PREVALENCE OF SICKLE CELL DISEASE IN THARU COMMUNITIES OF RUPANDEHI DISTRICT, NEPAL

Manoj Kumar Tharu
T.U. Registration No: 5-2-52-53-2009
T.U. Examination Roll No: Zoo 326/072
Batch: 2072

A thesis submitted in partial fulfilment of the requirements for the award of the degree of Master of Science in Zoology with special paper Ecology.

Submitted to

Central Department of Zoology
Institute of Science and Technology
Tribhuvan University
Kirtipur, Kathmandu
Nepal

April, 2021

DECLARATION

I hereby declare that the work presented in this thesis has been done by myself, and has not been submitted elsewhere for the award of any degree. All sources of information have been specifically acknowledged by reference to the author(s) or institution(s).

Date....19 March 2021

.....
Manoj Kumar Tharu

RECOMMENDATION

This is to recommend that the thesis entitled "**PREVALENCE OF SICKLE CELL DISEASE IN THARU COMMUNITIES OF RUPANDEHI DISTRICT, NEPAL**" has been carried out by **Manoj Kumar Tharu** for the partial fulfilment of Master's Degree of Science in Zoology with special paper Ecology and environment. This is his original work and has been carried out under my supervision. To the best of my knowledge, this thesis work has not been submitted for any other degree in any institutions.

Date..19. March 2021

Dr. Nanda Bahadur Singh
Professor
Central Department of Zoology
Tribhuvan University
Kirtipur, Kathmandu, Nepal
Now,
Vice-Chancellor
Mid-Western University
Birendranagar, Surkhet

LETTER OF APPROVAL

On the recommendation of supervisor "**Prof. Dr. Nanda Bahadur Singh**" this thesis submitted by **Manoj Kumar Tharu** entitled "**PREVALECE OF SICKLE CELL DISEASE IN THARU COMMUNITIES OF RUPANDEHI DISTRICT, NEPAL.**" is approved for the examination in partial fulfilment of the requirements for Master's Degree of Science in Zoology with special paper Ecology.

Date...21 March 2021

.....

Prof. Tej Bahadur Thapa
Head of Department
Central Department of Zoology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

CERTIFICATE OF ACCEPTANCE

This thesis work submitted by Manoj Kumar Tharu entitled "**PREVALENCE OF SICKLE CELL DISEASE IN THARU COMMUNITIES OF RUPANDEHI DISTRICT, NEPAL**" has been accepted as a partial fulfillment for the requirements of Master's Degree of Science in Zoology with special paper Ecology.

EVALUATION COMMITTEE

.....
Supervisor
Nanda Bahadur Singh, PhD
Professor
Central Department of Zoology
Tribhuvan University
Kirtipur, Kathmandu
Now,
Vice-Chancellor
Mid-Western University
Birendranagar, Surkhet

.....
Head of Department
Tej Bahadur Thapa, PhD
Professor
Central Department of Zoology
Tribhuvan University
Kirtipur, Kathmandu

.....
External examiner
Dr. Rohit Kumar Pokharel
Professor
Dept. of Orthopaedic
Institute of Medicine
T.U. Teaching Hospital
Maharajgunj, Kathmandu, Nepal

.....
Internal Examiner
Dr. Hari Prasad Sharma
Associate Professor
Central Dept. of Zoology
Tribhuvan University
Kathmandu, Nepal

Date of Examination: 05/04/2021

ACKNOWLEDGEMENTS

First, I'm thankful to my supervisor Prof. Dr. Nanda Bahadur Singh, CDZ, TU for his due support, suggestion, encouragement and help during my thesis work and lab work. I thank to our ex head of department Prof. Dr. Ranjana Gupta and current head of department Prof. Dr. Tej Bahadur Thapa, CDZ, TU including all faculty members and staffs for their due help from their places.

I am very much thankful to all those Tharu peoples of Rupandehi district, Nepal for providing me their DNA samples, supporting me to do research on their genomics and sharing their ethnic, cultural, religious and socio-economic practices with me to investigate and prepare my research design. They are credited at most regarding this thesis work.

I am also thankful to University Grant Commission, Sanothimi, Bhaktapur for providing me fund for my research work and District Public Health Office, Rupandehi and local governances of Rupandehi district for providing me the permission to do such an ethno-biological research.

At last I thank my all friends and colleagues to share their ideas helps to write this report. I feel myself very lucky to receive blessings, supports, economic helps and encouragement to proceed ahead from my parents Mr. Ram Bahadur Tharu and Mrs. Surya Kumari Tharu, my brother Mr. Ram Shankar Tharu and all remaining family members.

Manoj Kumar Tharu

CONTENTS

Pages

LIST OF TABLES

LIST OF FIGURES

LIST OF PHOTOGRAPHS

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ACS	Acute Chest Syndrome
ASS	Acute Splenic Sequestration
CBS	Central Bureau of Statistics
d.f.	degree of freedom
DDC	District Development Committee
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
GBD	Global Burden of Disease Study
Glu	Glutamic acid
Hb	Haemoglobin
HbA	Haemoglobin A
HbC	Haemoglobin C
HbE	Haemoglobin E
HbF	Haemoglobin F
HbS	Sickle Haemoglobin
HLA	Human Leukocyte Antigen
HSCT	Hematopoietic Stem Cell Transplantation
MCH	Mean Corpuscular Haemoglobin
MCV	Mean Corpuscular Volume
NBS	New Born Screening
NHLBI	National Heart, Lung and Blood Institute
NHRC	National Health Research Council
NLM	National Library of Medicine
NSAIDs	Non-Steroidal Anti-inflammatory Drugs
PAH	Pulmonary Arterial Hypertension
RBCs	Red Blood Corpuscles
RM	Rural Municipality
SCA	Sickle Cell Anemia
SCD	Sickle Cell Disease
SCT	Sickle Cell Trait
SM	Sub Metropolitan
SNP	Single Nucleotide Polymorphism
TRV	Tricuspid Regurgitant Jet Velocity
USA	United States of America
VDC	Village Development Committee
VOC	Vaso Occlusive Crisis
WHO	World Health Organization

ABSTRAT

The general objective of this work is to find the prevalence of Sickle Cell Disease in Tharu communities of Rupandehi district, Nepal and specific objectives are to compare prevalence of sickle cell disease among different sub-ethnic groups of Tharu community of Rupandehi district, Nepal; to find out genotype and allele frequency of SCD and sickle cell allele respectively in Tharu community of Rupandehi district, Nepal and to compare the preference between sickling test and sickle cell solubility test method for screening of SCD.

The blood samples were collected randomly from members of Tharu community of Rupandehi district aged above 3 years.. Sickling test and sickle cell solubility test were conducted on all samples for screening of presence of sickle haemoglobin(HbS) (*HBB*: c.20A>T) in the blood. The samples detected to be positive from either of these two methods were undertaken for molecular test by PCR-RFLP method for confirmation and genotyping of samples. The test preference was compared between sickling and solubility test by setting four criteria; cost, time, reliability and field-based feasibility.

Total 386 samples were collected from eight sub-ethnic groups of Tharu ethnic group. Both male and female from different nine municipalities of Rupandehi participated in this study. Out of them, 67 were found to be positive for sickling test and 68 for solubility test. Total 69 positive samples from these two methods were prescribed for molecular test and out of them two were found to be homozygous dominant(AA), 57 heterozygous(AS) and remaining 10 recessive for sickle cell trait. Kanwar Tharu showed highest prevalence(53.94%), Dadaha 46.87%, Kathariya 18.40%, Khausiya 9.52%, Dangoriya 6.94%, Baatar 6.09% and Marchaha Tharu showed lowest prevalence with no occurrence of sickle haemoglobin. The genotypic frequency was found 82.64%, 14.77% and 2.52% for wild, hetero and mutant respectively. The allele frequency of HbA and HbS was 0.9 and 0.1 respectively. Out of three methods, sickling test was found to be cheapest, short time-consuming, reliable and field-based method for screening of Sickle Cell Disease.

The prevalence of sickle cell allele is not common in Tharu people of Rupandehi district, Nepal, $\chi^2(2, N = 386) = 686.389$, $p < .00001$). The distribution of sickle cell allele is cast-dependent, $\chi^2(7, N = 386) = 58.3823$, $p < 0$) and Kanwar Tharu and Dadaha Tharu are most severely affected sub-ethnic groups with nearly 50% occurrence of HbS allele in their gene pool. The sickling test is cost-efficient and reliable field-based method for mass screening of sickle haemoglobin in the blood.

1. INTRODUCTION

1.1 Background

Sickle Cell Disease (SCD) is a group of blood disorders inherited from parents. The most common type is Sickle Cell Anemia (SCA). It results in an abnormality in the oxygen-carrying protein haemoglobin found in RBCs. This leads to a rigid sickle like shape under certain circumstances (NHLBI 2015). After maturation RBC acquire the remarkable ability of being deformable in response to external forces (Huisjes et al. 2018) and use this in order to pass through the narrowest blood capillaries (Viallat and Abkarian 2014). The importance of this characteristic becomes more evident when defects and abnormalities related to RBC shape and/or deformability lead to drastic and premature cell clearance. These changes can provide key information in establishing a differential diagnosis and categorizing different diseases (Ford 2013).

Sickle cell disease is caused by a variant of beta-globin gene called sickle haemoglobin (HbS). Inherited autosomal recessively, either two copies of HbS or one copy of HbS plus another beta globin variant (such as HbC) are required for disease expression (Ashley- Koch et al. 2000). The gene defect is a known SNP causing mutation of a single nucleotide (GAG codon changing to GTG) of the beta globin gene, which results in glutamic acid (Glu) being substituted by valine(Val) at position six (Wellstein and Pitschner 1988). The allele responsible for SCA isfound on the short arm of chromosome 11, more specifically11p15.5 (Allison 2009). Currently, the most common screening techniques include sickle solubility testing, hemoglobin electrophoresis, high-performance liquid chromatography (HPLC), and isoelectric focusing (IEF) (Bonham et al. 2010).

Symptoms include chronic anemia, acute chest syndrome, stroke, splenic and renal dysfunction, pain crisis and susceptibility to bacterial infection. Pediatric mortality is primarily due to bacterial infection and stroke (Ashley-Koch et al. 2000). Sickle cell disease is a heterogeneous disorder, with clinical manifestations including chronic hemolysis, an increased susceptibility to infectious and vaso-occlusive complications often requiring medical care. Patients with sickle cell disease can develop specific and sometimes life-threatening complications, as well as extensive organ damage reducing both quality of life and their life expectancy (Schnog et al. 2004).

Hydroxyurea is a potent inducer of fetal hemoglobin, and evidence over the past 25 years has documented its laboratory and clinical efficacy for both adults and children with SCA. The potential utility of hydroxyurea for all patients with SCA is clear and indisputable (McGann and Ware 2011). Hydroxyurea was introduced as a treatment for sickle cell anemia (homozygosity for HBB glu6val) more than 25 years ago based on its ability to increase fetal hemoglobin (HbF) level (Steinberg 1999). Hydroxyurea therapy is also monitored by a number of laboratory parameters which include increased HbF levels, mean corpuscular volume (MCV), and reduction in WBC count. Hydroxyurea has been found to be effective in the prevention of brain injury due to cerebrovascular disease (Ware et al. 1995). long-term use of hydroxyurea is safe and might decrease mortality (Steinberg et al. 2010). Pediatric studies also have shown that hydroxyurea can safely be used in children (Anderson 2006). A consensus exists that hydroxyurea should be more widely used (Brawley et al. 2008).

The only cure that is available for sickle cell disease is stem cell transplantation (SCT), which replaces the host's bone marrow with stem cells containing normal -globin genotype. Since the first successful transplant reported in 1984 (Johnson et al. 1984), there has been significant reduction in risks due to SCT and increasing success, with the best results, of up to 85% event free survival, occurring with HLA-matched sibling donors and transplantation early in the course of the disease before end-organ damage occurs (Walters et al. 2000). One limitation of SCT is the availability of sibling donors (Krishnamurti et al. 2003) , and therefore there have been attempts to improve survival for unrelated stem-cell donors (Adamkiewicz et al. 2004). The second limitation of SCT is that this line of treatment requires tremendous resources, and it becomes increasingly difficult for transplant physicians practicing in the developing world to reconcile the difference between what is possible and what is available (Makani et al. 2013).

Sickle cell disease has been reported in Hispanic, Caucasian, Native American, and Asian, presumably owing to immigration patterns (Shafer et al. 1996). It is also endemic among Tharu people of Nepal and India; however, they have a sevenfold lower incidence of malaria despite living in a malaria infested zone (Nepali Times 2015). There is no yet documented record available to demonstrate the exact number of people in Nepal affected by SCA, trait or disease (NHRC 2018). Recently at Bir hospital, people from other casts were diagnosed with SCA and NHRC is conducting research on SCA among Tharu community, but the research should include other communities also living in the Terai (The Himalayan Times 2018).

The Tharu people are an ethnic group indigenous to the southern foothills of the Himalayas; most of the Tharu people live in the Nepal Terai (Guneratne 2002). The Tharu people are an ethnic group indigenous to the Terai in southern Nepal and northern India (Bista 1971). They are recognized as an official nationality by the Government of Nepal (Lewis et al. 2014). As of 2011, the Tharu population of Nepal was censused at 1,737,470 people, or 6.6% of the total population (CBS 2012). There are several endogamous sub-groups of Tharu that are scattered over most of the Terai (McLean 1999). Contemporary medical research comparing Tharu and other ethnic groups living nearby found an incidence of malaria nearly seven times lower among Tharu (Terrenato et al. 1988).

SCA is documented to be prevalent in Terai belt of Nepal specially in Tharu community (NHRC 2018). SCA, mostly seen in people of Tharu community of Nepal, has been seen in people of other communities also and is not limited to western Terai but also seen in eastern Terai. As per study conducted by the Institute of Medicine in 2012, the prevalence of SCA among Tharu community is 80 to 90 percent and that of in other communities is two to four percent (The Himalayan Times 2018). Although, SCD individuals have an element of protection against malaria; with a lower prevalence of malaria infection (Kotila et al. 2007) and a lower parasite density (Awotua-Efebo et al. 2004) , the risk of mortality when SCD patients get malaria is significantly higher (Makani et al. 2010).

