2010). Hla polymorphisms in sindhi community in mumbai, india. *International Journal of Immunogenetics*, 37(5), 373–377.

[283] Bamshad, M. J., Watkins, W. S., Dixon, M. E., Jorde, L. B., Rao, B. B., Naidu, J. M., Prasad, B. R., Rasanayagam, A., and Hammer, M. F. (1998). Female gene flow stratifies hindu castes. *Nature*, 395(6703), 651–652.

[284] Wang, H.-W., Li, Y.-C., Sun, F., Zhao, M., Mitra, B., Chaudhuri, T. K., Regmi, P., Wu, S.-F., Kong, Q.-P., and Zhang, Y.-P. (2012). Revisiting the role of the himalayas in peopling nepal: insights from mitochondrial genomes. *Journal of human genetics*, 57(4), 228–234.

[285] Balakrishnan, K., Rathika, C., Kamaraj, R., Subashini, R., Saravanan, M., Asha, K., Kananan, M., VinothKumar, R., Manikandan, T., Dhivakar, M., et al. (2012). Gradients in distribution of hla–drb1 alleles in castes and tribes of south india. *International Journal of Human Genetics*, 12(1), 45–55.

[286] Malavige, G., Rostron, T., Seneviratne, S., Fernando, S., Sivayogan, S., Wijewickrama, A., and Ogg, G. (2007). Hla analysis of sri lankan sinhalese predicts north indian origin. *International journal of immunogenetics*, 34(5), 313–315.

[287] Quelvennec, E., Bera, O., Cabre, P., Alizadeh, M., Smadja, D., Jugde, F., Edan, G., and Semana, G. (2003). Genetic and functional studies in multiple sclerosis patients from martinique attest for a specific and direct role of the hla-dr locus in the syndrome. *Tissue antigens*, 61(2), 166–171.

[288] Garavito, G., Yunis, E. J., Egea, E., Ramirez, L. A., Malagon, C., Iglesias, A., De La Cruz, O. F., Uribe, O., Navarro, E., Martinez, P., et al. (2004). Hla-drb1 alleles and hla-drb1 shared epitopes are markers for juvenile rheumatoid arthritis subgroups in colombian mestizos. *Human immunology*, 65(4), 359–365.

[289] Manoussakis, M. N., Georgopoulou, C., Zintzaras, E., Spyropoulou, M., Stavropoulou, A., Skopouli, F. N., and Moutsopoulos, H. M. (2004). Sjögren's syndrome associated with systemic lupus erythematosus: clinical and laboratory profiles and comparison with

primary sjögren's syndrome. *Arthritis & Rheumatism*, 50(3), 882–891.

[290] Gonzalez-Galarza, F. F., McCabe, A., Santos, E. J. M. d., Jones, J., Takeshita, L., Ortega-Rivera, N. D., Cid-Pavon, G. M. D., Ramsbottom, K., Ghattaoraya, G., Alfirevic, A., et al. (2020). Allele frequency net database (afnd) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic acids research*, 48(D1), D783–D788.

[291] Grifoni, A., Weiskopf, D., Arlehamn, C. S. L., Angelo, M., Leary, S., Sidney, J., Frazier, A., Phillips, E., Mallal, S., Mack, S. J., et al. (2018). Sequence-based hla-a, b, c, dp, dq, and dr typing of 714 adults from colombo, sri lanka. *Human immunology*, 79(2), 87–88.

[292] Narayan, S., Maiers, M., Halagan, M., Sathishkannan, A., Naganathan, C., Madbouly, A., and Periathiruvadi, S. (2018). Human leucocyte antigen (hla)-a,-b,-c,-drb1 and-dqb1 haplotype frequencies from 2491 cord blood units from tamil speaking population from tamil nadu, india. *Molecular biology reports*, 45, 2821–2829.

[293] Debnath, M., Palanichamy, M. G., Mitra, B., Jin, J.-Q., Chaudhuri, T. K., and Zhang, Y.-P. (2011). Y-chromosome haplogroup diversity in the sub-himalayan terai and duars populations of east india. *Journal of human genetics*, 56(11), 765–771.

[294] Kivisild, T., Rootsi, S., Metspalu, M., Mastana, S., Kaldma, K., Parik, J., Metspalu, E., Adojaan, M., Tolk, H.-V., Stepanov, V., et al. (2003). The genetic heritage of the earliest settlers persists both in indian tribal and caste populations. *The American Journal of Human Genetics*, 72(2), 313–332.

[295] Bamshad, M., Kivisild, T., Watkins, W. S., Dixon, M. E., Ricker, C. E., Rao, B. B., Naidu, J. M., Prasad, B. R., Reddy, P. G., Rasanayagam, A., et al. (2001). Genetic evidence on the origins of indian caste populations. *Genome research*, 11(6), 994–1004.

[296] Gayden, T., Cadenas, A. M., Regueiro, M., Singh, N. B., Zhivotovsky, L. A., Underhill, P. A., Cavalli-Sforza, L. L., and Herrera, R. J. (2007). The himalayas as a directional barrier to gene flow. *The American Journal of Human Genetics*, 80(5), 884–894.

[297] Fornarino, S., Pala, M., Battaglia, V., Maranta, R., Achilli, A., Modiano, G., Torroni, A.,

Semino, O., and Santachiara-Benerecetti, S. A. (2009). Mitochondrial and y-chromosome diversity of the tharus (nepal): a reservoir of genetic variation. *BMC Evolutionary Biology*, 9, 1–16.

[298] Thangaraj, K., Chaubey, G., Kivisild, T., Selvi Rani, D., Singh, V. K., Ismail, T., Carvalho-Silva, D., Metspalu, M., Bhaskar, L., Reddy, A. G., et al. (2008). Maternal footprints of southeast asians in north india. *Human heredity*, 66(1), 1–9.

[299] Bhatia, K., Patel, M., and Gorogo, M. (1984). Type 2 (non-insulin-dependent) diabetes mellitus and hla antigens in papua new guinea. *Diabetologia*, 27, 370–372.

[300] Omar MAK, S. M., Hammond MG and AC, A. (1985). Hla antigens and non-insulin-dependent diabetes mellitus in young south african indians. *South African Medical Journal*, 67(4), 130–132.

[301] Williams, R., Muller, Y., Hanson, R., Knowler, W., Mason, C., Bian, L., Ossowski, V., Wiedrich, K., Chen, Y., Marcovina, S., et al. (2011). Hla-drb1 reduces the risk of type 2 diabetes mellitus by increased insulin secretion. *Diabetologia*, 54, 1684–1692.

[302] Chinniah, R., Vijayan, M., Sivanadham, R., Ravi, P. M., Panneerselvam, D., and Karuppiah, B. (2016). Association of hla-a, b, drb1* and dqb1* alleles and haplotypes in south indian t2dm patients. *Gene*, 592(1), 200–208.

[303] Hatmal, M. M., Alshaer, W., Mahmoud, I. S., Al-Hatamleh, M. A., Al-Ameer, H. J., Abuyaman, O., Zihlif, M., Mohamud, R., Darras, M., Al Shhab, M., et al. (2021). Investigating the association of cd36 gene polymorphisms (rs1761667 and rs1527483) with t2dm and dyslipidemia: Statistical analysis, machine learning based prediction, and meta-analysis. *Plos one*, 16(10), e0257857.

[304] Jan, A., Saeed, M., Afridi, M. H., Khuda, F., Shabbir, M., Khan, H., Ali, S., Hassan, M., Akbar, R., et al. (2021). Association of hla-b gene polymorphisms with type 2 diabetes in pashtun ethnic population of khyber pakhtunkhwa, pakistan. *Journal of diabetes research*, 2021.

[305] Vidan-Jeras, B., Gregorič, A., Jurca, B., Jeras, M., and Bohinjec, M. (2000). Possible influence of genes located on chromosome 6 within or near to the major histocompatibility complex on development of essential hypertension. *Pflügers Archiv*, 439, r060–r062.

[306] Shao, J., Hu, D., Chen, A., Wang, W., Chen, J., Sun, B., and Cai, X. (2004). Alleles of hla class ii drb1 of patients with essential hypertension in yunnan hans. *Zhonghua yi xue yi Chuan xue za zhi= Zhonghua Yixue Yichuanxue Zazhi= Chinese Journal of Medical Genetics*, 21(3), 286–287.

[307] Diamantopoulos, E. J., Andreadis, E. A., Vassilopoulos, C. V., Tabacila-Kakou, M. G., Sfakianakis, M. E., Tarassi, K. E., and Papasteriades, C. A. (2000). Distribution of different hla antigens in greek hypertensives according to the angiotensin-converting enzyme genotype. *American journal of hypertension*, 13(4), 438–441.

[308] Shkhvatsabaia, I., Rudnev, V., IuI, S., GIu, D., and Osipov, S. (1988). Various immunological aspects of essential and symptomatic hypertension. *Biulleten'Vsesoiuznogo kardiologicheskogo nauchnogo tsentra AMN SSSR*, 11(1), 7–12.

[309] Gladman, D. D., Kung, T. N., Siannis, F., Pellett, F., Farewell, V. T., and Lee, P. (2005). Hla markers for susceptibility and expression in scleroderma. *The Journal of rheumatology*, 32(8), 1481–1487.

[310] Basu, D., Datta, S. S., Montemayor, C., Bhattacharya, P., Mukherjee, K., and Flegel, W. A. (2018). Abo, rhesus, and kell antigens, alleles, and haplotypes in west bengal, india. *Transfusion Medicine and Hemotherapy*, 45(1), 62–66.

[311] Glass, R. I., Holmgren, J., Haley, C. E., Khan, M., Svennerholm, A., Stoll, B. J., Hossain, K. B., Black, R. E., Yunus, M., and Barua, D. (1985). Predisposition for cholera of individuals with o blood group possible evolutionary significance. *American journal of epidemiology*, 121(6), 791–796.

[312] Rowe, J. A., Opi, D. H., and Williams, T. N. (2009). Blood groups and malaria: fresh insights into pathogenesis and identification of targets for intervention. *Current opinion in hematology*, 16(6), 480–487.

[313] Shrestha, L., Malla, U., and Mahotra, N. B. (2013). Abo and rh blood groups and their ethnic distribution in a teaching hospital of kathmandu, nepal. *Journal of Nepal Medical Association*, 52(190).

[314] Garcia-Hejl, C., Martinaud, C., De Rudnicki, S., Sanmartin, N., Clavier, B., Chianea, D., and Vest, P. (2014). Blood group antigens frequencies in kabul, afghanistan. *Transfusion and Apheresis Science*, 50(2), 307–308.

[315] Chandimal, K., Yasawardene, S., and Adikari, G. (2011). The distribution of abo and rhesus (rh) blood groups in purana inhabitants of sigiriya, sri lanka. 12th Annual Research Symposium, University of Kelaniya.

[316] Zaman, R., Parvez, M., Jakaria, M., and Sayeed, M. A. (2015). Study of abo and rh-d blood group among the common people of chittagong city corporation area of bangladesh. *Journal of Public Health and Epidemiology*, 7(9), 305–310.

[317] Kumar, S., Modak, P. K., Ali, S. H., Barpanda, S., Gusain, V. S., and Roy, R. (2018). A retrospective study: Abo and rh phenotype blood group distribution among blood donors in hnb base hospital, srinagar, uttarakhand, india. *Journal of Family Medicine and Primary Care*, 7(1), 34–38.

[318] Sigamani, K. and Gajulapalli, S. P. (2022). An insight into the distribution of allele frequency of abo and rh (d) blood grouping system among blood donors in a tertiary care hospital in chengalpattu district of south india. *Cureus*, 14(4).

[319] Sidhu, S. (2003). Distribution of the abo blood groups and rh (d) factor among the scheduled caste population of punjab. *The Anthropologist*, 5(3), 203–204.

[320] Shrivastava, S., Gahine, R., and Kapse, V. (2015). Abo, rhesus blood group and allele frequency in and around raipur (chattisgarh state), india. *International Journal of Current Research and Review*, 7(17), 52.

[321] Al-Ani, L., Mahmood, H. M., and Abdulhaleem, N. (2020). Genetically determined abo and (rh) rhesus blood groups and their associations with diabetes mellitus. *Sys Rev Pharm*, 11(4), 604–608.

[322] Zafar, M., Masud, S., Malik, M. M., Latif, M. H., and Muhammad Sharjeel Noor, A. (2021). Gene frequency distribution of abo and rh-d blood group alleles in multan division (punjab), pakistan. *Pure and Applied Biology (PAB)*, 10(4), 1173–1180.

[323] Amin, M. R., Susanto, D., and Naher, L. (2015). Distribution pattern of abo and rh blood groups and their allelic frequencies among different ethnic groups in malaysia. *Asian Journal of Medical Sciences*, 6(4), 25–29.

[324] Allawati, M., Al-Kalbani, L., and Al-Balushi, T. (2021). Abo and rh (d) phenotypes, allele frequencies and estimated genotypes in omanis: a retrospective study from armed forces hospital. *JMSCR*, 9(4), 1–7.

6 माइल, सामदुर, तादोंग -737102
गंगटोक, सिक्किम, भारत
फोन-03592-251212, 251415, 251656
टेलीफैक्स -251067
वेबसाइट - www.cus.ac.in

6th Mile, Samdur,Tadong -737102
Gangtok, Sikkim, India
Ph. 03592-251212, 251415, 251656
Telefax: 251067
Website: www.cus.ac.in

सिक्किम विश्वविद्यालय
SIKKIM UNIVERSITY

(भारत के संसद के अधिनियम द्वारा वर्ष 2007 में स्थापित और नैक (एनएएसी) द्वारा वर्ष 2015 में प्रत्यायित केंद्रीय विश्वविद्यालय)
(A central university established by an Act of Parliament of India in 2007 and accredited by NAAC in 2015)

Ref. No: SU/2012/REG-03/IEC/2323/95 **Date: 16 April 2018**

TO WHOM IT MAY CONCERN

This is to certify that the proposal for issuing of Ethical clearance certificate for synopsis for Ph.D Registration entitled ***"Study of Human Leukocyte Antigen gene complex and ABO system among the Indo-aryan Kami population of Sub-Himalayan belt of Darjeeling and adjoining areas"*** by Mr. Dependra Chamlagai, Ph.D Student, Deptt. of Zoology has been examined by the Institutional Ethical Committee of Sikkim University in its 6th meeting held on 9th April 2018 and has approved for issue of Ethical clearance certificate subject to following guidelines issued by Medical Council of India.

Tej Kishen Kaul
(T.K.Kaul)
Member Secretary
Institutional Ethical Committee

Informed Consent Form for Participant

Part-I: Information Sheet

It is with reference to the research project sponsored by DST, Govt of India, I am here to give you information and to invite you to voluntarily participate in this research. I am Dependra Chamlagai, working for Sikkim University to carry out this research project under the guidance of Dr. Bisu Singh, Assistant professor, Department of Zoology, Sikkim University. I am trained and competent in all manners to conduct required works pertaining to this research.

Purpose of the research: This study is entitled "Study of Human Leukocyte Antigen (HLA) among the Indo Aryan Kami population of Sub- Himalayan belt of Darjeeling and adjoining area".Till date there are no studies done to find out the descendency of Kami population. This HLA based study will provide us with discernible data to establish the phylogeny of Kami population. Since HLA is also associated with protection and susceptibility to many diseases. We want to find out the prevalent HLA associated diseases in this population and also the diseases to which the population is more susceptible.

Type of Research Intervention: This research involves drawing out of 2 ml blood. Your participation in this research is entirely voluntary. It is your choice whether to participate or not.

Procedures and Protocol:

Description of the Process- For this study blood samples would be collected by the vein puncture method from the basilic/Median cubital or Cephalic veins in the arm from each participant.

Side Effects and Risks- Process does not have any side effects and does not involve any risk to the health of the participant.

Benefits - The study may help to generate HLA data for Kami community which may serve as a preliminary reference in transplantation procedures. It will also help us to know about diseases prevalent in the population which are associated with HLA complex. This study may also suggest us on presence of HLA alleles in population, which are known to provide protection to some diseases.

Reimbursements- This research does not involve any reimbursements.

Confidentiality: The result of the study will be kept fully confidential and it will be used only for the research purpose and not for the medical or therapeutic purpose and the data will be presented at the population level only.

PART II: Demographic details:

Place of blood collection:

Serial No.-……. Date: ……

Name: ……..

Address:……

DEMOGRAPHIC DETAILS

Date of birth: …… Age: ……

Sex: M/F: …… Ethnicity: ……

Subcast: …….. Occupation: ……

Income:

Height: ……. Weight: …….

Education: <class VIII/Class VIII/Madhyamik/H.S./Graduation/Masters/other:

Marital status: Married/Unmarried/Widow/Divorced

Migration (if any): ……..

Any autoimmune disease in the family: ……

Any sort of substance abuse: …….

Any medical condition: …….

Under Medication (If Any): ..... Blood Group :

Food habit:

Surgeries involving arm - if any (Mastectomy etc.,)

Special remarks:

Pedigree relationships:

Part III: Certificate of Consent

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions that I have asked have been answered to my satisfaction. I consent voluntarily to participate as a participant in this research.

Name of Participant_____

Signature of Participant _____

Date _____

If illiterate

I have witnessed the accurate reading of the consent form to the potential participant, and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

Print name of witness_____ AND Thumb print of participant

Signature of witness _____

Date _____

Statement by the researcher/person taking consent

I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understands that the following will be done:

1.

2.

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Print Name of Researcher/person taking the consent_____

Signature of Researcher /person taking the consent_____

Date _____

HLA-Ready Gene ABDR CE 0123 IVD LOT G954018 2019-12

Auswerteschema / Worksheet-1

Name / *Name*	Vorname / *First name*	Geburtsdatum / *Date of birth*
Proben-Nr. / *Sample I.D.*	Herkunft / *Origin*	Diagnose / *Diagnosis*
Test-Durchführung / *-Performance*	Datum / *Date*	Unterschrift / *Signature*
Test-Überprüfung / *-Review*	Datum / *Date*	Unterschrift / *Signature*

Bemerkungen / *Comments*

HLA-A
HLA-B
HLA-DRB
Ergebnis / *Result*

Photo

Mix # 4 (multiplex): 230 bp: **A*03**:01,...,11:130, 30:89, 32:04, 36:02; 90 bp: **A*02:53N**, (24:240N)
Mix # 8 (multiplex): 195 bp: ..., **A*23**:01,... *crossreactive to* 18:27, weak crossreactive to B*15:84, (B*35:23, 35:258), 40:100, (B*44:10), 47:01-03,06-08, (47:09), 82:01,02 // (C*04:13), C*04:34,(C*04:58,61,68),113,122,212, (C*04:160, C*07:08,108) 95 bp: ..., **A*23**:01..., **24**:02,...
Mix # 12 (multiplex): 225 bp: ..., **A*03**:01,..., ...; 170 bp: ..., **A*25**:01,..., **26**:01,..., **43**:01, **66**:01,04,...
Mix # 13 (multiplex): 200 bp: A*11:11, **43**:01; 170 bp: **A*66:02**,03
Mix # 16 (multiplex): 240 bp: **A*01**:01,... , *02:146, *03:41,57,107, *26:20, **36**:01..., *74:10, **80**:01,03 ; 90 bp: A*03,43,82, *11:113, **30**:01...
Mix # 17 (multiplex): 245 bp: **A*02:17**,..., *11:139, *23:04, **24:03**,..., *29:07,49, *31:29 *crossreactive to* B*18:27; 125 bp: A*01:07, **02**:01,..., *03:01:03,..., *23:21, *24:124, *26:07:02,..., *29:14, **30**:01,..., **31**:01,..., *33:01:02,..., *34:04, *68:30, *74:04 *crossreactive to* B*15:82
Mix # 22 (multiplex): 165 bp: A*02:01,..., *24:14,291, **69**:01,...; 100 bp: **A*24:09N** (C*06:46N ?) 165bp: kreuzreaktiv mit / *crossreactive to* B*18:59
Mix # 23 kreuzreaktiv mit / *crossreactive to* C*16:67
Mix # 24 (multiplex): 155 bp: **A*80**:01; 70/ 75 bp: A*02:55, *26:03,05,...,*68:05,...

Mix # 26 (multiplex): 210bp: **B*13:02,..,*39:06,..,*40:06,..,*41..,*45..,*51:05,..,*54..,*55:02,..,*56..,*59..,*73..** ; 120bp: **B*07:02..** 120bp: kreuzreaktiv mit / *crossreactive to* C*03:260, *04:77
Mix # 27 kreuzreaktiv mit / *crossreactive to* C*06:147
Mix # 29 (multiplex): 265bp: **B*14..** ; 175bp: **B*08..,*14..,*15:03,..,*18..,*37..,*38..,*39..,*48..** 265bp: kreuzreaktiv mit / *crossreactive to* C*01:30, *12:87
175bp: kreuzreaktiv mit / *crossreactive to* C*06:147
Mix # 32 kreuzreaktiv mit / *crossreactive to* C*01:73
Mix # 33 kreuzreaktiv mit / *crossreactive to* C*07:46
Mix # 34: kreuzreaktiv mit / *crossreactive to* A*30:64 / C*01:32, *02:56, *03:102,263, *06:20, *07:81,168,450, *12:50
Mix # 36 kreuzreaktiv mit / *crossreactive to* C*02:62, *03:14,26,40,42,43,119,175, 240, 250, *03:279, *04:80,100,110, *05:10,*06:14,72,143, *07:10,43,196,367, *08:44,81, 82, *12:44, *15:04,05,09,19,23,25,27,29,30,36,46 ,54,59,61,66,69,70,76,77,90,104,105Q,107,108,110,111,115N-117
Mix # 37 (multiplex): 395 bp: B*15:16,17,...; 190 bp: **B*15:12**,14,... 190 bp: kreuzreaktiv mit / *crossreactive to* A*24:174 , C*02:82, *06:165, *14:61
Mix # 38 (multiplex): 435 bp: B*15:13,16,17,...; 185 bp: B*51:11N
Mix # 39 kreuzreaktiv mit / *crossreactive to* A*24:02:37,*26:01:16
Mix # 41 kreuzreaktiv mit / *crossreactive to* A*02:598, *26:68,100, C*02:56, *03:102,263, *07:81,168
Mix # 43 gelegentlich schwacher Hintergrund / *occasional weak background*
Mix # 45 (multiplex): 210 bp: B*38..,*39..,*67..; 130 bp: **B*40..,*41..,*45..,*50..** 210 bp: kreuzreaktiv mit / *crossreactive to* C*07:137
Mix # 46 kreuzreaktiv mit / *crossreactive to* A*23:01:18, *25:01:08
Mix # 49 kreuzreaktiv mit / *crossreactive to* C*17:01:10
Mix # 50 (multiplex): 310 bp: B*13..,**40:02**,..,**44..,*47..**; 125 bp: **B*40:02**,..
Mix # 52 (multiplex): 320 bp: **B*07..**,...,**27..**,...,**42..**,...,**46..**,**54..**,**55..**,**56..**,**67..**,**73..**,**81..**,**82..** ; 180 bp: B*08:21,...,**15:09..**,**35..**,...,**40:05..**,...,**44:10..**,...,**51..**,**52..**,...,**56..**,**57..**,...,*78...,*82..
Mix # 53 kreuzreaktiv mit / *crossreactive to* C*15:25
Mix # 55 kreuzreaktiv mit / *crossreactive to* C*01:73
Mix # 56 (multiplex): 230bp: B*40:100,...,**47**:01,... kreuzreaktiv mit / *crossreactive to* C*04:32,77,*18:03 105bp: B*15:08,11,...
Mix # 61 kreuzreaktiv mit / *crossreactive to* C*06:147 gelegentlich schwacher Hintergrund / *occasional weak background*
Mix # 62 kreuzreaktiv mit / *crossreactive to* C*16:85
Mix # 65 kreuzreaktiv mit / *crossreactive to* A*23:01:14
Mix # 66 (multiplex): 175bp: B*45..,*49..,*50..,*56..,*82.. ; 95bp: B*15:16,..,*57.. 175bp: evt. kreuzreaktiv mit / *might be crossreactive to* C*15:15
Mix # 70 (multiplex): 540 bp: B*44:06,..., *51:23,...; 245 bp: **B*82:01**,02,... 245 bp: evt. kreuzreaktiv mit / *might be crossreactive to* C*14:61
Mix # 72 gelegentlich schwacher Hintergrund / *occasional weak background*