1.2 Objectives

The general objective of this research is to study prevalence of sickle cell disease in Tharu communities of Rupandehi district, Nepal and specific objectives are:

- To compare prevalence of sickle cell disease among different sub-ethnic groups of Tharu community of Rupandehi district, Nepal.
- To find out genotype and allele frequency of SCD and sickle cell allele respectively in Tharu community of Rupandehi district, Nepal.
- To test preference between sickling test and sickle cell solubility test for screening of SCD.

1.3 Research hypothesis

The distribution of HbS allele is common among the Tharu people of Rupandehi district, Nepal as they are living in malarial zone with the common practice of endogamous marriage and prevalence of HbS allele is cast independent within Tharu society.

1.4 Rational of the study

Though there are records of incidence of SCD in the district in government hospital and provincial hospital, but exact data and degree of prevalence of the disease are lacking in Rupandehi district. The victimized families are facing financial constrain due to such incurable disease requiring advanced medical care and frequent blood transfusion. Tharus still believe in in their traditional treatment dependent on herbs and supernatural power of so called 'Guruwa'. Due to lack of knowledge of causes and inheritance of this disease, peoples are practicing endogamy without premarital test so the prevalence of this disease is not seem to be decreasing within Tharu society. This study aims to find the prevalence of SCD in different sub-ethnic groups of Tharu community in Rupandehi district. The major limitations of the study are:

- The collection of sample was random and uneven. Sufficient no. of samples from all sub-ethnic groups could not be taken.
- PCR-RFLP test on all samples was not run. So false negative samples could not be identified if present.
- The sequencing of genomic DNA was not performed due to financial constrain, lacking of technology in Nepal and global Covid19 pandemics.
- Tharus nowadays have started inter-caste marriage that remarkably alters the gene frequency of HbS in a given community due to genetic drift and this result may have suffered from the same.

2. LITERATURE REVIEW

2.1 Haemoglobin

2.1.1 Haemoglobin: general overview

The blood's red color is due to the spectral properties of hemic iron ions in haemoglobin. Each human red blood cell contains approximately 270 millions of haemoglobin molecules (Angelo 2017). Hemoglobin is the iron-containing oxygen-transport metalloprotein in the red blood cells (Bruce and O'Brien 2006). Haemoglobin is a respiratory protein of the RBCs that transports oxygen gas from the lungs to the tissues and facilitates the return transport of carbon dioxide gas from the tissues to the lungs (Perutz et al. 1987).

In mammals, the protein makes up about 96% of the red blood cells' dry content (by weight), and around 35% of the total content including water (Weed et al. 1963). Hemoglobin is essential for oxygen transport (Hsia 1998). Hemoglobin comprises four subunits, each having one polypeptide chain and one heme group. All haemoglobin carry the same prosthetic heme group iron protoporphyrin associated with a polypeptide chain of 141 (alpha) and 146 (beta) amino acid residues (Marengo-Rowe 2006). The alpha and beta chains have different sequences of amino acids but fold up to form similar three dimensional structure. Each chain harbors one heme. The heme consists of a ring of carbon, nitrogen and hydrogen atoms called porphyrin, with an atom of iron, like a jewel, at its center. A single polypeptide chain combined with a heme is called a sub-unit of hemoglobin or a monomer of the molecule. In the complete molecule, four sub-units are closely joined to form a tetramer (Perutz 1978).

The mammalian haemoglobin molecule can bind (carry) up to four oxygen molecules (Dominguez de Villota et al. 1981). Haemoglobin is involved in the transport of other gases. It carries some of the body's respiratory carbon dioxide (about 20–25% of the total) as carbaminohaemoglobin, in which CO_2 is bound to the heme protein. The molecule also carries the important regulatory molecule nitric oxide bound to a globin protein thiol group, releasing it at the same time as oxygen (Epstein and Hsia 1998). Haemoglobin is present in blood at concentrations of 13.5–18.0 g dl^{-1} in men and 11.5–16.0 g dl^{-1} in women (Thomas and Lumb 2012).

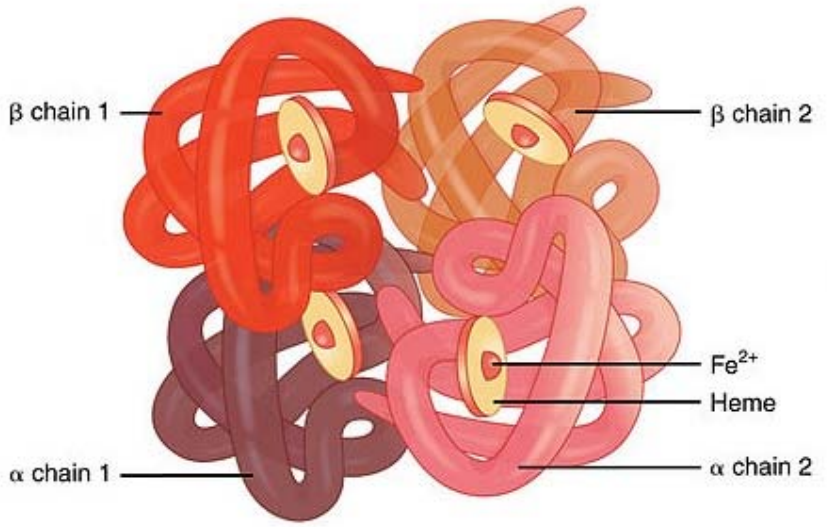


Fig. 2.1: An illustration showing structure and organization of a haemoglobin molecule. A typical adult haemoglobin, HbA is a tetramer and contains two alpha and two beta globin chains. Each globin polypeptide chain has a centrally located heme prosthetic group containing an iron atom.(source: https://pediaa.com/difference-between-hemoglobin-and-myoglobin/).

2.1.2 Haemoglobin synthesis and its common variants

The two main components of hemoglobin synthesis are globin production and heme synthesis. Globin chain production occurs in the cytosol of erythrocytes and occurs by genetic transcription and translation. Many studies have shown that the presence of heme induces globin gene transcription. Genes for the alpha chain are on chromosome 16 and genes for the beta chain are on chromosome 11. Heme synthesis occurs in both the cytosol and the mitochondria of erythrocytes. It begins with glycine and succinyl coenzyme A and ends with the production of a protoporphyrin IX ring. Binding of the protoporphyrin to an Fe2+ ion forms the final heme molecule (Chiabrando et al. 2014).

More than 1000 naturally occurring human hemoglobin variants with single amino acid substitutions throughout the molecule have been discovered (Thom et al. 2013). Normal adult red blood cells have three different types of haemoglobin: HbA ($\alpha_2\beta_2$) ~95%; HbA_2 ($\alpha_2\delta_2$) ~2.5%; and HbF ($\alpha_2\gamma_2$) 2.5% (Wilson et al. 2010). The most common type of hemoglobin in the adult is HbA, which comprises two alpha-globin and two beta-globin subunits. Different globin genes encode each type of globin subunit (Hafen and Sharma 2020).

HbF has a stronger oxygen affinity than HbA, allowing oxygen to flow from maternal to fetal circulation through the placenta. Production of HbF drops significantly after birth, reaches low, near-adult, levels by two years, and ultimately makes up 2-3% of hemoglobin in adults. HbA, the most common adult form of hemoglobin, comprises two alpha and two beta-globin subunits. Inversely to HbF, HbA production explodes after birth and ultimately makes up 95-98% of hemoglobin in adults. HbA2 is a less common adult form of hemoglobin. It comprises two alpha and two delta-globin subunits and makes up 1-3% of hemoglobin in adults (Harewood and Azevedo 2020).

2.2 Mutation

A mutation is an alteration in the nucleotide sequence of the genome of an organism, virus, or extrachromosomal DNA(Nature 2018). A mutation is a change in a DNA sequence. Mutations can result from DNA copying mistakes made during cell division, exposure to ionizing radiation, exposure to chemicals called mutagens, or infection by viruses(National Human Genome Research Institute).

Mutations are one of the fundamental forces of evolution because they fuel the variability in populations and thus enable evolutionary change. Based on their effects on fitness, mutations can be divided into three broad categories: the 'good' or advantageous that increase fitness, the 'bad' or deleterious that decrease it and the 'indifferent' or neutral that are not affected by selection because their effects are too small (Loewe and Hill 2010).Mutations result from errors during DNA or viral replication, mitosis, or meiosis or other types of damage to DNA (such as pyrimidine dimers caused by exposure to ultraviolet radiation (Rodgers and McVey 2016) or may also result from insertion or deletion of segments of DNA due to mobile genetic elements (Bertram 2000).

Mutations can involve the duplication of large sections of DNA, usually through genetic recombination (Hastings et al. 2009). Changes in chromosome number may involve even larger mutations, where segments of the DNA within chromosomes break and then rearrange (Zhang et al. 2004). Mutations are of several types. Changes within genes are called point mutations. The simplest kinds are changes to single base pairs, called base-pair substitutions. Many of these substitute an incorrect amino acid in the corresponding position in the encoded protein, and of these a large proportion result in altered protein function. Mutations that span more than one gene are called chromosomal mutations because they affect the structure, function, and inheritance of whole DNA molecules. Often these chromosome mutations result from one or more coincident breaks in the DNA molecules of the genome or gain of whole chromosomes results in a condition called aneuploidy (Griffiths 2020). Mutations can be subdivided into germline mutations, which can be passed on to descendants through their reproductive cells, and somatic mutations, also called acquired mutations, which involve cells outside the dedicated reproductive group and which are not usually transmitted to descendants (Genome dictionary 2007).

A 2007 study on genetic variations between different species of *Drosophila* suggested that, if a mutation changes a protein produced by a gene, the result is likely to be harmful, with an estimated 70% of amino acid polymorphisms that have damaging effects, and the remainder being either neutral or marginally beneficial (Sawyer et al. 2007). Humans on average pass 60 new mutations to their children but fathers pass more mutations depending on their age with every year adding two new mutations to a child (Jha 2012). Due to the damaging effects that mutations can have on genes, organisms have mechanisms such as DNA repair to prevent or correct mutations by reverting the mutated sequence back to its original state (Bertram 2000).

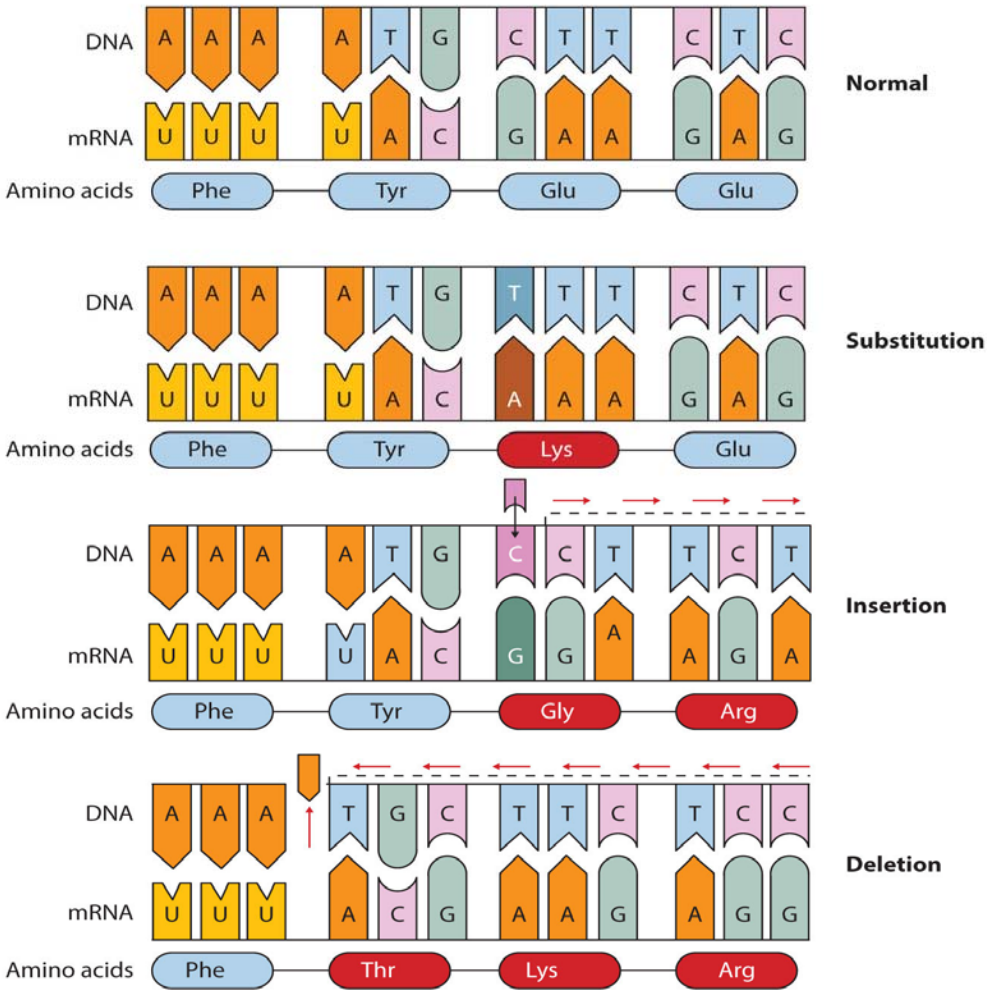


Fig. 2.2: An illustration showing different types of point mutation. A substitution mutation occurs when a nucleotide of a gene is replaced by another, an insertion mutation occurs when a new nucleotide is inserted into a gene and a deletion mutation occurs when a nucleotide from a gene is deleted. All these mutations cause change in mRNA sequence and hence synthesis of a new amino acid in the protein chain. (Source: Wikimedia Commons).

2.3 Haemoglobin S

The most common abnormal variant of hemoglobin is HbS (sickle cell hemoglobin). HbS results from a substitution of the sixth amino acid in the beta-globin subunits. The genetic mutation results in replacement of glutamic acid with valine and occurs most frequently in African Americans. Heterozygous individuals have a mutation in only one of the two beta chains, resulting in sickle cell trait. Resistance to falciparum malaria infection and complications are benefits of a sickle cell trait.