Mix # 79 (multiplex): 175 bp: DRB1*03:10, **04**:05,..., **13**:03,..., **14:01**,07,...; 85 bp: DRB1**03:01**,04,...,*11:59, *13:27,...
Mix # 82 (multiplex): 240 bp: DRB1**07:01**...; 150 bp: DRB1*11:16..., **13:01**..., **14:24**
Mix # 84 (multiplex): 195 bp: DRB1**09:01:02**..., *13:27..., DRB5*01:13; 70 bp: DRB1*04:15..., *08:31..., **11:01:01**..., *12:04..., *16:16...
Mix # 87 (multiplex): 235 bp: DRB1**14:01**; 110 bp: DRB1*08:32, **12**:01-13,...
Mix # 93 (multiplex): 225 bp: DRB1*08:20, *11:23,..., *13:13,..., *14:03,12,27,...; 150 bp: DRB1*08:09,..., *14:15,...
Mix # 95 (multiplex): 215 bp: **DRB4*01:01**,..., **02:01N**, *03:01N; 130 bp: **DRB4*01:03:01:02N**
Mix # 96 (multiplex): 265 bp: **DRB5***01...*02...(DRB1*09:07); 180 bp: **DRB5*01:08N**

Inno-Train Diagnostik GmbH · Niederhöchstädter Straße 62 · D-61476 Kronberg/Taunus · Germany
Tel.: 0049-(0)6173 / 6079-30 · Fax: 0049-(0)6173 / 6079-50 · E-mail: info@inno-train.de · www.inno-train.de

HLA-Ready Gene ABDR

CE 0123 IVD

LOT G954018

2019-12

Auswerteschema / *Worksheet-2*

	H	G	F	E	D	C	B	A
HLA-A*	**1**	**2**	**3**	**4***	**5**	**6**	**7**	**8***
1	**Negative Control** 430 bp NC	**01**.. / ... 800 bp IC 80 bp	**02**.. / ... 800 bp IC 135 bp	L: **03** / s: **02:53N** 800 bp IC 230 bp **90 bp**	... / **11**.. / ... 800 bp IC 215bp	02:17,... / **23**:01:03,... / ... 800 bp IC 210 bp	/ **24**.. / ... 800 bp IC 70 bp	L: ... / **23** B*47... (w) / B*82 (w) s: ... / **23**.., / **24**.., / 31:08 / ... 800 bp IC **195 bp** **95 bp**
	9	**10**	**11**	**12***	**13***	**14**	**15**	**16***
2	... / **25**.. / **26**.. / **34**.. / **66**.. / **68**.. / **69**.. 800 bp IC 200 bp	**02:03**.., / **25**.. / .. **66:02**.. 1070 bp IC 475, 455 bp	... / **26** / ... 800 bp IC 75 bp	L: ... / **03**.. /... s: ... / **25**.., / **26** / .. / **43** / **66**:01:04,... 800 bp IC 225 bp 170 bp	L: ... / **43** s: **66**:02... 800 bp IC 230 bp 170 bp	... / **11**.. / ... / **25**:02... / ... / **34**.. / **66**:01:04,... 800 bp IC 95 bp	**29**.. 800 bp IC 200 bp	L: **01**.. / 26:20 / **36**.. / 74:10 / **80**:01,03 s: .. / **30**.. 800 bp IC **240 bp** **90 bp**
	17*	**18**	**19**	**20**	**21**	**22***	**23**	**24***
3	L: 02:17... / ... / **24**:03 / ... s: ... / **02**,.. / ... / **30**.. / **31**.. / ... 800 bp IC **245 bp** **125 bp**	... / **32**.. 800 bp IC 260 bp	**33**.. / ... 1070 bp IC 200 bp	... / **74**.. 800 bp IC 155 bp	02:10,... / ... / **68**.. 800 bp IC 140 bp	L: **02**.. / ... / **69** s: **24:09N** 800 bp IC **165 bp** **100 bp**	.. / **03**.. / 11:04... / **29**.. / **31**:01:02.. / **32**.. / **33**.. / **34**.. / **36**.. / **68**.. / **74**.. 800 bp IC 185 bp	L: **80**.. s: ... / 26:03,05 .. / ... / 68:05,... 800 bp IC **155 bp** **70, 75 bp**
HLA-B*	**25**	**26***	**27**	**28**	**29***	**30**	**31**	**32***
4	1070 bp	L: .. / **13**:02,../ .. / **39**:06,../ **40**:06.../ **41**.. / .. / **45**.. / .. / **51**:05,../ .. / **54**,../ **55**:02,.. **56**.. / **59**.. / **73**.. / .. s: **07**... 800 bp IC **210 bp** **120 bp**	**08**.. / ... 800 bp IC 210 bp	**13**.. / ... 800 bp IC 145 bp	L: **14**.. / ... s: .. / **08**.. / **14**.. / **15**:03.../ **18**.. / .. / **37** / **38**.. / **39**../ .. / **48**../ ...	... / **14**:01.07N,.../ ... / **38**:01 .../ **39**:01,03,.../ ..	**14**:02.../ ... 38:05 / 39:04 800 bp IC 185 bp	.. / **15**:01-02,04-.../ ... / **46**:01 .../ ... 800 bp IC 130 bp
	33*	**34**	**35**	**36***	**37***	**38***	**39***	**40**
5	**15**:03,09,10,18,23,.../ ... 800 bp IC 325 bp	.. / **15**.., / **18**.. / **27:06**.. / .. / **39**:**15**.. / **44**.. / .. / **46**.. .. / **53**.. / .. / **57**.. / **58**.. 1070 bp IC 85 bp	... / **08**:04,17,54 / ... / **15**:03,... / .. / **39**:**02**,08,.. / ... / **48**.. 800 bp IC 105 bp	... / **35**../ .../ **53**../.../ **57**../ **58**.. 800 bp IC 100 bp	L: **15:16,17**.... s: **15:12,14**,.. / ... 800 bp IC **395 bp** **190 bp**	L: **15**:13,16,17,... / s: **51:11N** 800 bp IC **435 bp** **185 bp**	**18**.. / ... 800 bp IC 130 bp	**27**... / ... 800 bp IC 155 bp
	41*	**42**	**43***	**44**	**45***	**46**	**47**	**48**
6	.. / **13**../.. / **15**.. / **18**../ .. / **35**.. / **37**.. / **38**.. / **39**.. / .. / **44**.. / **46**../ **47**../ **51:06**.. / **53**.. / .. / **58**.. / **67**.. / **82**.. / 83 800 bp IC 140 bp	.. / **18**../ **35**../ .../ 78.. 800 bp IC 130 bp	40../ 41.../ .../ 45../ 49../ 50... 1070 bp IC 315 bp	.../ **37**.. / ... 800 bp IC 130 bp	L: ... **38** ... / ... **39** s: ... **40** ... / ... **41** ... / ... **45** ... / **50** .. 800 bp IC **210 bp** **130 bp**	**38**.. 800 bp IC 110 bp	.../ **14**.../ **38**.../ **39**.. 800 bp IC 400 bp	**07**../ **08**../ **14**../ **39**../ **40**../ **41**../ **42**../ **45**../ **48**../ **50**../ **67**../ 73../ 800 bp IC 425 bp
	49*	**50***	**51**	**52***	**53***	**54**	**55***	**56***
7	**40**:01,07,.../ .../ **48**:01(w),03,.../ **81**.. (w) 800 bp IC 165 bp	L: **13**../ .../ **40**:02../ **44**../ **47**../ ... s: ... / **40**:02,../ ... 800 bp IC **310 bp** **125 bp**	../ **08**.. / ../ **41**.. / **42**.. / .. 800 bp IC 280 bp	L: 07,/ 27,/ 42,/ 46,/ 54,/ 55,/ 56,/ 67,/ 73,/ 81,/ 82,/ **83**, s: .. / 15:09 / .. / 35.. / .. / 40:05,.. / .. / 44:10.. / .. / 51.. / 52.. / .. / 56.. / 57.. / .. / 78.. / 82.. 1070 bp IC 320 bp 180 bp	../ **44**../ .../ 83 800 bp IC 255 bp	.../ **45**../ ... / **50**:**02** / .. 800 bp IC 250 bp	**46**.. 800 bp IC 125 bp	L: **47**.. s: **15:08**,11.. 800 bp IC **230 bp** **105 bp**
	57	**58**	**59**	**60**	**61***	**62**	**63**	**64**
8	.../ **46**.. / **81**... 800 bp IC 175 bp	.../ **49**.. 800 bp IC 140 bp	../ **49**../ **50**../ **51**../ **52**../ **55**./ .. / 78 800 bp IC 205 bp	.../ **27**:**02**,30 /.../ **51**../ **52**./ **53**../ **57**../ **58**.. 800 bp IC 430 bp	../ **08**../ **14**:**02**../ **15**:**02**../ **35**../ **51**../ **53**../ **54**../ **55**../ **56**../ **59**../ **78**../ **82**.. 1070 bp IC 195 bp	**13**../ **15**../ .. / **40** / **41**../ **44**../ **45**../ **47**../ **49**../ **50**../ **52**.. / ... 800 bp IC 200 bp	.../ **54**... 800 bp IC 215 bp	.../ **54**../ **55**.../.... / **59**../ ... 800 bp IC 235 bp
	65*	**66***	**67**	**68**	**69**	**70***	**71**	**72**
9	.. / **51**../ **52**:**01**:**02**../ .. 800 bp IC 205 bp	L: .. / **45**../ .../ **49**../ **50**../ .. / **56**../.. / **82**.. s: 15:16,17,... / **57**.. 800 bp IC **175 bp** **95 bp**	**50**... 800 bp IC 110 bp	.../ **51**../ .../ **53**../ **59** 800 bp IC 85 bp	**73**:01 800 bp IC 295 bp	L: 44:06../ ... s: **82**.. 1070 bp IC 540 bp 245 bp	**Bw4** 800 bp IC 355 bp	**Bw6** 800 bp IC 355 bp
DRB1*	**73**	**74**	**75**	**76**	**77**	**78**	**79***	**80**
10	800 bp	**01**:01,02,04,... 430 bp IC 195 bp	**01**:**03**,39N,42 430 bp IC 195 bp	**15**.. 430 bp IC 210 bp	*15:50N, 15:80N **16**.. 430 bp IC 200, 205, 220 bp	**03**.. / ... 430 bp IC 225 bp	L: **04**:05 ... / 13:03... / **14**:01,07,...,54, s: **03**:01,04,.../ ... 430 bp IC **175 bp** **85 bp**	**03**:02,03,.../ ... / **14**:02,... 430 bp IC 90 bp
	81	**82***	**83**	**84***	**85**	**86**	**87***	**88**
11	... / **04**.. / ... 430 bp IC 205 bp	L: **07**.. s: ... / **13**:01.., / **14**:**24** 800 bp IC 240 bp 150 bp	**08**.. / ... / **14**:**04**.. 430 bp IC 220, 165 bp	L: **09**.. s: .. / **11**.. 430 bp IC 195 bp 70 bp	**10**.. 430 bp IC 205 bp	.../ **11**.. 430 bp IC 185, 180, 170 bp	L: **14**:**01** s: ... / **12**.. 430 bp IC **235 bp** **110 bp**	.../ **11**.. / **13**.. / **14**:**03**,... 430 bp IC 215 bp
	89	**90**	**91**	**92**	**93***	**94**	**95***	**96***
12	**08**.. / **11**.. / **12**:**02**,... **13**.. / **14**:15,... 430 bp IC 205, 215, 195 bp	**03**.. / ... / **13**.. / **14**:**02**,... 430 bp IC 180 bp	.../ **14**:**01**,**04**,...:**54**,... 800 bp IC 210, 205 bp	.../ **14**:**02**,**06**,.../ ... 430 bp IC 150, 145 bp	L: .... / **14**:**03**,... s: 08 09,... / 14:15,... 430 bp IC **225 bp** **150 bp**	**DRB3*** 430 bp IC 230 bp	L: **DRB4*** / **02**:**01N** / 03:01N s: DRB4*01:03:01:02N 430 bp IC 215 bp 130 bp	L: **DRB5*** s: DRB5*01:08N 430 bp IC **265 bp** **180 bp**