Homozygous individuals have mutations in both beta chains, resulting in sickle cell disease. When deoxygenated, HbS causes deformation of erythrocytes from a biconcave disc, to a crescent or "sickle" shape. This change in shape causes damage to erythrocyte membranes, premature destruction of erythrocytes, and chronic hemolytic anaemia. Sickled erythrocytes can obstruct blood flow and cause tissue hypoxia, which can cause severe ischemic pain or even stroke. These patients also have functional asplenia and are at risk for infections with encapsulated organisms (Kato et al. 2018).

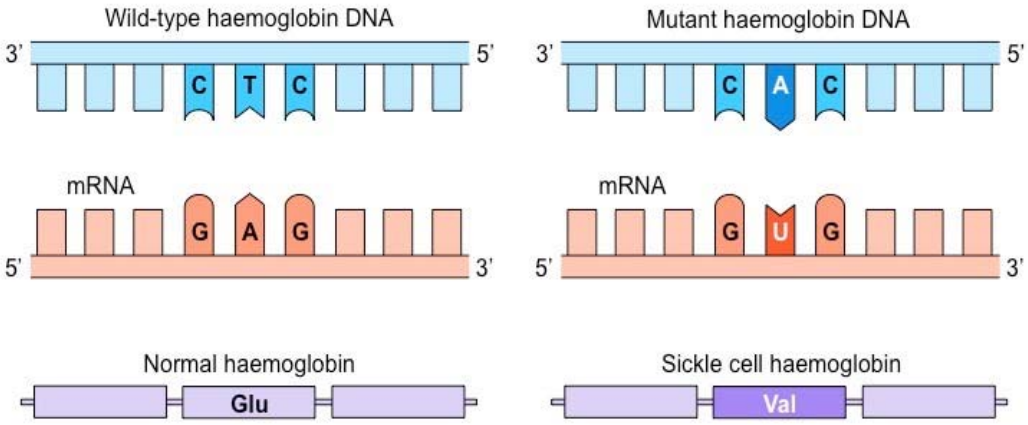


Fig. 2.3: An illustration showing single point mutation in haemoglobin beta gene(HBB) of 11th chromosome resulting in the formation of defecvtive sickle haemoglobin(HbS). Mutation occurs at 6th codon of HBB gene where thymine is replaced by adenine forming valine amino acid in the substitution of glutamic acid. (Source: BioNinja. https://ib.bioninja.com.au/standard-level/topic-3-genetics/31-genes/mutations.html).

2.4 Resistance to malaria

Malaria, the parasitic disease of the red blood cell caused by *Plasmodium sp.* (Kwiatkowski 2005). The strong selective pressure that malaria exerts has been shaping the human genome for at least 6,000 years (Rich et al. 1998) and the disease is the driving force behind several genetic disorders, such as the sickle cell trait, glucose-6-phosphate dehydrogenase deficiency and thalassemia (Kwiatkowski 2005). Malaria is also endemic in South Asian countries such as India, Nepal, Bhutan, and Sri Lanka (Kondrashin 1992). In Nepal, for example, the WHO reported 4,637 cases in 2004, from a population of 27.3 million; 7.9% of the cases were caused by *P. falciparum* (Caetano et al. 2006)). Few studies of resistance have been performed in this region (Wattavidanage et al. 1999). The lack of data on malaria-protective genes in South Asia is surprising since some ethnic groups, such as the Tharu in Nepal, have long been known to be resistant to the disease (Terrenato et al. 1988). A large part of Nepal consists of mountains where malaria cannot be transmitted (Gillies 1988), but the southern part of the country, the Terai area where malaria cases are concentrated is flat and exposed to the malaria parasites *P. vivax* and *P. falciparum* (Sherchand et al. 1995). In the late 1950s, the World Health Organization supported the Nepalese government in eradicating malaria in the forests of the central Terai (Brydon et al. 1961).

Malaria is perhaps the most important parasitic infection and strongest known force for selection in the recent evolutionary history of the human genome. Genetically-determined resistance to malaria has been well-documented in some populations, mainly from Africa. The disease is also endemic in South Asia, the world's second most populous region, where resistance to malaria has also been observed, for example in Nepal (Caetano et al. 2006). Resistance to malaria is widespread among the Tharu (Modiano et al. 1991). This may support an origin for the Tharu within the Tarai jungles as prolonged exposure to swampy, that is, malaria-rich environments which could have led to natural selection favouring resistance (Terrenato et al. 1998). Reports suggest a high SCD burden among the indigenous Tharu population of Nepal, who for centuries have inhabited regions where malaria is endemic. Unfortunately, health care resources are limited and often inaccessible for Tharu individuals suffering from SCD (Marchand et al. 2017). The Tharu were famous for their ability to survive in the most malarial parts of the Terai that were deadly to outsiders. In 1902, a British observer noted:" Plainsman and Pahari generally die if they sleep in the Terai before November 1 or after June 1" (Guneratne 2002).

Contemporary medical research comparing Tharu and other ethnic groups living nearby found an incidence of malaria nearly seven times lower among Tharu (Terrenato et al. 1988). The researchers believed such a large difference pointed to genetic factors rather than behavioral or dietary differences. This was confirmed by follow-up investigation finding genes for thalassemia in nearly all Tharu studied (Luzzatto 2012). In India, tribal communities are highly vulnerable to hereditary diseases. The sickle cell haemoglobinopathy and G6PD enzyme deficiency are important genetic and public health problems in Central-Eastern part of India (Balgir 2006).

2.5 Sickle cell disease

2.5.1 History of sickle cell disease

There are no clear historical reports elucidating the discovery of SCD. At the middle of 18th century (1846), SCD was reported from adult African slaves struggling for life and death. While at the end of 19th century, it was revealed that individuals with SCD acquires immunity against malaria (Ballas 2015). The first peculiar, elongated and sickle shaped RBCs have been reported jointly by Ernest E. Irons and James B. Herrick in 1910 during microscopic examination of human blood and the term sickle cell anemia was first used by Verne mason (Serjeant 2010). The molecular cause of SCD was first discovered by Ingram in1956. He reported substitution of glutamic acid by valine at sixth codon of β- globin of haemoglobin (Ingram 1956). During the period of 1960 to 1970, the mechanism of abnormal deoxy HbS polymerization and its various patho-physiological schemes were extensively elucidated (Bunn and Forget 1986).

2.5.2 Geographical distribution and morbidity of SCD

Approximately 3 00,000 infants are born per year with SCD globally (Azar and Wong 2017). It is one of the most common serious genetic disease in childhood, affecting approximately one in 2,500 births and 100,000 individuals in the USA (Strouse 2016). As of 2015, about 4.4 million people have SCD, while an additional 43 million have sickle cell trait. In 2015, it resulted in about 114,800 deaths (GBD 2015). About 80%of SCD cases are believed to occur in Sub-Saharan Africa (Rees et al. 2010). It also occurs relatively frequently in parts of India, the Arabian Peninsula and among people of African origin living in other parts of the world (Elzouki 2012).

Malaria remains one of the most prevalent parasitoses worldwide. About 350 to 500 million febrile episodes are observed yearly in African children alone and more than one million people die because of malaria each year (Hassell et al. 2010). Three-quarters of sickle cell cases occur in Africa (WHO 2011). It is thus not surprising that the highest prevalence of Hb S in the world remains in sub-Saharan Africa with the Middle East and Indian subcontinents as close seconds (Hassell 2010).

The maintenance of high frequencies of the sickle cell trait in the presence of almost obligatory losses of homozygotes in Equatorial Africa implied that there was either a very high frequency of HbS arising by fresh mutations or that the sickle cell trait conveyed a survival advantage in the African environment (Serjeant 2001). SCD has been reported in persons who self-identify as Hispanic, Caucasian, Native American, and Asian, presumably owing to immigration patterns (Shafer et al. 1996).

For each World Health Organization region, annual rates of neonates born with SCD are estimated at 230,000 in the African Region, 43,000 in the South-East Asia Region, 13,000 in the Region of the Americas, 10,000 in the Eastern Mediterranean Region,3500 in the European Region, and 4 in the Western Pacific Region per year (Piel 2016). It is concluded that the incidence of SCA in Central India is high (Erazo et al. 2012). The prevalence has ranged from 9.4 to 22.2% in endemic areas of Madhya Pradesh, Rajasthan and Chhattisgarh (Awasthy et al. 2008). It is also endemic among Tharu people of Nepal (Nepali Times 2015). Despite some level; of progress in the sickle cell projections, there is no yet documented record available to demonstrate the exact number of people in Nepal affected by SCA, trait or disease (NHRC 2018).

About 90% people with SCD survive to age 20, and close to 50% survive beyond age 50 (Kumar et al. 2009). In 2001, according to one study performed in Jamaica, the estimated mean survival for people with sickle cell was 53 years old for men and 58 years old for women with homozygous SCD (Wierenga et al. 2001). The specific life expectancy in much of the developing world is unknown (Costa and Conran 2016).

2.5.3 Patho-physiology of SCD

SCD is one of the most common severe monogenic disorders in the world. Haemoglobin polymorphism, leading to erythrocyte rigidity and vaso-occlusion, is central to the pathophysiology of this disease, although the importance of chronic anaemia, hemolysis, and vasculopathy has been established (Rees et al. 2010). SCD is a heterogeneous inherited disorder of haemoglobin that causes chronic hemolytic anaemia, vaso-occlusion, and endothelial dysfunction (Noronha et al. 2016). SCD is the most common haemoglobinopathy worldwide, particularly in Africa and among people of African descent (Marchand et al. 2017). SCD can cause severe pain, significant end-organ damage, pulmonary complications, and premature death (Estcourt et al. 2016). SCD is an inherited monogenetic disease characterized by misshapen RBCs that causes vaso-occlusive disease, vasculopathy, and systemic inflammations.

Acute, chronic and acute-on-chronic complications contribute to end-organ damage and adversely affect quantity and quality of life (Azar and Wong 2017). Pulmonary arterial hypertension (PAH), once considered a rare complication of sickle cell disease (SCD) and thalassemia, appears to be more common in adults with haemoglobinopathy than previously appreciated.

On prospective screening of adults with SCD, approximately one-third of adults are found on echocardiography to have a tricuspid regurgitant jet velocity (TRV) of 2.5 m/s or higher, many of whom are asymptomatic (Vichinsky 2004). Based on a prospective cohort study of 1056 patients with SCA predicted an increased likelihood of developing organ damage and death: 77% patients with chronic lung disease, 75% of those with renal insufficiency, and 51% of those with stroke had a prior chronic condition. Of the 232 patients who died, 73% had one or more clinically recognized forms of irreversible organ damage.

By the fifth decade, nearly one-half of the surviving patients (48%) had documented irreversible organ damage. End-stage renal disease (glomerulosclerosis), chronic pulmonary disease with pulmonary hypertension, retinopathy, and cerebral micro infarctions are manifestations of arterial and capillary microcirculation obstructive vasculopathy (Powars et al. 2005). CNS injury is the most debilitating frequent complication of SCD and includes stroke, silent cerebral infarcts (SCI) and cognitive impairment. Among children with SCA, 11% had stroke by age of 18 years and 27% had SCI by their fifth birthday (Strouse 2016).

One of the main problem of SCD in children is the development of cerebrovascular disease and cognitive impairment. Recurrent episodes of vaso-occlusion and inflammation result in progressive damage to most organ including the brain, kidneys, lungs, bones and cardiovascular system, which become apparent with increasing age (Rees et al. 2010).

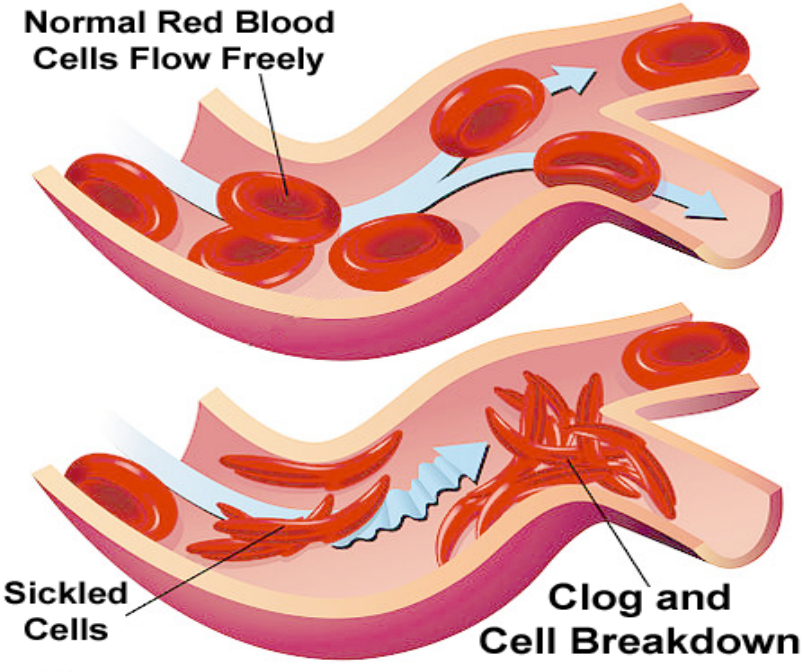


Fig. 2.4: An illustration showing the flow of normal and sickled erythrocytes in the fine blood vessels. The upper one has normal erythrocytes trhat are flexible and thus can squeeze throuth narrow blood vessels and flow smoothly. The lower has sickled erythrocytes that clog together blocking the smaller blood vessels/capillaries causing a serious condition called vaso-occlusion. (Source: https://bioinformatics.org/jmol-tutorials/jtat/hemoglobin/images4all/sickle4.png).

2.5.4 Inheritance of SCD

Normally, humans have haemoglobin A, which consists of two alpha and two beta chains, haemoglobin A_2, which consists of two alpha and two delta chains, and haemoglobin F, consisting of two alpha and two gamma chains in their bodies. Out of these three types, HbF dominates until six weeks of age. Afterwards, HbA dominates throughout the life. In people diagnosed with SCD, at least one of the beta-globin subunits in HbA is replaced with what's known as HbS. In SCA, a common form of SCD, HbS replaces both beta globin subunits in the haemoglobin (NLM 2018).

Sickle cell conditions have an autosomal recessive pattern of inheritance from parents. The types of haemoglobin a person makes in the red blood cells depend on what haemoglobin genes are inherited from his or her parents. If one parent has SCA and other has sickle cell trait, then the child has a 50% chance of SCD and a 50% chance of sickle cell trait. When both parents have sickle cell trait, a child has a 25% chance of SCD, 25% do not carry any sickle cell alleles, and 50% have the heterozygous condition (Genetics Home 2016). A person who receives defected genes from both father and mother develops the disease; a person who receives one defective and one healthy allele remains healthy but can pass on the disease and is known as the carrier or heterozygote (Allison 2009).

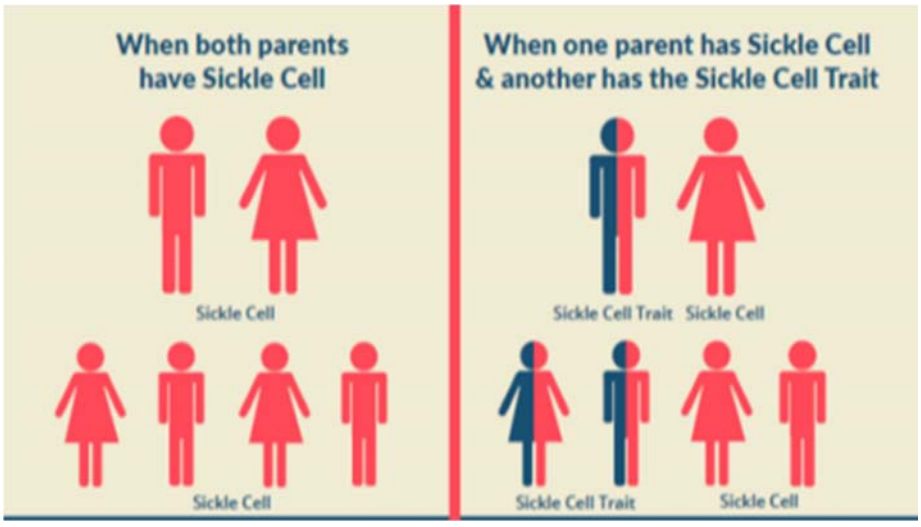


Fig. 2.5.1: An illustration showing inheritance pattern of sickle cell anemia. If both parents are sickle cell anaemic, all their chioldren will suffer from same disease. If one parent has SCA and another has SCT, the ratio of having homozygous and heterozygous children is 1:1. (source: https://blog.scdhec.gov/2018/09/19/raising-the-awareness-bar-on-sickle-cell-trait-and-sickle-cell-disease/).

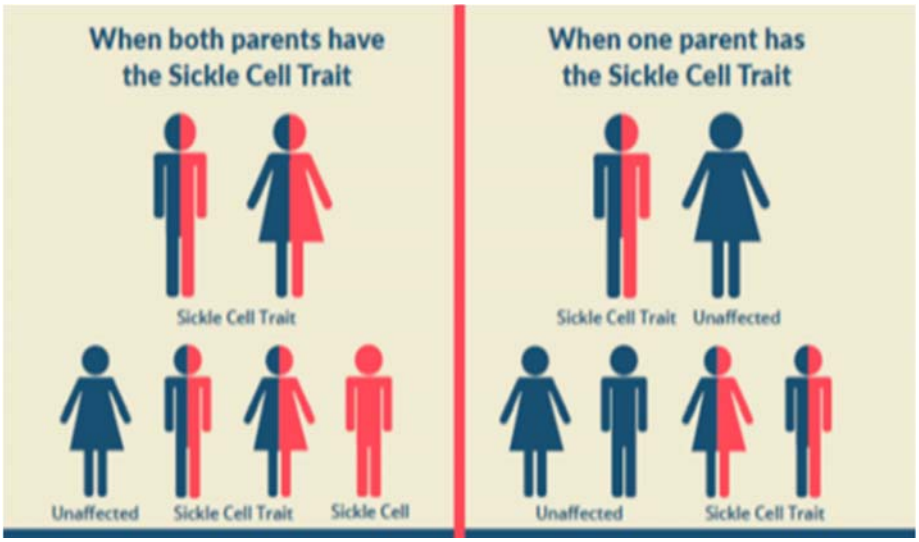


Fig. 2.5.2: An illustration showing inheritance pattern of sickle cell anemia. If both parents have SCT, the genotypic ratio of wild,hetero and mutant of F1will be 1:2:1. If one of the parent has SCT while another being wild, the genotypic ratio of wild, hetero and mutant of F1 will be 1:1:0. SCA is a homozygous recessive disease and follow typical Mendelian inheritance. (Source:https://blog.scdhec.gov/2018/09/19/raising-the-awareness-bar-on-sickle-cell-trait-and-sickle-cell-disease/).

2.5.5 SCD at molecular basis

The molecule, DNA, is the fundamental genetic material that determines the arrangement of amino acid building blocks in all proteins. Segments of DNA that code for particular proteins are called genes. The gene that controls the production of beta globin subunit of haemoglobin is located in chromosome 11. Ordinarily, the Hb molecules exist as single, isolated unit in RBCs., whether they have oxygen bound or not. Sickle haemoglobin exists as single unit when they have oxygen bound but when they release the oxygen in the peripheral tissue, the molecules tend to stick together and form long chains or polymers. The abnormal valine amino acid at position 6 in the beta globin chain interacts weakly with beta globin in an adjacent sickle haemoglobin molecule. These rigid polymers distort the cell and cause it to bend out of shape. Most distorted cells are shaped irregularly, a few have a crescent-like shape under microscope. These crescent-like or sickle shaped red cells gave the disorder its name (www.sickle.bwh.harward.edu).

The deoxy form of haemoglobin exposes a hydrophobic patch on the protein between E and F helices (Phe85, Leu88). The hydrophobic side chain of the valine residue at position 6 of the beta chain in haemoglobin is able to associate with the hydrophobic patch, causing HbS molecules to aggregate and form fibrous precipitates (Wellstein and Pitschner 1988).

2.5.6 Diagnosis of SCD

The diagnostic methods can be categorized into three basic types, namely screening test, confirmatory test and molecular diagnosis. In most hospitals, sickling test is most commonly done for screening of SCA. It involves preparation of thin blood film under hypoxic condition by treatment with sodium metabisulphite that causes sickling of erythrocytes and is observed under microscope(Makani et al. 2013). Currently, the most common screening techniques include sickle solubility testing, hemoglobin electrophoresis, high-performance liquid chromatography (HPLC), and isoelectric focusing (IEF). The sickle solubility test is a low-cost assay that relies on the relative insolubility of HbS in the presence of a reducing agent, such as sodium dithionite, by detecting turbidity or crystal formation from lysis of HbS-containing erythrocytes. Because it only detects the presence or absence of sickle hemoglobin, the solubility test cannot differentiate individuals with SCD and SCT and can be falsely negative in infants with high hemoglobin F or in individuals with very low percentage HbS (<10%), making confirmatory testing essential (Bonham et al. 2010).

Hemoglobin electrophoresis, HPLC, and IEF are methods used either for primary identification of SCT or as confirmatory tests. These techniques can provide discrimination and relative quantification of haemoglobins, allowing for differentiation of SCT from SCD syndromes. Hemoglobin electrophoresis, an inexpensive and frequently used technique, uses the principles of gel electrophoresis to separate hemoglobin molecules by size and charge. IEF is a highly sensitive, discriminatory pH-based electrophoresis technique that identifies haemoglobins by their isoelectric point. HPLC and capillary electrophoresis have also been adopted for haemoglobinopathy screening by many reference laboratories, owing to their ability to more precisely quantify hemoglobin components.

Molecular protocols for hemoglobinopathies are often used in the research setting to identify SCT carriers using banked DNA samples (Benson and Therrell. 2010). The most popular molecular diagnosis of sickle cell mutation, based on restriction enzyme digestion, is performed on HBB PCR products. The point mutation, which results in SCD, abolishes the restriction site for the restriction enzyme *DdeI* (Makani et al. 2013). Recent rapid advances in technology have allowed for detection of SCT from DNA through exome sequencing (Auer et al. 2012). Restriction Fragment Length Polymorphism (RFLP) is a technique that exploits variations in homologous DNA sequences. In RFLP analysis, the DNA sample is broken into pieces (digested) by restriction enzymes and the resulting restriction fragments are separated according to their lengths by gel electrophoresis (www.wikipedia.org). Restriction endonucleases occur ubiquitously among prokaryotic organisms (Raleigh and Brooks 1998). DNA isolation is a process of purification of DNA from sample using a combination of physical and chemical methods. The first isolation of DNA was done in 1869 by Friedrich Miescher (Dahm 2008).

Restriction enzymes are traditionally classified into four types on the basis of subunit composition, cleavage position, sequence specificity and cofactor requirements. Type I enzymes are complex, multisubunit, combination restriction-and-modification enzymes that cut DNA at random far from their recognition sequences. Type II enzymes cut DNA at defined positions close to or within their recognition sequences. They produce discrete restriction fragments and distinct gel banding patterns, and they are the only class used in the laboratory for routine DNA analysis and gene cloning. Type III enzymes are also large combination restriction-and-modification enzymes. They cleave outside of their recognition sequences and require two such sequences in opposite orientations within the same DNA molecule to accomplish cleavage; they rarely give complete digests. Type IV enzymes recognize modified, typically methylated DNA and are exemplified by the McrBC and Mrr systems of *E. coli* (www.international.neb.com). The Type II restriction enzymes are among the most valuable tools available to researchers in molecular biology. These enzymes recognize short DNA sequences (4–8 nucleotides) and cleave at, or close to, their recognition sites (Pingoud and Jeltsch, 2001).

2.5.7 Management of SCD

SCD is a chronic disease that needs management for recurrent periodic life threatening complications called sickle cell crisis characterized by Acute Chest Syndrome (ACS), Acute Splenic Sequestration (ASS), pain, ulceration, end organ damage, fever, anaemia, jaundice and asthma. Proven effective treatment options for sickle cell patients are limited to hydroxyurea, blood transfusions and bone marrow transplantation (Schnog et al. 2004). Nitric oxide (NO) is a critical factor in the pathophysiology of SCD and is a promising anti sickling agent with vasodilation properties. Although NO is difficult to administer, its precursor, L-arginine, is an oral supplement (Vichinsky 2002). Hematopoietic stem cell transplantation is the only cure available today, but is not feasible for the vast majority of the people suffering from SCD (Azar and Wong 2017). SCD complications begin with the polymerization of HbS. Thus, SCD therapies are focused on HbS production or reducing the circulating amount of HbS. Hydroxyurea treatment has become more widespread, whereas the number of evidence-based indications for erythrocyte transfusion is small. Hematopoietic stem cell transplantation is a curative option for SCD but less than 25% of patients have suitable donor (Meier 2018). The introduction of penicillin prophylaxis, conjugated pneumococcal and Haemophilus influenza type B vaccines have dramatically decreased the rate of life-threatening infections, while use of hydroxyurea in children has decreased pain and acute chest syndrome events. Use of transcranial Doppler coupled with regular blood transfusion has the rate of overt strokes and premature death associated with strokes. Regular blood transfusion reduces incidence of strokes, acute chest pain episodes but is associated with the burden of monthly visits and excessive iron stores (Kassim and DeBaun 2014). All health professionals should be aware of the common acute and chronic complications of SCD and the basic principles of their management (Olujohungbe and Howard 2008). Followings are the chief SCD management strategies accepted globally:

1. Neonatal and newborn screening
2. Prevention and treatment of infections
3. Blood transfusion
4. Pain management
5. Hydroxyurea
6. Stem cell transplantation

2.6 Tharu: history and distribution

The Rana Tharus claim to be of Rajput origin and to have migrated from the Thar Desert to Nepal's Far Western Terai region. Tharu people farther east claim to be descendants of the Sakya and Koliya peoples living in Kapilvastu. The word Tharu is thought to be derived from sthavir meaning follower of Theravada Buddhism (Skar 1995). According to Alberuni, Tharu people have been living in the eastern Terai since at least the 10th century (Krauskopff 1995). The Tharu were described by ethnographers in the nineteenth century as representing either an East Asian phenotype or a somatological amalgam of 'Indo-Chinese' or 'Mongoloid racial stock' and a 'negrito' aboriginal substrate (Buchanan-Hamilton 1886).

The Tharu people are an ethnic group indigenous to the Terai in southern Nepal and northern India (Bista 1971). They are recognized as an official nationality by the Government of Nepal (Lewis et al. 2014). As of 2011, the Tharu population of Nepal was censused at 1,737,470 people, or 6.6% of the total population (CBS 2012). In 2009, the majority of Tharu people were estimated to live in Nepal (CBS 2009). Tharu is an 'umbrella-ethnonym', uniting a variety of historically endogamous groups that speak different dialects and are also distinct in many socio-cultural aspects (Chaubey et al. 2014). There are several endogamous sub-groups of Tharu that are scattered over most of the Terai (McLean 1999).

Rana Tharu in the Kailali and Kanchanpur Districts of the far western Nepal Terai (Krauskopff 1995), also in India, in Udham Singh Nagar district, Uttarakhand and Kheri district in Uttar Pradesh. Rana Tharu claim Rajput origin (Lewis et al. 2014). Kathariya Tharu mostly in Kailali District and in India (Krauskopff 1995).

Sonha Tharu in Surkhet District (Lewis et al. 2014). Dangaura Tharu in the western Terai: Dang-Deukhuri, Banke, Bardia, Kailali, Kanchanpur, Rupandehi and Kapilvastu Districts (Lewis et al. 2014). Chitwan Tharu in central Terai: Sindhuli, Chitwan and Nawalparasi Districts (Krauskopff 1995). Kochila Tharu in eastern Terai: Saptari, Bara, Parsa, Rautahat, Sarlahi, Mahottari and Udayapur districts (Lewis et al. 2014). Danuwar in eastern Terai: Udayapur, Saptari and Morang Districts (Lewis et al. 2014). Lampucchwa Tharu in Morang and Sunsari District (Krauskopff 1995).