[1] einige seltenere Allele ausgenommen (vollständige Liste: Interpretation -/Specificity Table) / *some rare alleles excluded (complete list: Interpretation -/Specificity Table)*

Fettdruck: Allel(gruppen) mit Häufigkeiten von mind. 0,001 / ***Bold letters***: *parts of allele groups or alleles with min. frequency of 0,001* (Ref.: www.allelefrequencies.net)

IC: PCR Produkte der internen Kontrollprimermixe / *PCR products of the internal control primer mixes*

w: schwache Bande / *weak band*

Multiplex Mixes: 2 spezifische Banden / 2 *specific bands*: L(arge), s(mall)

Inno-Train Diagnostik GmbH · Niederhöchstädter Straße 62 · D-61476 Kronberg/Taunus · Germany
Tel.: 0049-(0)6173 / 6079-30 · Fax: 0049-(0)6173 / 6079-50 · E-mail: info@inno-train.de · www.inno-train.de

HLA-Ready Gene ABDR CE 0123 IVD

List of common and well-documented HLA-A alleles (CWD)

Allele Specificity	CWD*	Ser. Type
A*01:01:01:01	++	A1
A*01:02	++	A1
A*01:03	++	A1
A*01:04N	+	Null
A*01:06	+	-
A01:09	+	-
A*01:12	+	-
A*01:17	+	-
A*01:25	+	-
A*02:01:01:01	++	A2
A*02:01:04	++	A2
A*02:02	++	A2
A*02:03:01	++	A203
A*02:04	++	A2
A*02:05:01	++	A2
A*02:06:01	++	A2
A*02:07:01	++	A2
A*02:08	++	A2
A*02:09	++	A2
A*02:10	++	A210
A*02:11:01	++	A2
A*02:12	+	A2
A*02:13	+	A2
A*02:14	++	A2
A*02:16	+	A2
A*02:17:01	++	A2
A*02:18	+	A2
A*02:19	+	-
A*02:20:01	++	A2
A*02:21	+	A2
A*02:22:01	++	A2
A*02:24:01	++	A2
A*02:25	+	A2
A*02:27	+	-
A*02:29	+	A2
A*02:30	++	-
A*02:33	+	-
A*02:34	+	A2
A*02:35:01	+	A2
A*02:36	+	-
A*02:38	+	-
A*02:39	+	-
A*02:44	+	-
A*02:45	+	-
A*02:49	+	A2
A*02:53N	++	Null
A*02:58	+	-
A*02:60:01	++	-
A*02:63	+	-
A*02:64	+	-
A*02:74:01	+	A2
A*02:84	+	-
A*02:85	+	-
A*02:86	+	-
A*02:87	+	-
A*02:93	+	-
A*02:96	+	-
A*02:119	+	-
A*02:122	+	-
A*02:123	+	-
A*02:131	+	-
A*02:137	+	-

Allele Specificity	CWD*	Ser. Type
A*03:01:01:01	++	A3
A*03:01:01:03	+	A3
A*03:01:02	+	A3
A*03:01:03	+	A3
A*03:02:01	++	A3
A*03:05:01	++	A3
A*03:06	+	-
A*03:07	+	-
A*03:08	+	-
A*03:09	+	-
A*03:17:01	+	-
A*03:21N	+	-
A*03:22:01	+	-
A*03:26	+	-
A*03:50	+	-
A*03:108	+	-
A*11:01:01	++	A11
A*11:02:01	++	A11
A*11:03	+	A11
A*11:04	++	A11
A*11:05	+	A11
A*11:08	+	A11
A*11:09	+	A11
A*11:10	+	A11
A*11:12	+	A11
A*11:19	+	A11
A*11:29	+	-
A*11:30	+	-
A*23:01:01	++	A23(9)
A*23:03:01	+	-
A*23:04	+	A23(9)
A*23:05	+	-
A*23:06	+	-
A*23:13	+	-
A*23:15	+	-
A*23:17	+	-
A*23:19Q	+	-
A*24:02:01:01	++	A24(9)
A*24:02:01:02L	++	-
A*24:02:02	+	A24(9)
A*24:03:01	++	A2403
A*24:04	+	A24(9)
A*24:05	+	A24(9)
A*24:06	+	A24(9)
A*24:07	++	A24(9)
A*24:08	+	A24(9)
A*24:09N	+	Null
A*24:10	++	A2403
A*24:11N	+	Null
A*24:13:01	+	A24(9)
A*24:14	+	A24(9)
A*24:15	+	-
A*24:17	++	A24(9)
A*24:20	++	A24(9)
A*24:22	+	A9
A*24:23	+	A2403
A*24:25	++	-
A*24:26	+	-
A*24:29	+	-
A*24:30	+	-
A*24:31	+	-
A*24:32	+	-

Allele Specificity	CWD*	Ser. Type
A*24:33	++	A2403
A*24:35	+	-
A*24:37	+	A24(9)
A*24:38	+	-
A*24:43	+	A24(9)
A*24:47	+	-
A*24:53	+	-
A*24:56	+	-
A*24:57	+	-
A*24:58	+	A24(9)
A*24:64	+	-
A*24:72	+	-
A*24:81	+	-
A*24:90N	+	-
A*24:95	+	-
A*25:01:01	++	A25(10)
A*25:02	+	A25(10)
A*26:01:01	++	A26(10)
A*26:02	+	A26(10)
A*26:03:01	++	A26(10)
A*26:07:01	+	A26(10)
A*26:07:02	+	A26(10)
A*26:08	++	A26(10)
A*26:09	+	A26(10)
A*26:12	+	-
A*26:13	+	-
A*26:14	+	-
A*26:15	+	A10
A*26:16	+	-
A*26:17	+	-
A*26:18	+	-
A*26:20	+	-
A*26:31	+	-
A*29:01:01:01	++	A29(19)
A*29:02:01:01	++	A29(19)
A*29:04	+	-
A*29:09	+	-
A*29:10	+	A29(19)
A*29:12	+	-
A*30:01:01	++	A30(19)
A*30:02:01	++	A30(19)
A*30:03	+	A30(19)
A*30:04:01	++	A30(19)
A*30:07	+	-
A*30:08	+	-
A*30:09	+	-
A*30:10	+	-
A*30:11:01	+	A30(19)
A*30:15	+	A30(19)
A*30:16	+	-
A*31:01:02	++	A31(19)
A*31:02	+	A31(19)
A*31:03	+	-
A*31:04	+	A31(19)
A*31:06	+	-
A*31:08	+	-
A*31:09	+	-
A*31:11	+	-
A*31:12	+	A31(19)
A*31:13	+	-
A*31:15	+	-
A*31:16	+	-

Allele Specificity	CWD*	Ser. Type
A*31:18	+	-
A*31:20	+	-
A*31:26	+	-
A*32:01:01	++	A32(19)
A*32:02	+	A32(19)
A*32:03	+	-
A*32:04	+	-
A*32:06	+	-
A*32:07	+	-
A*32:08	+	-
A*33:01:01	++	A33(19)
A*33:03:01	++	A33(19)
A*33:05	+	A33(19)
A*33:09	+	A19
A*34:01:01	++	A34(10)
A*34:02:01	++	A34(10)
A*34:05	++	-
A*36:01	++	A36
A*36:02	+	-
A*36:03	+	A36
A*36:04	+	A36
A*43:01	++	A43
A*66:01	++	A66(10)
A*66:02	++	A66(10)
A*66:03	+	A10
A*66:04	+	-
A*68:01:01:01	++	A68(28)
A*68:01:02	++	A68(28)
A*68:02:01:01	++	A68(28)
A*68:02:01:02	+	A68(28)
A*68:03:01	++	A68(28)
A*68:05	++	A68(28)
A*68:06	+	-
A*68:07	+	-
A*68:08:01	+	A68(28)
A*68:10	+	A28
A*68:11N	+	Null
A*68:12	+	A28
A*68:15	+	-
A*68:17	+	A28
A*68:20	+	-
A*68:22	+	-
A*68:24	+	-
A*68:25	+	-
A*68:27	+	-
A*68:30	+	-
A*68:31	+	-
A*68:35	+	-
A*68:37	+	-
A*68:40	+	-
A*69:01	++	A69(28)
A*74:01	++	A74(19)
A*74:03	++	A74(19)
A*74:05	+	-
A*74:06	+	A74(19)
A*74:09	+	-
A*74:10	+	-
A*74:11	+	-
A*74:13	+	-
A*80:01	++	A80

* ++ : *Common alleles (frequency min. 0,001);* + : *well-documented alleles*

The definition of the Common and well documented alleles is part of the kit documentation and based on Common and well documented alleles 2.0 from http://igdawg.org/cwd.html.

Angaben zu den Haplotyp-Häufigkeiten finden sich auf den folgenden Homepages: http://www.allelefrequencies.net/hla60033a.asp und http://bioinformatics.nmdp.org/em-haplotype/. / *Haplotype frequencies are stated on the following websites: http://www.allelefrequencies.net/hla60033a.asp and http://bioinformatics.nmdp.org/em-haplotype/.*

Serological equivalents are based on the EMBL EBI HLA Dictionary http://www.ebi.ac.uk/ipd/imgt/hla/dictionary.html with the WHO assigned settings.

Inno-Train Diagnostik GmbH · Niederhöchstädter Straße 62 · D-61476 Kronberg/Taunus · Germany
Tel.: 0049-(0)6173 / 6079-30 · Fax: 0049-(0)6173 / 6079-50 · E-mail: info@inno-train.de · www.inno-train.de

List of common and well-documented HLA-B alleles (CWD)

Allele Specificity	CWD*	Ser. Type
B*07:02:01	++	B7
B*07:03	+	B703
B*07:04	++	B7
B*07:05:01	++	B7
B*07:06	++	B7
B*07:07	+	B7
B*07:08	+	B7
B*07:09	+	B7
B*07:10	++	B7
B*07:12	+	B7
B*07:13	+	-
B*07:14	+	B7
B*07:15	+	B7
B*07:16	+	B7
B*07:20	+	-
B*07:23	+	-
B*07:26	+	B7
B*07:36	+	-
B*07:37	+	-
B*07:42	+	-
B*07:43	+	-
B*07:46	+	B7
B*07:51	+	-
B*08:01:01	++	B8
B*08:02	+	B8
B*08:03	+	B8
B*08:04	+	-
B*08:07	+	B8
B*08:09	+	B8
B*08:13	+	-
B*08:18	+	B8
B*08:20	+	-
B*08:23	+	-
B*08:35	+	-
B*13:01:01	++	B13
B*13:02:01	++	B13
B*13:03	+	-
B*13:04	+	-
B*13:08Q	+	-
B*13:10	+	-
B*13:12	+	-
B*14:01:01	++	B64(14)
B*14:02:01	++	B65(14)
B*14:03	++	B14
B*14:05	+	-
B*14:09	+	-
B*15:01:01:01	++	B62(15)
B*15:01:01:02N	+	Null
B*15:02:01	++	B75(15)
B*15:03:01	++	B72(70)
B*15:04	+	B62(15)
B*15:05:01	+	B62(15)
B*15:06	+	B62(15)
B*15:07:01	++	B62(15)
B*15:08	++	B75(15)
B*15:09	++	B70
B*15:10:01	++	B71(70)
B*15:11:01	++	B75(15)
B*15:12	++	**B76(15)**
B*15:13:01	++	**B77 (15)**
B*15:14	+	B76(15)
B*15:15	++	B62(15)
B*15:16:01	++	**B63(15)**
B*15:17:01:01	++	**B63(15)**
B*15:18:01	++	B71(70)
B*15:20	+	B62(15)
B*15:21	++	B75(15)
B*15:23	+	-
B*15:24	++	**B62(15)**
B*15:25:01	++	B62(15)
B*15:27:01	+	B62(15)
B*15:28	+	B62(15)
B*15:29	+	B15
B*15:30	++	B62(15)
B*15:31	+	B75(15)
B*15:32	++	B62(15)
B*15:33	+	B15
B*15:34	++	B62(15)
B*15:35	++	B62(15)
B*15:37	+	B70
B*15:38:01	+	-
B*15:39:01	++	B62(15)
B*15:40	+	-
B*15:42	+	-
B*15:44	+	-
B*15:45	+	B62(15)
B*15:46	+	B72(70)
B*15:47:01	+	-
B*15:48	+	B62(15)
B*15:52	+	B15
B*15:53	+	-
B*15:54	+	-
B*15:55	+	B15