Smaller numbers of Tharu people reside in the adjacent Indian districts Champaran of Bihar, Gorakhpur, Basti and Gonda districts in Uttar Pradesh and Udham Singh Nagar in Uttarakhand (Rajaure 1981). In 2001, Tharu people were the largest of five scheduled tribes in Uttarakhand, with a population of 256,129 accounting for 33.4% of all scheduled tribes (Office of the Registar General 2001). In the same year, they constituted 77.4% of the total tribal population of Uttar Pradesh with a population of 83,544 (Office of the Registar General 2001).

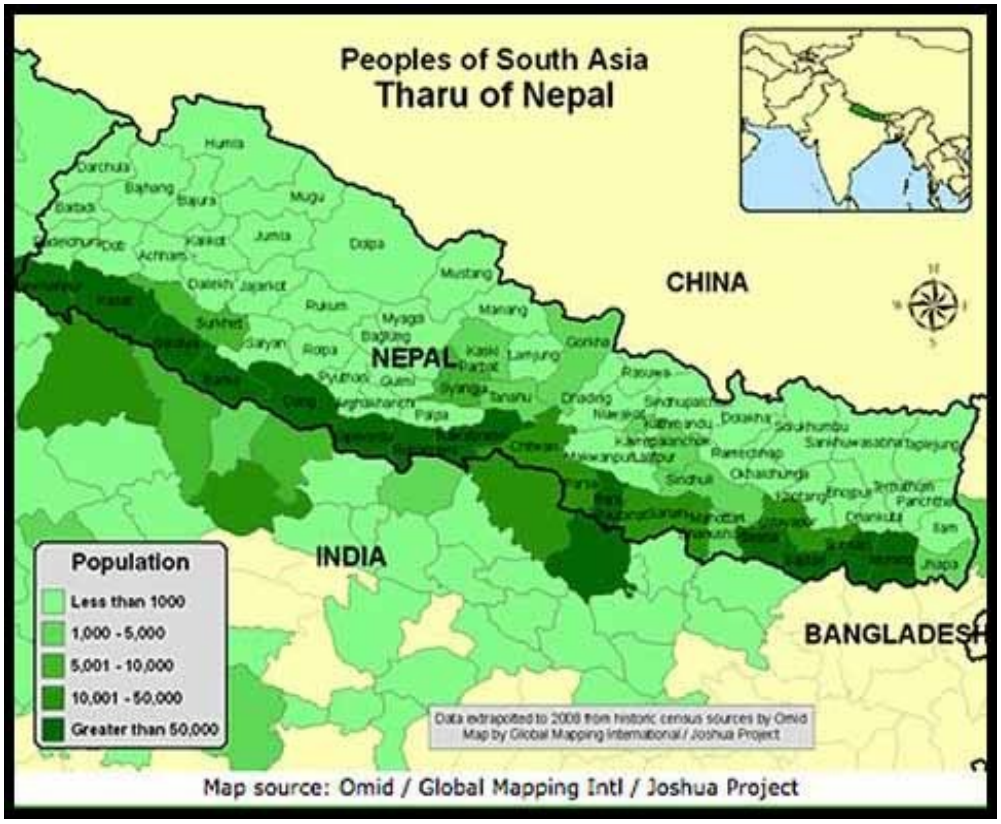


Fig. 2.7: A demograophic map showing distribution of Tharu people in different parts of Nepal and India. Tharu population coverage is colour coded.
(Source:https://www.researchgate.net/figure/The-geographical-distribution-of-the-Tharu-in-South-Asia)

3. MATERIALS AND METHODS

3.1 The study area

The study area of this research is in Rupandehi district(27°20'00"to 27°47'25"N, 83°12'16" to 83°38'16" E) which belongs to province no. 5,Nepal with total area 1360 km^2 and total population 880,196 in2011 (DDC, Rupandehi 2011). The total population in 2016 was 982851 with population density 720/km^2. Major residents belong to Tharu, Yadav, Kewat, Musalman, Magar etc. this district is populated by 86% Hindu. 36.6% people speak Bhojpuri, 6.4% Awadhi and 6.3% Tharu (CBS 2014).

The district lies on the southern and western part of Nepal. The elevation of the district lies between 100 m to 1229 m from sea level with two major cities, Bhairahawa and Butwal with headquarter Bhairahawa. 16.1% part of the district lies in Churia range and rest in Terai region. Rupandehi is surrounded by West Nawalparasi district from east, Palpa from north, Kapilbastu from west and Uttar Pradesh, India from south. 89.3% area lies in lower tropical zone (below 300 m), 10.5% part in upper tropical zone (300-1000 m) and rest only 0.2% part lies in subtropical zone (1000-2000 m) (DDC, Rupandehi 2011).

Fig 3.1: Political map showing position of Rupandehi district in Nepal and stydy area.(Source: My Republica-Nagarik Network)

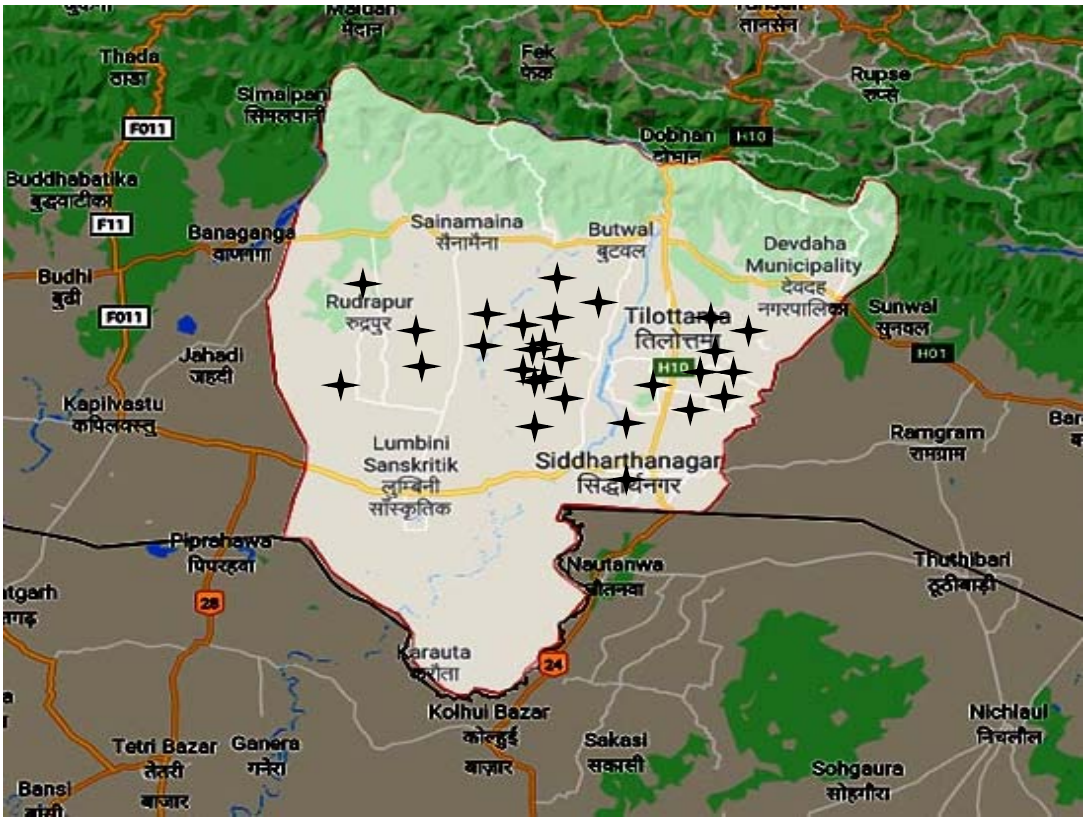


Fig 3.2: Geographical map showing distribution of Rupandehi district shown in light and different major and small towns in and around Rupandehi district. Blood sample collection sites are marked by star. (Source: The Himalayan times)

3.2 Field techniques

3.2.1 Acquisition of consent from parties

At first a prior informed consent was taken by each individuals after explaining and making them fully understand each and every terms, procedures, consequences and their right to refuse in their own Tharu language. The consent paper was written in Nepali. No individuals were forced to provide their blood samples for investigation. Children below sixteen years were included only after the permission from their parents. Permissions from all individuals were taken in written form. After this the next permission was taken from District Public Health Office, Rupandehi and respective municipalities.

3.3.2 Research design

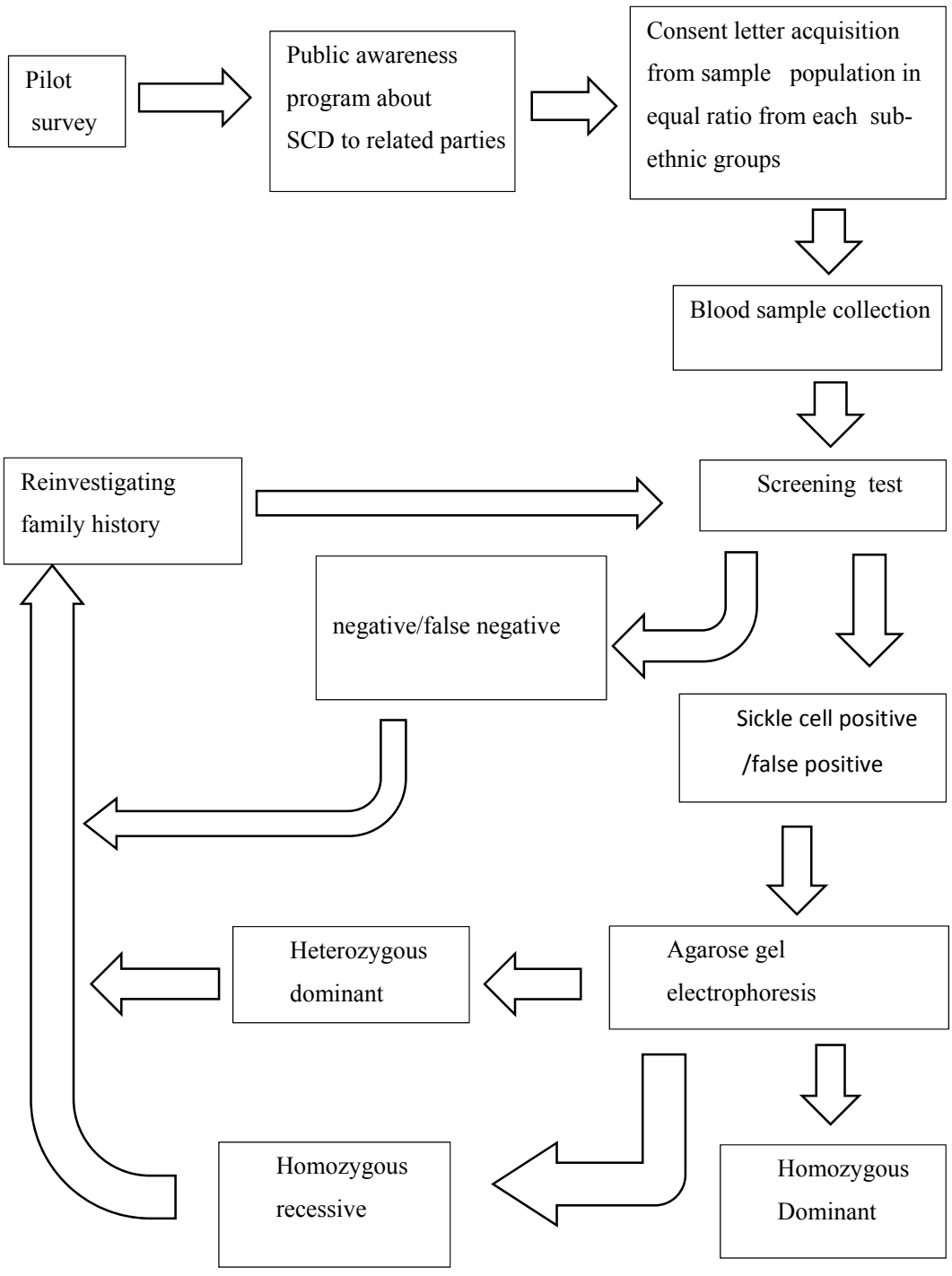


Fig 3.3 : Flow chart showing different steps in research design

3.3 Prevalence of sickle haemoglobin(HbS) in Tharu community

Three different methods, i. Sickle solubility test, ii. Sickling test and iii. RFLP test were used to detect sickle haemoglobin in the blood samples to find prevalence of sickle cell disease in different sub-ethnic groups of the Tharu community of Rupandehi district. The RFLP method was used to find genotype and gene frequencies. And test preference was compared by setting five criteria for comparison between sickle solubility test and sickling test. The criteria were time taken per test, total cost per test, feasibility in field-based detection, level of skill in reagent preparation and test procedure and false positive/negative outcomes.

Sickle haemoglobin in the blood samples were detected by three different techniques where sickle solubility test and sickling test were applied to all samples collected while RFLP method was applied to the samples diagnosed positive from either of above two methods.

3.3.1 Sickle cell solubility test

Sickle solubility test is based on the principle that when blood is mixed with the reducing agent, the HbS forms liquid crystals and give a cloudy appearance to the phosphate buffer solution. At first, the blood sample were collected from the participants with the help of health workers from local government health post. A tourniquet was tightened in the right forearm above the elbow to prevent venous return. The veins in antecubital fossa are cephalic vein, basilica vein and median cubital vein. Usually median cubital vein was selected for blood collection. The area was cleaned with sprit swab. A thin needle of 21 gauge was inserted in the vein of antecubital fossa. Total three ml of blood was taken. The tourniquet was removed immediately to regain venous flow. While withdrawing the prick point was gently pressed with cotton swab to prevent blood oozing. The collected blood sample was put into a vial containing 2 mg/ml of K_2 EDTA as anticoagulant. The vial was labelled with the serial number and name of the donor. A photograph was taken from each donor and name, sex, age, ethnic group, address, parents name, marriage state and name of children were noted.