Allele Specificity	CWD*	Ser. Type
B*15:56	+	-
B*15:58	+	B62(15)
B*15:63	+	-
B*15:65	+	-
B*15:71	+	B62(15)
B*15:73	+	B62(15)
B*15:75	+	-
B*15:82	+	B62(15)
B*15:113	+	-
B*15:125	+	-
B*15:132	+	-
B*15:141	+	-
B*18:01:01:01	++	B18
B*18:01:02	+	B18
B*18:02	++	B18
B*18:03	++	B18
B*18:04	+	B18
B*18:05	+	B18
B*18:06	+	B18
B*18:08	+	-
B*18:11	+	-
B*18:12	+	-
B*18:13	+	-
B*18:14	+	-
B*18:18	+	-
B*18:20	+	-
B*18:25	+	-
B*18:33	+	-
B*27:01	+	B27
B*27:02:01	++	B27
B*27:03	++	B27
B*27:04:01	++	B27
B*27:05:02	++	B27
B*27:05:03	+	B27
B*27:06	++	B27
B*27:07:01	++	B27
B*27:08	++	B2708
B*27:09	+	B27
B*27:10	+	B27
B*27:12	+	B27
B*27:14	+	B27
B*27:15	+	-
B*27:17	+	B27
B*27:19	+	B27
B*27:20	+	B27
B*27:24	+	-
B*27:27	+	-
B*35:01:01:01	++	B35
B*35:02:01	++	B35
B*35:03:01	++	B35
B*35:04:01	++	B35
B*35:05:01	++	B35
B*35:06	+	B35
B*35:08:01	++	B35
B*35:09:01	+	B35
B*35:09:02	+	B35
B*35:10	+	B35
B*35:11:01	+	B35
B*35:12:01	++	B35
B*35:13	+	B35
B*35:14:01	++	B35
B*35:15	+	B35
B*35:16	++	B35
B*35:17	++	B35
B*35:18	+	B35
B*35:19	++	B35
B*35:20:01	++	B35
B*35:21	++	-
B*35:22	+	-
B*35:23	++	B35
B*35:24:01	+	-
B*35:25	+	B35
B*35:26	+	-
B*35:27	+	B35
B*35:28	+	-
B*35:29:01	+	B35
B*35:30	+	B35
B*35:31	+	-
B*35:32:01	+	B35
B*35:34	+	-
B*35:37	+	-
B*35:41	++	B35
B*35:43:01	++	B35
B*35:47	+	-
B*35:48	+	-
B*35:49	+	-
B*35:55	+	B35
B*35:77	+	-
B*37:01:01	++	B37
B*37:02	+	-
B*37:03N	+	-
B*37:07	+	-

Allele Specificity	CWD*	Ser. Type
B*38:01:01	++	B38(16)
B*38:02:01	++	B38(16)
B*38:02:02	+	B38(16)
B*38:06	+	-
B*38:09	+	B16
B*39:01:01:01	++	B3901
B*39:02:01	+	B3902
B*39:02:02	++	B3902
B*39:03	++	B39(16)
B*39:04	+	B39(16)
B*39:05:01	++	B39(16)
B*39:06:01	+	B39(16)
B*39:06:02	++	B39(16)
B*39:07	+	-
B*39:08	++	B39(16)
B*39:09	+	B39(16)
B*39:10:01	++	B39(16)
B*39:11	++	B39(16)
B*39:12	++	B39(16)
B*39:13:01	+	B39(16)
B*39:14	+	-
B*39:15	++	B39(16)
B*39:20	+	-
B*39:22	+	-
B*39:24:01	++	B39(16)
B*39:26	+	-
B*39:31	+	-
B*39:37	+	-
B*39:39:01	+	-
B*40:01:02	++	B60(40)
B*40:02:01	++	B61(40)
B*40:03	+	B61(40)
B*40:04	+	B61(40)
B*40:05	++	B4005
B*40:06:01:01	++	B61(40)
B*40:07	+	B60(40)
B*40:08	++	-
B*40:09	+	B61(40)
B*40:10:01	+	B60(40)
B*40:11:01	+	B61(40)
B*40:12	+	-
B*40:13	+	-
B*40:16	++	B61(40)
B*40:19	+	-
B*40:20	+	B61(40)
B*40:22N	+	Null
B*40:23	+	-
B*40:25	+	-
B*40:26	+	B21
B*40:27:01	++	B61(40)
B*40:30	+	-
B*40:32	+	B44(12)
B*40:40	+	-
B*40:43	+	-
B*40:50	+	B61(40)
B*40:64	+	B40
B*40:90	+	-
B*40:94	+	-
B*40:114	+	-
B*40:129	+	-
B*41:01	++	B41
B*41:02:01	++	B41
B*41:03:01	++	B41
B*41:07	+	-
B*42:01:01	++	B42
B*42:02	++	B42
B*42:05:01	+	-
B*44:02:01:01	++	B44(12)
B*44:03:01	++	B44(12)
B*44:03:02	++	B44(12)
B*44:04	++	B44(12)
B*44:05:01	++	B44(12)
B*44:06	+	B44(12)
B*44:07	+	B44(12)
B*44:08	+	B44(12)
B*44:09	+	B45(12)
B*44:10	++	B44(12)
B*44:12	+	B44(12)
B*44:15	+	B12
B*44:18	+	-
B*44:20	+	-
B*44:21	+	-
B*44:22	+	-
B*44:23N	+	Null
B*44:24	+	-
B*44:25	+	-
B*44:27:01	++	B44(12)
B*44:29	+	B44(12)
B*44:50	+	-
B*45:01	++	B45(12)
B*46:01:01	++	B46
B*46:02	+	B46

Allele Specificity	CWD*	Ser. Type
B*47:01:01:01	++	B47
B*47:02	+	B47
B*47:03	+	B47
B*48:01:01	++	B48
B*48:02:01	+	B48
B*48:03:01	++	B48
B*48:04	+	B48
B*48:05	+	B48
B*48:06	+	-
B*48:07	+	B48
B*48:12	+	-
B*48:13	+	-
B*49:01:01	++	B49(12)
B*50:01:01	++	B50(21)
B*50:02	++	B45(12)
B*50:04	+	B50(21)
B*51:01:01	++	B51(5)
B*51:01:02	+	B51(5)
B*51:02:01	++	B5102
B*51:02:02	+	B5102
B*51:03	+	B5103
B*51:04	+	B51(5)
B*51:05	++	B51(5)
B*51:06:01	++	B51(5)
B*51:07:01	++	B51(5)
B*51:08	++	B51(5)
B*51:09:01	++	B51(5)
B*51:10	+	-
B*51:11N	+	Null
B*51:13:01	+	-
B*51:13:02	+	-
B*51:14	+	B51(5)
B*51:17	+	B51(5)
B*51:19	+	-
B*51:21	+	-
B*51:22	+	-
B*51:23	+	-
B*51:29	+	B51(5)
B*51:34	+	-
B*51:40	+	-
B*51:43	+	B51(5)
B*51:49	+	-
B*51:50	+	-
B*51:55	+	-
B*52:01:01:01	++	B52(5)
B*52:01:02	++	B52(5)
B*52:02	+	-
B*52:03	+	-
B*52:04	+	B52(5)
B*52:21	+	-
B*53:01:01	++	B53
B*53:02	+	-
B*53:04	+	-
B*53:05	+	-
B*53:06	+	-
B*54:01:01	++	B54(22)
B*55:01:01	++	B55(22)
B*55:01:03	+	B55(22)
B*55:02:01	++	B55(22)
B*55:04	+	B55(22)
B*55:07	+	B54(22)
B*55:12	+	B22
B*55:23	+	-
B*56:01:01	++	B56(22)
B*56:02	+	B56(22)
B*56:03	+	B22
B*56:04	+	B56(22)
B*56:10	+	B55(22)
B*56:15	+	B56(22
B*56:16	+	-
B*57:01:01	++	B57(17)
B*57:02:01	++	B57(17)
B*57:03:01	++	B57(17)
B*57:04	+	B57(17)
B*57:05	+	-
B*58:01:01	++	B58(17)
B*58:02	++	-
B*58:04	+	-
B*58:06	+	-
B*59:01:01:01	++	B59
B*67:01:01	++	B67
B*67:02	+	-
B*73:01	++	B73
B*78:01:01	++	B78
B*78:02:02	+	B78
B*78:04	+	-
B*81:01	++	B81
B*81:02	+	B81
B*82:01	++	B82
B*82:02	+	-

* ++: Common alleles (frequency min. 0,001); +: well-documented alleles

The definition of the Common and well documented alleles is part of the kit documentation and based on Common and well documented alleles 2.0 from http://igdawg.org/cwd.html.

Angaben zu den Haplotyp-Häufigkeiten finden sich auf den folgenden Homepages: http://www.allelefrequencies.net/hla6003a.asp und http://bioinformatics.nmdp.org/em-haplotype/. / *Haplotype frequencies are stated on the following websites: http://www.allelefrequencies.net/hla6003a.asp and http://bioinformatics.nmdp.org/em-haplotype/.*

Serological equivalents are based on the EMBL EBI HLA Dictionary http://www.ebi.ac.uk/ipd/imgt/hla/dictionary.html with the WHO assigned settings.