The reagents were prepared by adopting the following protocol (Nalbandian et al. 1971):

Stock solution

- K_2HPO4, anhydrous 216g.
- KH_2PO_4, crystals 169g.
- Saponin 10g
- Makeup with deionized water to make 1 liter (PH 6.4-6.8)
- Store reagent at 4°C for 1 month

Working solution

- Prepare sufficient quantity for the day by adding 5 mg of sodium dithionite to 1 ml of stock solution.

Procedure

1. Allow reagents and specimens to warm at room temperature prior to performing this test.
2. Pipette 2 ml of working solution into a labelled 12 X 75 mm test tube.
3. Add 0.02 ml of whole blood to the appropriately labelled test tube.
4. Mix the content thoroughly.
5. Incubate the tubes for five minutes in a test tube holder at room temperature. In the paper board test tube holder, reading card must have 14-point or 18-point black type in straight lines on a white background, approximately 0.5 cm apart. Tubes should be held 2.5 cm from the reading card.
6. Read for the turbidity.
7. When blood is mixed with the reducing agent, the HbS forms liquid crystals and give a cloudy appearance to the phosphate buffer solution.
8. A positive result is indicated by a turbid suspension through which the ruled lines are not visible.
9. A negative result is indicated by a transparent suspension through which the ruled lines are visible.

Limitations

- Does not distinguish the difference between SS and AS.
- To confirm the presence of HbS and differentiate between the two states, a gel electrophoresis should be performed.
- Other abnormal haemoglobin variants (HbC) also cause sickling.

Solubility test errors: Technical sources of error include:

- Inactive or outdated reagent
- Reagents below room temperature
- Improper mixing of reagents with specimen
- Improper interpretation of result

Physiologic sources of errors include:

- False positive: erythrocytosis, hyperglobulinemia, extreme leukocytosis, or hyperlipidemia.
- False negative: anemic individual with Hb of < 7.0 g/Dl. Use of packed erythrocyte 0.01 ml for solubility test will correct for this error.
- False negative may occur due to low concentration of HbS such as an infant younger than six months or SS individuals with history of recent transfusion.

3.3.2 Sickling test

Carrying out the sickling test is part of the diagnostic work up in patients suspected of having a sickle cell syndrome. Standard laboratory protocols were adopted from John Old et al. (2012).

Principle: Sodium metabisulphite reduces the oxygen tension inducing the typical sickle-shape of red blood cells.

Sample: Fresh blood in any anticoagulant.

Reagents: 0.2 g of sodium metabisulphite in 10 ml of distilled water. Stir until dissolved. Prepare fresh each time.

Protocol :

1. Mix 1 drop of blood with 1 drop of 2% sodium metabisulphite solution on a microscope slide.
2. Cover with a cover slip and seal the edge with wax/Vaseline mixture or with nail varnish. Allow to stand at room temperature for 1 to 4 hours.
3. Examine under a microscope with the dry objective.

Interpretation: In positive samples the typical sickle-shaped red blood cells will appear. Occasionally the preparation may need to stand for up to 24 hours. In this case put the slides in a moist Petri dish. False negative results may be obtained if the metabisulphite has deteriorated or if the cover slip is not sealed properly. A positive test does not distinguish the sickle cell trait from sickle cell disease. It is important to examine the preparation carefully and in particular near the edge of cover slip.

3.3.3 PCR-RFLP test

PCR-RFLP test was confirmatory test and was used for genotyping of the samples. All the samples collected were not prescribed for this test but those which were positive from either solubility test or sickling test or both were only prescribed for molecular test. The molecular test was carried out in following steps:

3.3.3.1 Blood DNA extraction and purification

The genomic DNA was extracted by spin column method using Quick DNA Miniprep Plus Kit 10 preps. (Zymo Research Corporation, USA) following manufacturers manual. The following protocol was adopted:

1. Reagent preparation**:** add 260 µl Proteinase K Storage Buffer to each Proteinase K (5 mg) tube prior to use. The final concentration of Proteinase K is ~ 20 mg/ml. store at -20°C after mixing.
2. Add up to 200 µl sample to a microcentrifuge tube and add:
   200 µl Biofluid & Cell Buffer(red)
   20 µl Proteinase K.

3. Mix thoroughly or vortex 10 – 15 seconds and then incubate the tube at 55 °C for 10 minutes.
4. Add 1 volume Genomic Binding Buffer to the digested sample. Mix thoroughly or vortex 10-15 seconds.
5. Transfer the mixture to a Zymo-Spin IIC-XLR Column in a Collection tube. Centrifuge at ≥ 12,000 X g for 1 minute. Discard the collection tube with the flow through.
6. Add 400 µl DNA Pre-Wash Buffer to the spin column in a new Collection Tube. Centrifuge at ≥ 12,000 X g for 1 minute. Empty the collection tube.
7. Add 700 µl g-DNA Wash Buffer to the spin column. Centrifuge at ≥ 12,000 X g for 1 minute. Empty the collection tube.
8. Add 200 µl g-DNA Wash Buffer to the spin column. Centrifuge at ≥ 12,000 x g for 1 minute. Discard the collection tube with the flow through.
9. Transfer the spin column to a clean microcentrifuge tube. Add ≥ 50 µl DNA Elution buffer or water directly on the matrix. The eluted DNA can be used immediately for molecular based applications or stored ≤ -20 °C for future use.

3.3.3.2 DNA quantification

The extracted DNA was quantified by using Biospec-nano Micro-volume UV-Spectrophotometer (Shimadzu Corporation, Tokyo, Japan) following manufacturers manual.

3.3.3.3 DNA amplification

The primers were adopted from Adhikari (2017) that target amplification of 6th codon of β-globin gene. Thus designed primer amplifies the 539bp region including 5' UTR, 1st exon, 1st intron and 2nd exon of β-globin gene of 11th chromosome.

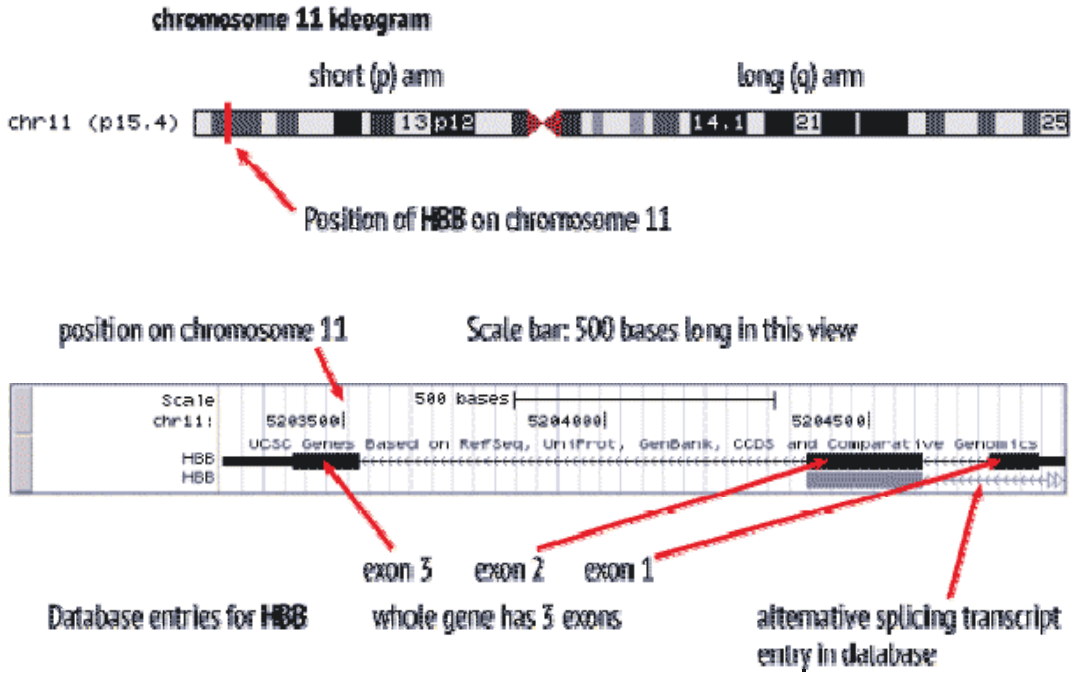


Fig.3.4 : Genetic map showing position and structure of HBB gene in human genome. Upper figure shows location of HBB gene in short arm of chromosome 11. Lower figure shows composition of HBB gene including three exons and two introns.
Source: http://johnhawks.net/explainer/laboratory/genetic-map-haemoglobin

The primer set used to amplify the desired region is shown in following table:

Table 3.1 : Table showing nucleotide sequence of forward and reverse primer for amplification of 539bp of HBB gene in PCR. The following PCR components were used for amplification:

Gene	**Primer**	**Nucleotide sequence**
Human β-globin(HBB)	Forward	5'-AGTCAGGGCAGAGCCATCTA-3'
	Reverse	5'-AGGGTCCCATAGACTCACCC-3'

Table 3.2 : Table showing components of PCR mixture with quantity and concentration for amplification of HBB.

S.N.	Components	Stock con.	Final con.	Dilution factor	Single Reaction (25ul)	For 5+1 reactions	For 10+1 reactions
1	H_2O	-	-	-	13.75 ul	76.5 ul	140.25 ul
2	PCR buffer	10X	1X	10	2.5 ul	15 ul	27.5 ul
3	dNTP mix	2 mM	0.2 mM	10	2.5ul	15 ul	27.5 ul
4	Forward primer	10 uM	250 nM	40	0.625 ul	3.75 ul	6.875 ul
5	Reverse primer	10 uM	250 nM	40	0.625 ul	3.75 ul	6.875 ul
6	$MgCl_2$	25 mM	1.5 mM	16.66	1.5 ul	9 ul	16.5 ul
7	Taq polymerase	5 U/ul	2.5 U/ul	-	0.5 ul	3 ul	5.5 ul
8	Template DNA	100 ng/u	250-1000 Ng	-	3 ul	-	-

The Taq polymerase, dNTPs, primers and template DNA were put in ice to prevent degradation. A master mix was prepared by mixing components 1 to 7 from above table. A thin walled PCR tube was placed in ice and 22 µl of master mix was transferred in it. Then 3 µl of template DNA was transferred and the tube was labelled with distinct sample code. A total 25 µl of reaction mixture was then put in the thermal cycler and the sample DNA was amplified under following condition:

Table 3.3: Table showing optimized conditions during PCR for amplification of HBB.

Step	Temperature °C	Time	Number of cycle
Initial denaturation	95	4 min	1
Denaturation	95	30 sec	38
Annealing	55	30 sec	
Extension	72	1 min	
Final extension	72	7 min	1
Hold	4	-	-

3.3.3.4 Agarose gel electrophoresis of PCR product

Thus PCR amplified product was run in 1.2% agarose gel. The whole procedure was carried out as following steps:

1. Measure 1.2 g of agarose.
2. Mix agarose powder with 100 ml 1X TAE buffer in a 500 ml Erlenmeyer flask.
3. Microwave for 1-3 min until the agarose is completely dissolved. At every 30 s intervals, remove the flask and swirl the contents to mix well.
4. Let agarose solution cool down to 50°C.
5. Pour the agarose into a gel tray with a well comb in place.
6. Place newly poured gel at 4°C for 10-15 min or let sit at room temperature for 20-30 min, until it has completely solidified.
7. Add 1/5 volume loading dye 6X to each of DNA sample.
8. Once solidified, place the agarose gel into the gel box.
9. Fill gel box with 1X TAE until the gel is covered.
10. Carefully load the 20 µl 100bp molecular weight ladder into the first lane of the gel.
11. Carefully load 20µl of samples into the additional wells of the gel.

12. Run the gel at 100V until the dye line is approximately 75% of the way down the gel. A typical run time is 1-1.5 hr. the sample well must be along the negative electrode so that DNA runs towards the positive electrode since DNA is negatively charged.
13. Turn off power, disconnect the electrodes from the power source and then carefully remove the gel from gel box.
14. Place the gel into a container filled with 100 ml of TAE running buffer and 5 µl of EtBr, place on a rocker for 20-3- min, replace EtBr solution with water and distain for 5 min.
15. Using any device that has UV light, visualize the DNA fragments. Take photographs.
16. Using the DNA ladder in the first lane as a guide, infer the size of the DNA in sample lanes.

3.3.3.5 Restriction digestion of amplified DNA template

Restriction endonuclease DdeI (Vivantech, Malasiya) has four restriction sites (C ↓T N A G) on amplified product that yields six fragments of 31bp/37bp, 88bp/89bp/93bp and 201bp in wild genotype. On the contrary RFLP product of homozygous mutant will give five fragments of 31bp/37bp, 88bp/89bp, and 294bp whereas heterozygous mutant results in seven fragments of 31bp/37bp, 88bp/89bp/93bp, 201bp and 294bp upon restriction digestion.

Table 3.4: Table showing quantities of different components used foe restriction digestion of amplified DNA template using DdeI restriction endonuclease.

S.N.	Components	Reaction volume
1	Nuclease free water	18µl
2	PCR product	10µl
3	10X buffer	2µl (1X)
4	DdeI (10U/µl)	2µl
	Total volume	32µl

The components given in above table were transferred in a microcentrifuge tube in the same order as given in the table then mixed properly followed by short spin for few seconds. The reaction mixture was incubated at 37°C overnight. The enzyme was then inactivated by incubating the reaction mixture at 65°C for 20 minutes.