Inno-Train Diagnostik GmbH · Niederhöchstädter Straße 62 · D-61476 Kronberg/Taunus · Germany
Tel.: 0049-(0)6173 / 6079-30 · Fax: 0049-(0)6173 / 6079-50 · E-mail: info@inno-train.de · www.inno-train.de

HLA-Ready Gene ABDR CE 0123 IVD

List of common and well-documented HLA-DRB1, DRB3, DRB4 & DRB5 alleles (CWD)

Allele Specificity	CWD*	Ser. Type
DRB1*01:01:01	++	DR1
DRB1*01:01:02	++	DR1
DRB1*01:02:01	++	DR1
DRB1*01:03	++	DR103
DRB1*01:04	+	DR1
DRB1*01:07	+	DR1
DRB1*01:11	+	-
DRB1*03:01:01:01	++	DR17(3)
DRB1*03:01:02	+	DR17(3)
DRB1*03:02:01	++	DR18(3)
DRB1*03:02:02	+	DR18(3)
DRB1*03:03	+	DR18(3)
DRB1*03:04:01	+	DR17(3)
DRB1*03:05:01	+	DR3
DRB1*03:06	+	DR3
DRB1*03:07	+	DR3
DRB1*03:15	+	DR3
DRB1*03:16	+	-
DRB1*03:22	+	-
DRB1*03:23	+	DR3
DRB1*04:01:01	++	DR4
DRB1*04:02:01	++	DR4
DRB1*04:03:01	++	DR4
DRB1*04:03:02	+	DR4
DRB1*04:04:01	++	DR4
DRB1*04:05:01	++	DR4
DRB1*04:06:01	++	DR4
DRB1*04:07:01	++	DR4
DRB1*04:08:01	++	DR4
DRB1*04:09	++	DR4
DRB1*04:10:01	++	DR4
DRB1*04:11:01	++	DR4
DRB1*04:12	+	-
DRB1*04:13	+	DR4
DRB1*04:14	+	DR4
DRB1*04:15	+	DR4
DRB1*04:16	+	DR4
DRB1*04:17:02	++	DR4
DRB1*04:18	+	-
DRB1*04:19	+	DR4
DRB1*04:22	+	DR4
DRB1*04:23	+	DR5
DRB1*04:25	+	DR6
DRB1*04:26	+	DR4
DRB1*04:33	+	-
DRB1*04:35	+	-
DRB1*04:38	++	-
DRB1*04:40	+	-
DRB1*04:41	+	-
DRB1*04:50	+	-
DRB1*04:51	+	-
DRB1*04:54	+	-
DRB1*07:01:01:01	++	DR7
DRB1*07:03	+	DR7
DRB1*07:05	+	-
DRB1*07:07	+	-
DRB1*07:11	+	-
DRB1*07:13	+	-
DRB1*08:01:01	++	DR8
DRB1*08:02:01	++	DR8
DRB1*08:02:03	+	DR8
DRB1*08:03:02	++	DR8
DRB1*08:04:01	++	DR8
DRB1*08:04:02	+	DR8
DRB1*08:04:04	++	DR8
DRB1*08:05	+	DR8
DRB1*08:06	++	DR8
DRB1*08:07	++	DR8
DRB1*08:08	+	-
DRB1*08:09	+	DR8
DRB1*08:10	++	DR8
DRB1*08:11	++	DR8
DRB1*08:12	+	DR8
DRB1*08:13	+	DR8
DRB1*08:14	+	DR8
DRB1*08:17	+	DR8

Allele Specificity	CWD*	Ser. Type
DRB1*08:18	+	DR8
DRB1*08:19	+	-
DRB1*08:20	+	-
DRB1*08:26	+	-
DRB1*08:28	+	-
DRB1*09:01:02	++	DR9
DRB1*09:04	+	-
DRB1*09:06	+	-
DRB1*10:01:01	++	DR10
DRB1*11:01:01	++	DR11(5)
DRB1*11:01:02	++	DR11(5)
DRB1*11:02:01	++	DR11(5)
DRB1*11:03	++	DR11(5)
DRB1*11:04:01	++	DR11(5)
DRB1*11:04:02	++	DR11(5)
DRB1*11:05	+	DR11(5)
DRB1*11:06:01	++	DR11(5)
DRB1*11:07	+	DR11(5)
DRB1*11:08:01	+	DR11(5)
DRB1*11:09	++	DR11(5)
DRB1*11:10:01	++	-
DRB1*11:11:01	++	DR11(5)
DRB1*11:12:01	++	-
DRB1*11:13:01	+	DR11(5)
DRB1*11:14:01	+	DR11(5)
DRB1*11:15	++	DR11(5)
DRB1*11:16	+	
DRB1*11:17	++	-
DRB1*11:18	+	-
DRB1*11:19:01	++	DR11(6)
DRB1*11:20	+	DR11(5)
DRB1*11:24	+	-
DRB1*11:25	+	DR11(5)
DRB1*11:29	+	DR11(5)
DRB1*11:32	+	DR11(5)
DRB1*11:28:01	+	-
DRB1*11:34	+	-
DRB1*11:36	+	-
DRB1*11:37:01	+	DR11(5)
DRB1*11:39	+	-
DRB1*11:42	+	-
DRB1*11:43	+	-
DRB1*11:45	+	-
DRB1*11:47	+	-
DRB1*11:56	+	-
DRB1*11:66	+	-
DRB1*11:69	+	-
DRB1*12:01:01	++	DR12(5)
DRB1*12:02:01	++	DR12(5)
DRB1*12:02:02	++	DR12(5)
DRB1*12:04	+	-
DRB1*12:05	+	DR12(5)
DRB1*12:10	+	-
DRB1*13:01:01	++	DR13(6)
DRB1*13:01:03	+	DR13(6)
DRB1*13:02:01	++	DR13(6)
DRB1*13:03:01	++	DR13(6)
DRB1*13:03:02	+	DR13(6)
DRB1*13:04	++	DR13(6)
DRB1*13:05:01	++	DR13(6)
DRB1*13:06	+	DR13(6)
DRB1*13:07:01	+	DR13(6)
DRB1*13:07:02	++	DR13(6)
DRB1*13:08	+	DR13(6)
DRB1*13:09	+	-
DRB1*13:10	++	DR13(6)
DRB1*13:11:01	+	DR13(6)
DRB1*13:12:01	++	DR13(6)
DRB1*13:13	+	DR13(6)
DRB1*13:14:01	++	DR13(6)
DRB1*13:15	+	-
DRB1*13:16	+	DR13(6)
DRB1*13:17	+	DR13(6)
DRB1*13:18	+	DR13(6)
DRB1*13:19	+	DR13(6)
DRB1*13:20	++	DR13(6)

Allele Specificity	CWD*	Ser. Type
DRB1*13:22	+	DR13(6)
DRB1*13:25	+	-
DRB1*13:26:01	+	-
DRB1*13:27	+	DR13(6)
DRB1*13:29	+	DR13(6)
DRB1*13:31	+	-
DRB1*13:36	+	DR13(6)
DRB1*13:37	+	-
DRB1*13:38	+	-
DRB1*13:39	+	-
DRB1*13:40	+	-
DRB1*13:41	+	-
DRB1*13:42	+	DR13(6)
DRB1*13:49	+	-
DRB1*13:56	+	-
DRB1*13:59	+	-
DRB1*13:60	+	-
DRB1*13:76	+	-
DRB1*13:81	+	-
DRB1*13:82	+	-
DRB1*14:01:01	++	DR14(6)
DRB1*14:02	++	DR14(6)
DRB1*14:03:01	++	DR1403
DRB1*14:04	++	DR1404
DRB1*14:05:01	++	DR14(6)
DRB1*14:06:01	++	DR14(6)
DRB1*14:07:01	++	DR14(6)
DRB1*14:08	+	DR14(6)
DRB1*14:09	+	-
DRB1*14:10	+	DR14(6)
DRB1*14:11	+	DR14(6)
DRB1*14:12:01	+	DR14(6)
DRB1*14:13	+	DR14(6)
DRB1*14:14	+	DR14(6)
DRB1*14:15	+	DR8
DRB1*14:16	+	DR6
DRB1*14:17	++	DR6
DRB1*14:18	++	DR6
DRB1*14:19	+	DR14(6)
DRB1*14:20	+	DR14(6)
DRB1*14:21	+	DR14(6)
DRB1*14:22	+	DR14(6)
DRB1*14:24	++	-
DRB1*14:25	+	-
DRB1*14:28	+	-
DRB1*14:29	+	DR14(6)
DRB1*14:33	+	-
DRB1*14:48	+	-
DRB1*14:54:01	++	DR14(6)
DRB1*14:61	+	-
DRB1*14:70	+	-
DRB1*15:01:01:01	++	DR15(2)
DRB1*15:02:01	++	DR15(2)
DRB1*15:02:02	++	DR15(2)
DRB1*15:03:01:01	++	DR15(2)
DRB1*15:04	++	DR15(2)
DRB1*15:06:01	++	DR15(2)
DRB1*15:07:01	+	DR15(2)
DRB1*15:10	+	-
DRB1*15:11	+	-
DRB1*15:14	+	-
DRB1*15:18	+	-
DRB1*15:20	+	-
DRB1*15:22	+	-
DRB1*15:23	+	-
DRB1*15:24	+	-
DRB1*15:38	+	-
DRB1*16:01:01	++	DR16(2)
DRB1*16:02:01	++	DR16(2)
DRB1*16:04	++	DR16(2)
DRB1*16:05:01	+	DR16(2)
DRB1*16:07	++	-
DRB1*16:10	+	-
DRB1*16:12	+	-

Allele Specificity	CWD*	Ser. Type
DRB3*01:01:02:01	++	DR52
DRB3*01:02	+	DR52
DRB3*01:03	+	-
DRB3*02:01	++	DR52
DRB3*02:02:01:01	++	DR52
DRB3*02:02:02	+	DR52
DRB3*02:03	+	DR52
DRB3*02:06	+	-
DRB3*02:10	++	DR52
DRB3*02:11	+	DR52
DRB3*02:17	+	-
DRB3*03:01:01	++	DR52
DRB4*01:01:01:01	++	DR53
DRB4*01:02	++	DR53
DRB4*01:03:01:01	++	DR53
DRB4*01:03:01:02N	++	Null
DRB4*01:03:02	+	DR53
DRB4*01:03:03	++	DR53
DRB4*02:01N	++	Null
DRB4*03:01N	+	Null
DRB5*01:01:01	++	DR51
DRB5*01:01:02	+	DR51
DRB5*01:02	++	DR51
DRB5*01:03	++	-
DRB5*01:08N	++	Null
DRB5*01:10N	+	Null
DRB5*02:02	++	DR51
DRB5*02:03	+	-

* ++ : Common alleles (frequency min. 0,001); + : well-documented alleles

The definition of the Common and well documented alleles is part of the kit documentation and based on Common and well documented alleles 2.0 from http://igdawg.org/cwd.html.

Angaben zu den Haplotyp-Häufigkeiten finden sich auf den folgenden Homepages: http://www.allelefrequencies.net/hla6003a.asp und http://bioinformatics.nmdp.org/em-haplotype/. / *Haplotype frequencies are stated on the following websites: http://www.allelefrequencies.net/hla6003a.asp and http://bioinformatics.nmdp.org/em-haplotype/.*

Serological equivalents are based on the EMBL EBI HLA Dictionary http://www.ebi.ac.uk/ipd/imgt/hla/dictionary.html with the WHO assigned settings.

Inno-Train Diagnostik GmbH · Niederhöchstädter Straße 62 · D-61476 Kronberg/Taunus · Germany
Tel.: 0049-(0)6173 / 6079-30 · Fax: 0049-(0)6173 / 6079-50 · E-mail: info@inno-train.de · www.inno-train.de

Human Immunology 81 (2020) 395–396

Contents lists available at ScienceDirect

Human Immunology

journal homepage: www.elsevier.com/locate/humimm

Short population report

HLA-A -B and -DRB1 distribution in Kami: A caste population of Gorkha community from the sub-Himalayan region of West Bengal, India

Bisu Singh*, Dependra Chamlagai, Jiwan Gurung

Department of Zoology, School of Life Sciences, Sikkim University, Gangtok, Sikkim, India

ARTICLE INFO

Keywords:
Kami
HLA
Darjeeling
West Bengal
Gorkha

ABSTRACT

The present investigation is the first of its kind to understand the HLA profile of Kami population from the Indian Gorkha community of sub-Himalayan West Bengal, India. A total of 158 individuals from Kami population were genotyped at first field resolution by HLA ABDRB1 PCR SSP typing kit. The genotype assignment to the individuals was performed by Ready Gene V.1.0.0.0′ software. The data were analysed by PopWin32-0.7.0 software. All the loci typed were in Hardy-Weinberg equilibrium. The genotype data is accessible at Allele Frequencies Net Database with the name India, sub-Himalayan West Bengal, Kami number 3720.

1. Introduction

The sub-Himalayan region of West Bengal, India extends between 27°2′ N latitude and 88°15′E longitude. It is bordered by Nepal, Sikkim and Bhutan and covers a surface area of 3149 km^2 with a population of around 18.47 lakh mainly composed Gorkha (also called Gurkha), Lepcha and Bhutia community. The current demographic pattern of Gorkhas in the sub-Himalayan West Bengal is largely due to migration across the Indo-Nepal border since 1700 CE. The principal factors for migration were extremely severe policies in Nepal, developing tea industry, the establishment of Darjeeling Himalayan Railways, development of Darjeeling into a hill resort by Britishers and recruitment of Gorkha soldiers in British Army from outside the borders of Nepal [1–4]. Historically the word Gorkha is associated with khasas from northern India [5]. Moreover, after the establishment of the Gorkha dynasty by Drabya Shah the word Gorkha was used for the inhabitants of the principality where the Gorkhas ruled [6]. However, nowadays the term 'Gorkha' is used to define a community or is used to differentiate Indian citizens of Nepali ethnicity from citizens of Nepal [6]. The Gorkhas as a community can be divided into three ethnic groups; (i)'Kiratis' which include Rai, Magar, Limbu, Lepcha, Tamang, (ii)'Newaris' or 'Newars', and (iii) 'Tagadharis' who are the Nepali counterpart of Indian Hindus including Bahuns, Chettri, Kami, Damai and Sarki [5]. Each of these community is unique in their language, script, culture and customs.

Kami are Indo-Aryan language speaking Hindu socio-ethnic caste group from Gorkha community mostly confined to the sub-Himalayan region in the state of West Bengal, and sporadically to the other states of India. They are essentially artisans involved in metalwork and makers of famous 'Khukuri' knives used by the Gorkha army. Besides, they are also involved in making traditional Nepali drum called 'Madal' and exponent of 'Maruni Nritya', a traditional folk dance. They are mostly monogamous however; polygamy was also evident in the past. Kami are classified into 54 exogamous clans and the marriage among the different clans is a common phenomenon [7]. However, the marriage with other populations within and outside the Gorkha community is rare. According to 2011 census Kami population is about 52,178 in West Bengal. Very little is known about the origin of Kami population. However, it has been mentioned that predecessors of Kami have emigrated from different places of India such as Punjab, Rajputana, Kashmir and neighbouring places between 10th and 15th Centuries AD to Nepal [7].

The 158 individuals belonging to Kami population were recruited from Darjeeling, Kalimpong and Jalpaiguri district of sub-Himalayan West Bengal, India. After explaining the study procedure, written consent was taken from the individuals for voluntarily donating their blood samples and participation in the study. The three-generation pedigree charts were prepared for each individual to assure their unrelatedness, and those having the history of inter-caste marriage within the studied pedigree were excluded. The study was approved by the Institutional Ethical Committee, Sikkim University. All the laboratory experiment was conducted at the Molecular Biology Laboratory, Department of Zoology, Sikkim University, Gangtok, India. The extraction of genomic DNA was performed by a kit (Qiagen) and the first field HLA Class I and Class II genotyping were performed by HLA ABDRB1 PCR SSP typing kit (Inno-Train Diagnostik, Germany) as per allele list based on release

* Correspondence author at: Department of Zoology 5th Mile, Sikkim University Gangtok, 737101 East Sikkim, India.
E-mail address: bisusingh22@yahoo.co.in (B. Singh).

https://doi.org/10.1016/j.humimm.2020.06.007
Received 10 June 2020; Received in revised form 14 June 2020; Accepted 14 June 2020
Available online 29 June 2020

Human Heredity

Brief Report

Hum Hered
DOI: 10.1159/000514220

Received: July 10, 2020
Accepted: December 23, 2020
Published online: February 16, 2021

HLA Profile of Kami Population Refutes the Earlier Proposition of Exclusive Closer Genetic Affinity of All the Gorkhas to Mongoloids

Bisu Singh Dependra Chamlagai Jiwan Gurung

Department of Zoology, School of Life Sciences, Sikkim University, Gangtok, India

Keywords
Human leucocyte antigen · Kami population · Gorkha population · Khukuri knife · Indo-Aryan population

Abstract
Objective: Based on the HLA profile of Indian Gorkhas, Debnath and Chaudhuri (2006) proposed that Gorkhas are genetically closer to Mongoloids, and they may have originated from Mongolians or Tibetan stocks. However, the major limitation of the earlier study was that Gorkhas comprise 2 broad groups, i.e. Tibeto-Burmans and Indo-Aryans. Besides, Gorkhas have an assemblage of many sociocultural and linguistically distinct populations such as Rai, Magar, Limbu, Tamang, Newar, Bahun, Kami, and so on. Thus, the generalization of the findings on Gorkhas by considering them as a single homogenous population may not be free from biases. Therefore, the present study aims to understand the genetic affinity of a constituent population from the Gorkha community, i.e. Kami, based on HLA polymorphism. ***Methods:*** First field HLA typing was performed among 158 Kami individuals by PCR-SSP methods. ***Results:*** The most frequent genes observed were HLA-A*11, HLA-B*15, HLA-DRB1*15. The frequency of HLA-DRB1*15 reported here is the highest recorded among the North Indian population to date, which is a noteworthy finding of the study. The hierarchical cluster analysis and principal component analysis showed that the Kami population lies within the cluster of the Indian subcontinental population. ***Conclusion:*** The study refutes the earlier proposition of exclusive belongingness of all the Gorkhas to Mongoloids.

Introduction

Kamis are an Indo-Aryan language-speaking socio-ethnic caste group from the Gorkha community. Presently the Indian Kami population is mostly confined to the sub-Himalayan region of West Bengal, India, and sporadically in the other parts of the country. Kamis are essentially Hindu and worship different Hindu deities of the pantheon. However, few of them have converted to Christianity or other religions. By tradition, Kamis are artisans involved in metalwork and hold the legacy of making the famous "Khukuri" knives used by the Gorkhas. Monogamy is prevalent among the Kamis, even though polygamy was also common in the past. There are 54 exogamous clans among the Kamis, and marriage among the different clans is a usual phenomenon [1]. Sporadic intercaste marriage of Kamis with other populations from the Gorkha community is evident. However,

karger@karger.com
www.karger.com/hhe

Karger

Bisu Singh
Department of Zoology
Sikkim University
5th Mile, Gangtok 737101, East Sikkim (India)
bisusingh22@yahoo.co.in

Full text open access online (Since 2001)

PRINT: ISSN 0972-3757 ONLINE: ISSN 2456-6330
Int J Hum Genet, 24(2): 199-204 (2024)
DOI: 10.31901/24566322.2024/24.02.884

HLA Association with Type 2 Diabetes Mellitus and Hypertension among the Kami Population from Sub-Himalayan West Bengal

Chamlagai Dependra[1,3], Jiwan Gurung[2,3] and Bisu Singh[3,*]

[1]Department of Zoology, Nar Bahadur Bhandari Government College, Government of Sikkim, Sikkim, India
E-mail: dependrachamlagai@gmail.com
[2]Department of Zoology, Sikkim Alpine University, Namchi, Sikkim, India
E-mail: geminejeevan@gmail.com
[3]Department of Zoology, School of Life Sciences, Sikkim University, Gangtok, Sikkim, India
E-mail: bisusingh22@yahoo.co.in

KEYWORDS Association. Genetics. Immunology. Polymorphism. Susceptibility

ABSTRACT The Kami population is an Indo-Aryan caste group from the Gorkha community of the Sub-Himalayan region. This study aims to investigate the association of Human Leukocyte Antigen (HLA) genes with Type 2 Diabetes Mellitus (T2DM) and hypertension (HT) among the Kami population. One hundred fifty unrelated Kami individuals were recruited from the sub-Himalayan region of West Bengal, India. Twenty of them had a history of T2DM, and 15 had HT. The enrolled individuals provided blood samples, which were then utilized to perform HLA typing using PCR-SSP A, B, and DR typing kits. The findings showed that T2DM cohort had a higher frequency of HLA-B*13 and HLA-B*15 and a lower frequency of HLA-A*33. Conversely, a significantly high frequency of HLA-A*24 and HLA-B*13 was observed among the HT cohort. The study demonstrates the presence of susceptibility genetic factors at HLA loci for T2DM and HT among the Kami population.

INTRODUCTION

The incidence and burden of Type 2 Diabetes Mellitus (T2DM) and hypertension (HT) have increased considerably in the last three decades among the world population (Danaei et al. 2011). Both these diseases have multifactorial aetiology with prevalence varying widely across geographical region and ethnicity. The susceptibility of T2DM varies across various populations with studies reporting higher risk for Asian Indians, Native Americans, and Pacific Islanders (Pradeepa and Mohan 2021). It is suggested that the aetiology of type II diabetes involves interactions between multiple genes, and gene-environment interactions (Chen et al. 2012). Since interactions between genes and the environment are involved, delineating the specific genetic and environmental factors in the etiopathology of T2DM has been a daunting task (Barroso 2005). Similarly, interactions between genetic, environmental, and demographic factors are suggested aetiological factors for hypertension (Kato 2012). In addition, ethnic and genetic factors are implicated in the disparity in the rate of prevalence of hypertension among the populations (Tomson and Lip 2005). Diabetes and hypertension are common in India across all ages and all regions of the country (Geldsetzer et al. 2018). Given the genetic diversity found in people with varying ethnic backgrounds, genetic factors that contribute to the aetiology of these two diseases may play a significant role. Many researchers have attempted to correlate type II diabetes and hypertension to different genetic loci across various populations of different ethnic origins (Sladek et al. 2007; Chauhan et al. 2010; Yamamoto et al. 2006; Shimodaira et al. 2010; Kamide et al. 2011; Zhang et al. 2013; Zhang et al. 2016).

Of all the gene loci in the human genome, the human leukocyte antigen (HLA) is among the most researched and has been connected to more disorders than any other genes (Trowsdale 2011). Due to the substantial amount of polymorphism and the notable variations in HLA gene frequencies among various ethnic groups, it is imperative that the ethnic composition of the population be taken into account when performing studies related to

*Address for correspondence:
Dr. Bisu Singh
Department of Zoology
Sikkim University
Gangtok 737 101, East Sikkim
Phone: +919733155848
E-mail: bisusingh22@yahoo.co.in

Original Article

Access this article online

Quick Response Code:

Website: www.ajts.org

DOI: 10.4103/ajts.ajts_158_23

Distribution of ABO and Rhesus (D) blood group antigens among the Indian Kami population from the Sub-Himalayan region of West Bengal and Sikkim, India

Dependra Chamlagai[1,2], Jiwan Gurung[3], Bisu Singh[1]

Abstract:

BACKGROUND: Limited information is available regarding the distribution of ABO and Rhesus (Rh) blood groups among the populations from the sub-Himalayan regions. The present study was undertaken to understand the phenotypic, genotypic, and allelic frequencies of ABO and Rh antigens among the Kami population.

METHODS: A total of 1420 Kami individuals were included in the study from the sub-Himalayan region of West Bengal and Sikkim. The ABO and RhD blood groups were determined among 158 individuals by standard slide agglutination method, and from 1262 individuals blood group, data were collected who had undergone routine blood group tests in diagnostic laboratories.

RESULTS: The frequency of blood group O (31.06%) was found to be the highest followed by B (29.15%), A (27.54%), and AB (12.25%). A total of 97.89% individuals were Rh (D) positive and 2.11% were Rh (D) negative. The allele frequencies of ABO were found to be 0.200 for I^A (p), 0.243 for I^B (q), and 0.557 for I^O (r). For Rh (D), the calculated gene frequencies were 0.855 for Rh positive (I^D) and 0.145 for Rh negative (I^d). Among the subcaste of Kami population, a significantly high frequency of blood group A was observed among Baraily and Lamgade, O among Sashankar, Dural, Diyali, and Dutraj, and B among Ramudamu, Koirala, and Silal.

CONCLUSION: The study is the first of its kind to comprehensively describe the ABO blood group distribution among the Kami population which shows the same trend as the general Indian population.