3.3.3.6 Agarose gel electrophoresis of digested product

The digested product was fed into an agarose gel 1% electrophoresis under following protocol:

1. Measure 1 g of agarose.
2. Mix agarose powder with 100 ml 1X TAE buffer in a 500 ml Erlenmeyer flask.
3. Microwave for 1-3 min until the agarose is completely dissolved. At 30 s intervals, remove the flask and swirl the contents to mix well.
4. Let agarose solution cool down to 50°C.
5. Pour the agarose into a gel tray with a well comb in place.
6. Place newly poured gel at 4°C for 10-15 min or let sit at room temperature for 20-30 min, until it has completely solidified.
7. Add 1/5 volume loading dye 6X to each of DNA sample.
8. Once solidified, place the agarose gel into the gel box.
9. Fill gel box with 1X TAE until the gel is covered.
10. Carefully load the 20 µl 100bp molecular weight ladder into the first lane of the gel.
11. Carefully load 20µl of samples into the additional wells of the gel.
12. Run the gel at 100V until the dye line is approximately 75% of the way down the gel. A typical run time is 1-1.5 hr. the sample well must be along the negative electrode so that DNA runs towards the positive electrode since DNA is negatively charged.
13. Turn off power, disconnect the electrodes from the power source and then carefully remove the gel from gel box.
14. Place the gel into a container filled with 100 ml of TAE running buffer and 5 µl of EtBr, place on a rocker for 20-3- min, replace EtBr solution with water and distain for 5 min.

15. Using any device that has UV light, visualize the DNA fragments. Take photographs.
16. Using the DNA ladder in the first lane as a guide, infer the size of the DNA in sample lanes.

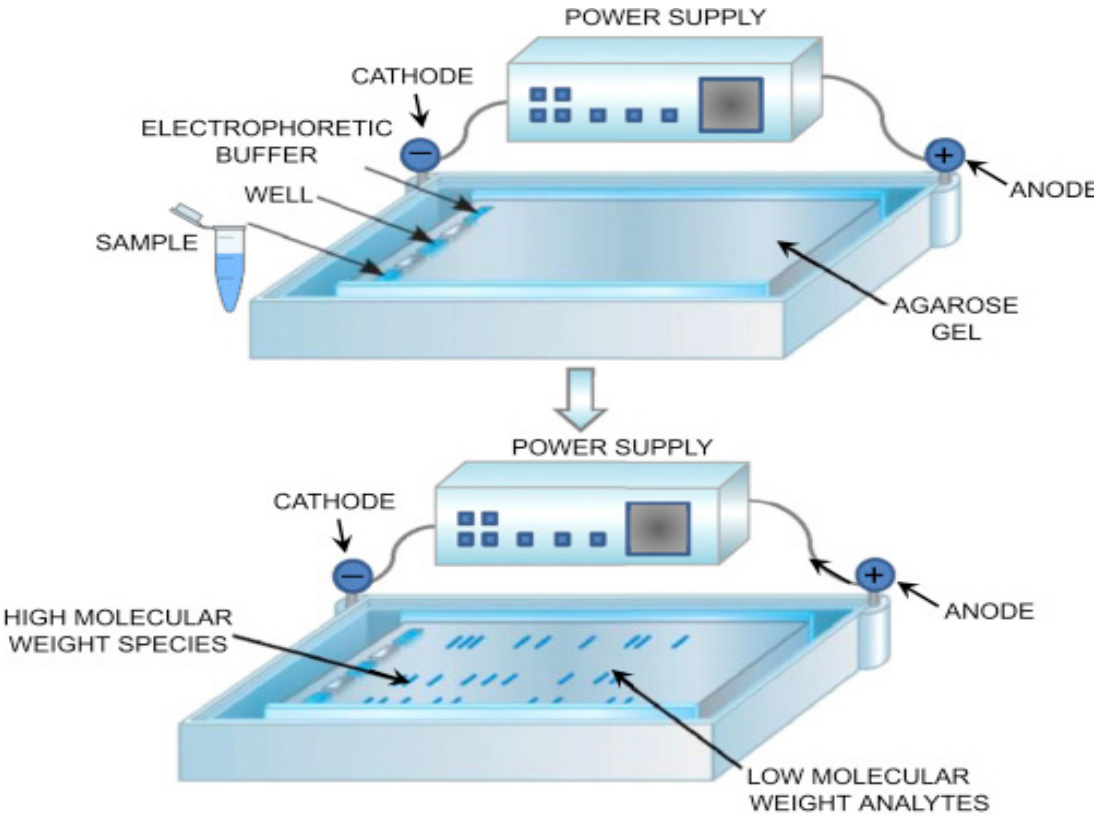


Fig. 3.5: Diagrammatic representation of agarose gel electrophoresis showing loading of sample along cathode and running of DNA fragments along gel in electrophoretic tank. (Source: https://laboratoryinfo.com/agarose-gel-electrophoresis/)

3.4 Gene and genotype frequency of SCA in Tharu community

The genotypes of the positive samples from either of sickling and solubility test were found out by PCR-RFLP method. The genotypes of screening test negative samples were considered to be homozygous dominant. After detecting of genotypes, the genotypic and allelic frequencies were calculated manually by using the relation (Templeton 2006):

1. Genotype frequency = no. of individuals with typical genotype in a population / total no. of population
2. Allele frequency of p = G_{AA} + ½ G_{Aa}
3. Allele frequency of q = G_{aa} + ½ G_{Aa}

Where p is the dominant allele and q is the recessive allele of a gene of a diploid organism. G_{AA} is homozygous dominant, G_{Aa} is heterozygous and G_{aa} is homozygous recessive genotypes.

3.5 Test preference between sickle cell solubility test and sickling test

Sickle cell solubility test and sickling test were used as preliminary test methods for detection of presence of sickle haemoglobin in the blood sample. And test preference was compared by setting four criteria for comparison between sickle solubility test and sickling test. The criteria were time taken per test, cost per test, feasibility in field-based detection, and false positive/negative outcomes or reliability. The average cost per test was calculated by diving total cost of materials including reagents to the total no. of samples. The average time was calculated by measuring time taken per test. The counting time was started after mixing the blood in sodium metabisulphite solution in case of sickling test and after mixing blood sample in concentrated phosphate working solution in solubility test. The time was being monitored until maximum output in the result was seen. Total 20 samples were selected randomly to calculate average time of test. Time consumed during reagent preparation was excluded. The no. of false positive or negative were found out by using RFLP test as confirmatory method. The method which is cost and time efficient, has field-based feasibility and has less number of false positive/negative results was selected as preferred method for HbS detection.

4. RESULTS

4.1 Prevalence of sickle haemoglobin (HbS) in Tharu community

Out of 386 samples, 67 showed positive result for sickling test, 68 for solubility test and 67 for RFLP test. 66 showed positive result for both sickling and solubility test, one positive for both sickling and RFLP test, two were positive for solubility test only and 317 negative for all three tests. Sickling test showed no false positive result while solubility test showed 2 false positive results.

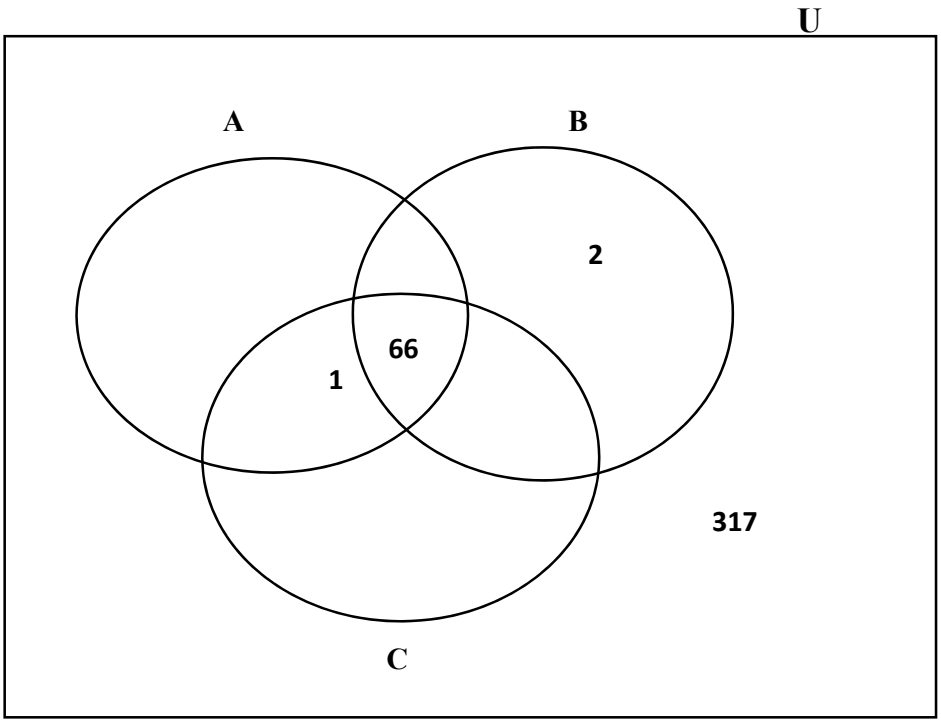


Fig. 4.1: Venn diagram showing comparative result of sickling, solubility and RFLP test where, A= no. of positive samples from sickling test. ; B= no. of positive samples from sickle cell solubility test. ; C= no. Of positive samples for HbS from RFLP test and U= total no. of samples tested.

4.1.1 Sickle cell solubility test

Out of total 386 samples, 68 showed positive result and remaining 318 showed negative result under sickle cell solubility test. The positive samples showed turbidity in concentrated phosphate solution and the black lines behind the test tube were not visible through the solution. The negative samples were completely soluble and made transparent solution where black lines were visible through them.

Table 4.1: sickle solubility test result

Total no. of samples	No. of positive samples	No. of negative samples
386	68	318

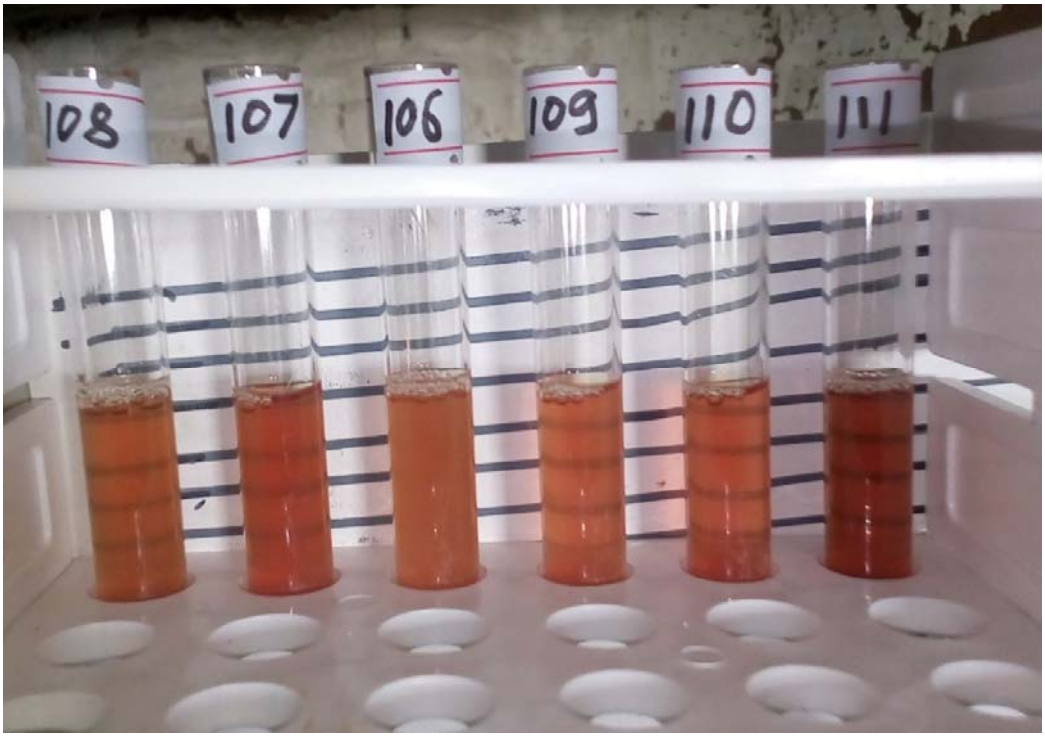


Photo. 4.1: Photo showing the result of sickle cell solubility test for the presence of sickle haemoglobin. sample no. '106' is a positive sample with turbidity in conc. Phosphate buffer where the black lines in a paper behind test-tube are not visible through it. Sample no. 107, 108, 109, 110 and 111 show negative result and they form transparent solution in conc. Phosphate buffer making black lines visible behind them.

4.1.2 Sickling test

All the samples were tested by sickling test method as screening procedure for presence of sickle haemoglobin. Out of total 386 samples, 67 were diagnosed to be positive and remaining 319 to be negative in sickling test. Very distinct crescent shaped erythrocytes were seen under microscope on examination under standard protocol.

Table 4.2: sickling test results

Total no. of samples	No. of positive samples	No. of negative samples
386	67	319

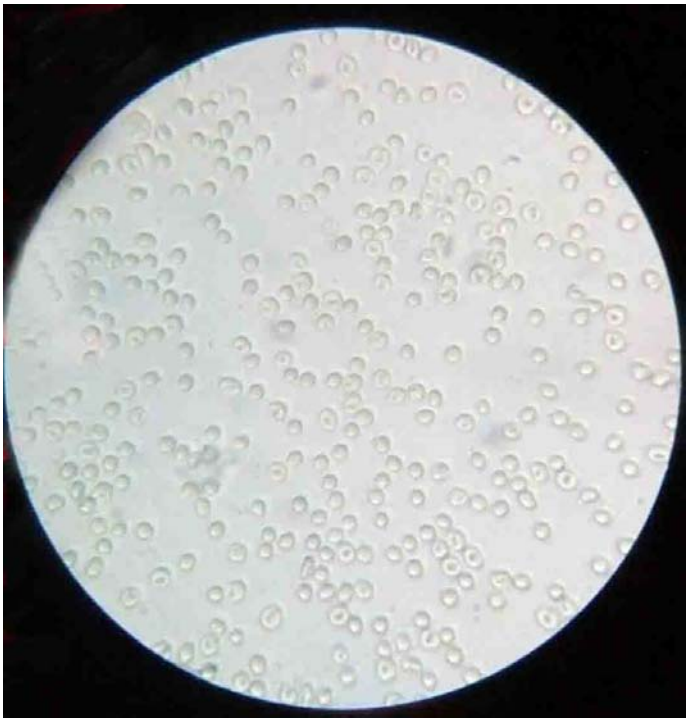


Photo. 4.2: Photograph showing normal blood sample under microscope in 2% sodium metabisulphite solution without stain. The evenly distributed circular findings are erythrocytes. the erythrocytesw show no sickling phenomenon due to absent of HbS in them causing sickling in deoxygeneted condition.