Keywords:

ABO, blood group, Kami, Sikkim, Sub-Himalayan West Bengal

[1]Immunology, Parasitology and Cytogenetics laboratory, Department of Zoology, School of Life Sciences, Sikkim University, [2]Department of Zoology, Nar Bahadur Bhandari Government College, Gangtok, [3]Department of Zoology, Sikkim Alpine University, Namchi, Sikkim, India

Address for correspondence:
Dr. Bisu Singh, Immunology, Parasitology and Cytogenetics laboratory, Department of Zoology, School of Life Sciences, Sikkim University, Gangtok - 737 101, East Sikkim, India.
E-mail: bisusingh22@yahoo.co.in

Submitted: 09-08-2023
Revised: 27-08-2023
Accepted: 10-09-2023
Published: 06-02-2024

Introduction

The Kami population is the socioethnic caste group belonging to the Gorkha community mostly confined to the sub-Himalayan region of West Bengal, India. However, their sporadic distribution can be found in North East and other parts of India. Kami population consists of 54 subcastes with a total population of 52,178 in West Bengal and 35,822 in Sikkim as per the 2011 census. They are primarily known for their work as blacksmith and metalsmith and belong to the schedule caste community as per the Indian constitution. There is not much information regarding their origin but their ancestry can be traced back to Punjab, Rajputana, and Kashmir.[1]

Karl Landsteiner's discovery of ABO blood group antigens in 1901 and the subsequent identification of Rhesus (Rh) blood group antigen in 1937, in collaboration with

How to cite this article: Chamlagai D, Gurung J, Singh B. Distribution of ABO and Rhesus (D) blood group antigens among the Indian Kami population from the sub-Himalayan region of West Bengal and Sikkim, India. Asian J Transfus Sci 0;0:0.

ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES
Knowledge to Innovation

Vol 9, Issue 3, 2016

Online - 2455-3891
Print - 0974-2441

Research Article

STUDY OF *IN VITRO* ANTI-INFLAMMATORY ACTIVITY OF ETHNOMEDICINAL PLANTS OF SIKKIM *VISCUM ARTICULATUM* AND *ACORUS CALAMUS*

DEPENDRA CHAMLAGAI, BISU SINGH*

Department of Zoology, School of Life Sciences, Sikkim University, Samdur, Tadong, Gangtok, Sikkim, India.
Email: bisusingh22@yahoo.co.in

Received: 22 December 2015, Revised and Accepted: 07 January 2016

ABSTRACT

Objective: The present study was carried out to evaluate the anti-inflammatory property of ethnomedicinal plants *Viscum articulatum* and *Acorus calamus.*

Methods: Human red blood cells membrane stabilization method was applied to assess the anti-inflammatory property of both the plants.

Results: It was observed that both plant extracts have the anti-inflammatory potential comparable with that of the drug indomethacin. However, *A. calamus* was found to provide slightly more inhibition of hemolysis (76.8%) than that of standard drug (indomethacin) (72.8%), whereas *V. articulatum* provided slightly lesser inhibition (68%) in comparison to the drug. There was the tendency of increase protection along with the increase in the concentration of the extract, and maximum protection was achieved at the maximum concentration of 5000 µg/ml.

Conclusion: The study proves the anti-inflammatory efficacy of both the plants and they hold a good prospect for drug development against inflammation.

Keywords: Anti-inflammatory, *Viscum articulatum*, *Acorus calamus*, Sikkim.

INTRODUCTION

Sikkim is part of the Himalayan region and with its exquisite ecosystem hosts varieties of medicinal plants. Over 400 plants possessing therapeutic properties have been recorded from this region [1]. The tremendous medicinal plant wealth of Sikkim Himalayas has been part of the traditional system of medicine (Ayurveda, Homeopathy, Naturopathy, and Unani and Siddha system of medicine) [2]. Ancient medicinal systems abound in Sikkim are popularly nurtured by Buddhist groups for their traditional Tibetan pharmacopeia [3]. In traditional medicine, many of the plants are used in combination as well as individually to produce the desired effect.

Inflammation has been associated with many diseases which bring about the hazardous effect on the patients sometimes causing a threat to his/her life. This has attracted the attention of researchers to this field. There are many anti-inflammatory drugs to treat the consequences of inflammation belonging to steroidal or non-steroidal anti-inflammatory drugs (NSAIDS). However, studies suggest that these drugs are not free from adverse effects, as they are responsible for gastrointestinal complications such as mucosal damage and bleeding [4]. Moreover, NSAIDS can also cause acute renal failure [5]. For these reasons, many researchers have shifted their focus on finding the medicinal plants, which have gotten the anti-inflammatory property and can serve as a potential ingredient for future drug development [6].

Many studies had been conducted on plants to investigate its anti-inflammatory potential, and also these plants have been screened for phytochemicals, which are actively associated with its anti-inflammatory property. Over the years, one of the most popular methods to study *in vitro* anti-inflammatory property of plant extract has been membrane stabilization method of red blood cells (RBC). In one such study, anti-inflammatory activity was evaluated on extracts of fresh leaves of *Clerodendrum paniculatum* by *in vitro* (human red blood cells [HRBC] membrane stabilization method) and *in vivo* methods (0.1 ml of 1% w/v carrageenan-induced rat paw edema model). Petroleum ether, chloroform, ethyl acetate, alcohol, and aqueous extracts of *C. paniculatum* were screened for *in vitro* anti-inflammatory activity. Among them, petroleum ether and chloroform extracts showed the best *in vitro* anti-inflammatory activity at the dose level of 200 and 400 mg/kg. The study demonstrated that petroleum ether and chloroform extracts possess statistically significant ($p<0.001$) anti-inflammatory potential which provides a scientific basis for the traditional claims of *C. paniculatum* leaves as an anti-inflammatory drug [7]. In another study evaluation of the anti-inflammatory property of the leaf exacts of *Gendarussa vulgaris* was carried out by Saleem *et al.* [8], the aqueous and alcoholic extracts of the leaves were used in both *in vitro* and *in vivo* methods. *In vitro* anti-inflammatory property was estimated by HRBC membrane stabilization method. The results concluded that the alcoholic extract at a concentration of 300 mg/ml showed potent activity on comparing with the standard drug diclofenac sodium. Another study was carried out to evaluate the *in vitro* anti-inflammatory activity of *Centella asiatica* by HRBC membrane stabilization. The maximum membrane stabilization of *C. asiatica* extracts was found to be 94.97% at a dose of 2000 µg/ml. The study revealed the potency of active constituents from *C. asiatica* in treating inflammations [9]. Further, phytochemical analysis of *C. asiatica* plant extracts revealed the presence of biochemical compounds such as triterpenoids and flavonoids, which have been observed to have remarkable anti-inflammatory activity. Recently, a study was conducted to evaluate the membrane stabilization activity of *Spilanthes paniculata* leaves. The results showed that at the concentration of 1 mg/ml, ethanol extract, and n-hexane and ethyl acetate soluble fractions significantly inhibited hypotonic solution-induced lysis of the HRBC (27.406±3.57, 46.034±3.251, and 30.72±5.679%, respectively), whereas standard drug acetylsalicylic acid (concentration 0.1 mg/ml) showed 77.276±0.321% inhibition. In the case of heat-induced HRBC hemolysis, the plant extracts also showed significant activity (34.21±4.72%, 21.81±3.08%, and 27.62±8.79% inhibition, respectively). The study showed n-hexane extract has better effects than the other extracts. It was concluded that the leaves of *S. paniculata* possess remarkable pharmacological effects and has the potential to act as an anti-inflammatory agent [6]. Further, Somlata (*Sarcostemma acidum*) highly used by the rural and tribal

NORDIC JOURNAL OF PSYCHIATRY, 2018
https://doi.org/10.1080/08039488.2018.1441438

Taylor & Francis
Taylor & Francis Group

BRIEF REPORT

Elevated levels of C-reactive protein and IL-6 among the antipsychotic medicating schizophrenia patients of Siliguri, West Bengal, India

Jiwan Gurung[a], Dependra Chamlagai[a], Nirmal Kumar Bera[b], Tapas Kumar Chaudhuri[c] and Bisu Singh[a]

[a]Department of Zoology, School of Life Sciences, Sikkim University, Sikkim, India; [b]Department of Psychiatry, North Bengal Medical College, and Hospital, Siliguri, India; [c]Department of Zoology, Cellular Immunology Laboratory, University of North Bengal, Siliguri, India

ABSTRACT

Purpose: Chronic, low-grade inflammation is a proposed etiological factor associated with schizophrenia. Thus, various studies have been conducted to understand the role of inflammatory process in schizophrenia by using inflammatory maker C-reactive protein (CRP) with conflicting findings. Inadvertently, studies of CRP among the Indian schizophrenia patients are very few. Therefore, the present study was undertaken to investigate the role of inflammatory process among Indian Bengalee schizophrenia patients of Siliguri, using the marker CRP and its stimulating cytokine interleukin-6 (IL-6). In addition, the study also intended to investigate the immunomodulatory effect of antipsychotic medication on serum levels of CRP and IL-6.

Materials and methods: The serum levels of CRP and IL-6 were measured by Enzyme-Linked Immunosorbent Assay (ELISA) among 67 antipsychotic medicating, 28 psychotropic medication-free schizophrenia patients, and it was compared with 72 age, sex and ethnicity-matched controls.

Results: A significantly higher level of CRP and IL-6 were recorded among the antipsychotic medicating patients. Although CRP was found to be higher among the psychotropic medication-free patients than the controls, it was not found to be significant. However, a significantly higher level of IL-6 was observed in this group.

Conclusions: The results provide the evidence for a possible immunomodulatory effect of antipsychotic drugs on CRP. Future investigations including the study of antipsychotics separately may help to understand the differential effects of individual antipsychotics on CRP level. Additional studies with a larger sample size of psychotropic medication-free patients may help to verify the role of inflammation in schizophrenia patients of this region.

ARTICLE HISTORY
Received 15 August 2017
Revised 24 December 2017
Accepted 11 February 2018

KEYWORDS
C-reactive protein; interleukin-6; schizophrenia; Siliguri

1. Introduction

Schizophrenia is a multifactorial psychiatric disorder characterized by positive, negative and cognitive symptoms [1]. Despite tireless research, etiology of schizophrenia is still concealed [2]. Increasing evidence suggests that dysregulation of the immune system and inflammatory processes, along with the genetic and environmental factors contribute to etiology and pathophysiology of schizophrenia [3–5]. Recently, the study of immunoinflammatory process in schizophrenia has received an increase attention from the researchers. One of the well-known inflammatory markers which have been intensively studied in schizophrenia is C-reactive protein (CRP). It is synthesized by hepatocytes during the inflammatory process [6] under the direct stimulation of interleukin-6 (IL-6) [7]. However, the findings of CRP in schizophrenia till date are not in agreement with each other. Thus, some of the studies reported its increased level [8–13], whereas others observed no associations [14–18]. The reasons for inconsistencies of the findings may be attributed to small sample size [8,9,19,20], the absence of controls [10,21–23], lack of simultaneous studies of CRP along with its direct stimulatory cytokine such as IL-6 [18,24,25], and none consideration of medication status [22,26,27] in the previous investigations.

Various studies have been carried out till date among the schizophrenia patients to understand the modulatory effect of antipsychotic drugs on the CRP levels. Thus, some of the studies reported higher levels of CRP after treatment with clozapine [28], olanzapine [14], quetiapine [22] and atypical antipsychotics [29]. Nonetheless, in most of the previous investigations, either the controls were not included in the study or else have considered only first-degree relatives of the schizophrenia patients. However, the comparison of CRP among the patients and their relatives may not be free from biases as studies have reported 40% heritability of the baseline CRP levels [30]. Moreover, the variation in gene encoding CRP was found to be associated with circulating CRP levels [31,32]. Furthermore, studies suggest that CRP gene polymorphisms are independently associated with increased or decreased levels of CRP [33].

Most of the studies of CRP in schizophrenia have been carried out in the Caucasian [13,34] and some Mongoloid population [24,35]. The literature review suggests that only two studies [12,20] have been conducted so far in India to

CONTACT Bisu Singh bisusingh22@yahoo.co.in Department of Zoology, Sikkim University, Gangtok 737101, East Sikkim, India