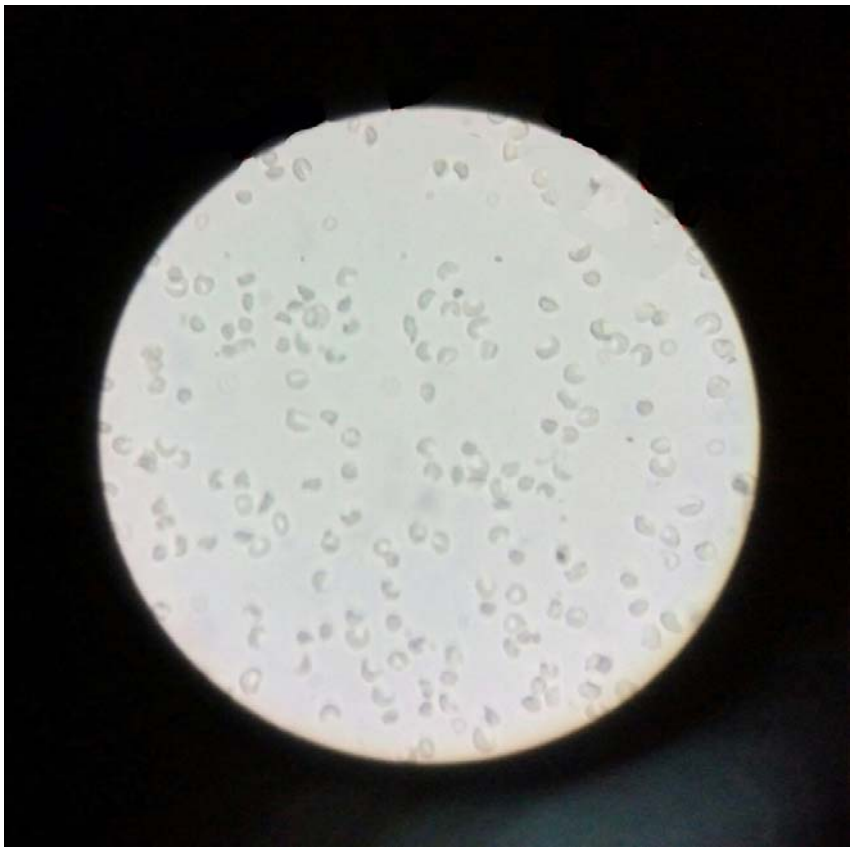


Photo. 4.3: Photographs taken with microscope during sickling test. The Photograph shows a positive sample with typical sickle-shaped erythrocytes. All the samples were treated by 2% sodium metabisulphite solution as reducing agent.

4.1.3 PCR-RFLP test

4.1.3.1 DNA qualification and quantificatioin

The quality and quantity of extracted gDNA was measured by spectrophotometry using BioSpec – nano (Shimadzu Biotech, Japan). The concentration of extracted DNA was ranging from 45 ng/µl to 180 ng/µl; the 260/280 ratio ranged from 1.2 to 1.94 and 260/230 ratio ranged from 0.5 to 1.51.

4.1.3.2 Confirmation of targeted region of β-globin gene

Amplification of β-globin gene that targets 6th codon which causes SCD was performed by polymerase chain reaction from genomic DNA sample. The 539 bp of desired amplicon size were confirmed by 1.2% agarose gel electrophoresis with 100bp DNA ladder.

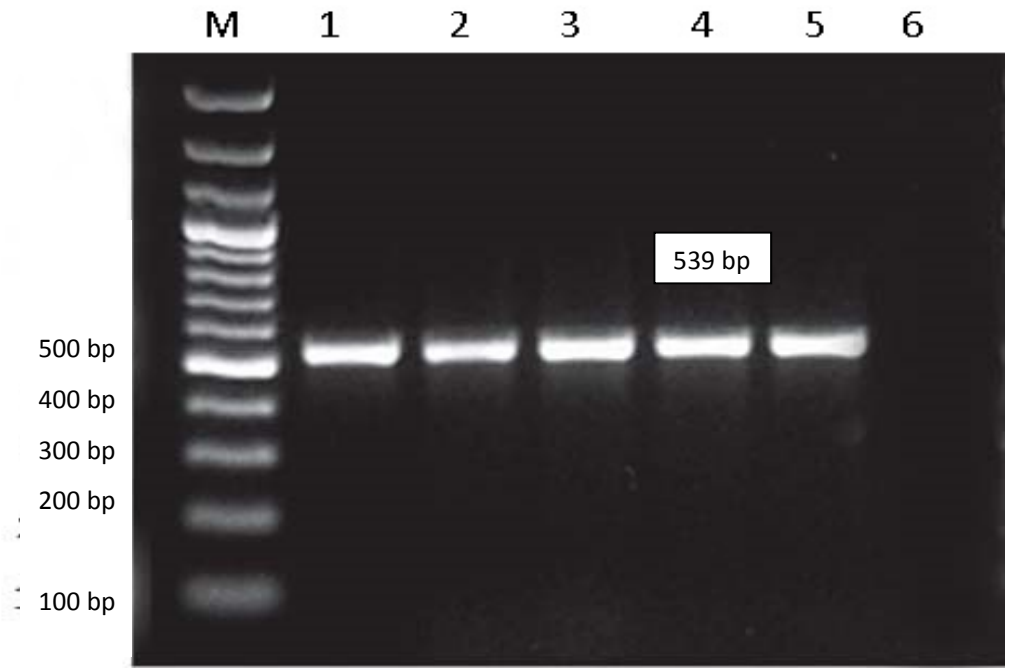


Photo. 4.4: A TAE agarose gel 1.2% electrophoresis image showing 539bp amplicons measured with 100bp molecular weight ladder. Lane M is 100bp DNA ladder; lane 1,2,3,4 and 5 are PCR amplified samples; lane 6 is a negative control (without PCR product containing DNA). The gel was run at 100V in TAE running buffer.

4.1.3.3 Confirmatory test by RFLP

Thus, PCR amplified product was subjected to endonuclease digestion by DdeI restriction enzyme and the digested product was run in 1% agarose gel to separate fragments by their size. The samples were able to be differenciated into three categories, namely, homozygous dominant, heterozygous dominant and homozygous recessive.

The homozygous dominant were normal samples without mutation at 6th codon of beta globin gene of homozygous pair of 11th chromosome. They had six bands upon electrophoresis with five restriction sites. The size of DNA fragments cut by restriction endonuclease were 31bp/37bp, 88bp/89bp/93bp and 201bp. 319 samples were found to be wild.

The heterozygous mutant samples had a mutant chromozome. It gave seven bands upon electrophoresis. The first restriction site is not recognized by restriction endonuclease due to mutation at 6th codon of β-globin gene of 11th chromosome thus an extra 294bp band is seen resulting from combined 93bp and 201bp of a chromosome. The band size were 31bp/37bp, 88bp/89bp/93bp, 201bp and 294bp. 57 samples were hetero with sickle cell trait.

The homozygous mutant samples had mutation on both the chromosomes. Hence the restriction sites at 6th codon of β-globin genes were absent and remained uncut by restriction enzyme resulting in only five bands by DdeI enzyme. The band size were 31bp/37bp, 88bp/89bp and 201bp. 10 samples were homozygous mutant with sickle cell anaemia.

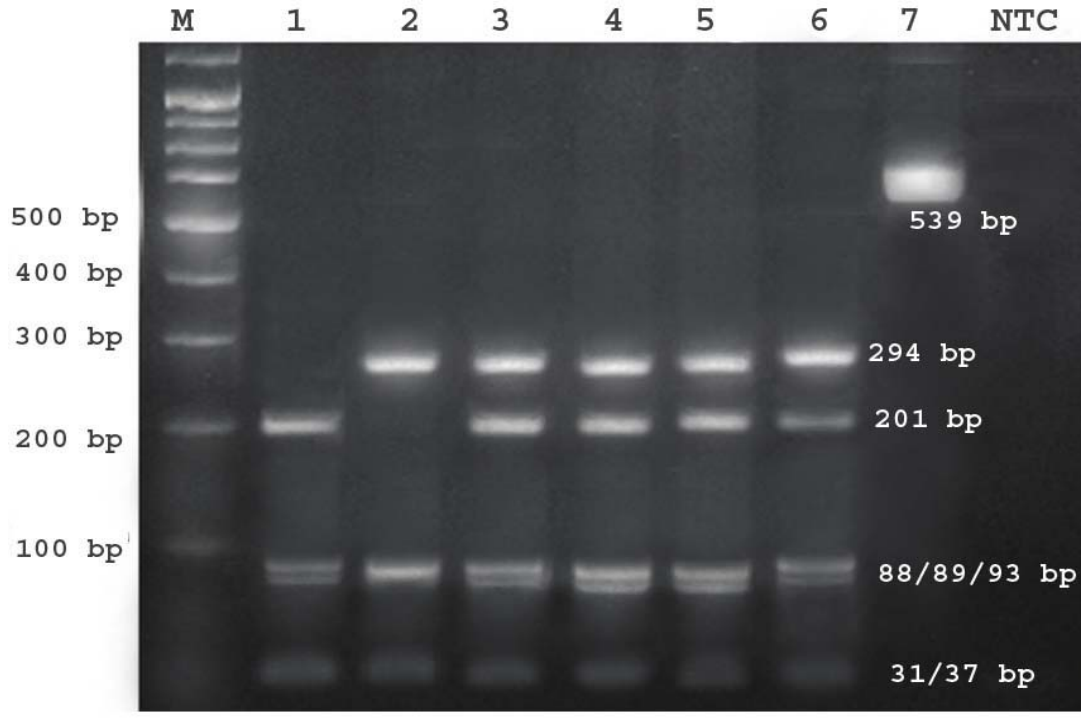


Photo. 4.5: A photograph of TAE agarose gel 1% electrophoresis showing different genotypes with different bandings cut by DdeI restriction endonuclease. Lane M is the 100 bp DNA ladder, lane 1 is a negative control(AA), lane 2 is a positive control(SS), lane 3,4,5 and 6 are trait(AS), lane 7 is restriction negative control with uncut template DNA and lane NTC is a no template control.

4.1.3.4 Ethnicity wise distribution of HbS

Total eight ethnic groups of Tharu had participated in this research. Baatar Tharu had 6.09%, Kathariya Tharu 18.4%, Dadaha Tharu 46.87%, Kanwar Tharu 53.94%, Purbiya Tharu 12%, Dangoriya Tharu 6.94%, Khausiya Tharu 9.52% and Marchaha Tharu had 0% incidence of HbS positive case. Out of them, Kanwar Tharu was at top position by bearing 53.84% of population with HbS while Marchaha Tharu was at lowest position with no any member with HbS at all.

Table 4.3: table showing ethnicity wise distribution of sickle haemoglobin among different ethnic sub-groups of Tharu community of Rupandehi district, Nepal.

Cast	**Total no. of samples**	**No. of positive samples**	**Prevalence(%)**
Kanwar	26	14	53.84
Dadaha	32	15	46.87
Kathariya	125	23	18.40
Purbiya	25	3	12.00
Khausiya	21	2	9.52
Dangoriya	72	5	6.94
Baatar	82	5	6.09
Marchaha	3	0	0
Total	386	67	

4.2 Gene and genotype frequency of SCD in Tharu community

Out of 386 samples, 319 were found homozygous dominant, 57 heterozygous and 10 were homozygous recessive for sickle cell disease. Thus, genotype frequency of homozygous dominant (HbAA) was 82.64%, heterozygous (HbAS) 14.77% and homozygous recessive (HbSS) was 2.59%.

The allele frequency of normal adult beta haemoglobin gen, HbA was 0.9 and allele frequency of mutated beta haemoglobin gene, HbS was 0.1 and the distribution of alleles were found to support the Hardy-Weinberg equilibrium theorem.

4.3 Test preference between sickle cell solubility test and sickling test.

Table 4.4: Comparison of preference between sickling and solubility test.

Diagnostic Method	Cost/test (NRs.)	Time/test	False positive result	False Negative Result	Field-based Feasibility
Sickling test	46.87	34 min.	0	0	Yes
Solubility test	113.33	8 min	2	1	Yes

The degree of preference of selection of diagnostic method for screening SCD was compared between two test methods used in this research work. The sickling test and haemoglobin solubility test methods were compared on the basis of four major criteria. The cost per test, time taken per test, accuracy of test and field-based feasibility were the selected criteria. The cost for sickling test was found to be NRs. 46.87 while that of solubility test was NRs. 113.33. The solubility test took 8 min. and sickling test took 34 min. to show the result. The sickling test was found to be 100% reliable method giving no false positive/negative result but solubility test gave two false positive and one false negative result. The technical sources of errors were neglected in both methods.

From above comparison, sickling test method was found to be preferred due to its low cost, reliability and possibility of field-based interpretation of result. Haemoglobin solubility test method was found to be less reliable than sickling test method due to interpretation of false positive and false negative results but is also a field-based technique.

The molecular method, RFLP was the only confirmatory test that gives no false results, if technical errors are not taken under consideration. But it was the most costly and time taking method requiring highly skilled manpower. Provided that, the general pathological symptoms can confirm the SCD, sickling test is preferred method for mass screening of SCD in the field.

4.4 Hypothesis test

The first research hypothesis was tested against alternative hypothesis and the critical value of test statistics was found to be 686.389. χ^2 at 5% level of significance at 2 degree of freedom is 5.99. since, calculated χ^2 is > tabulated χ^2 0.05,2. Null hypothesis is rejected. i.e. distribution of sickle cell allele is not found to be common at 95% level of confidence. There is a significant difference between observed and expected frequencies, $\chi^2(2, N = 386) = 686.389, p < .00001)$. Sickle cell allele is not common in Tharu people of Rupandehi district.

Likewise the second hypothesis was tested against alternative hypothesis and the critical of test statistics was found to be 58.3823. χ^2 at 5% level of significance at 7 d.f. is 14.07. Since, value of calculated chi-square is greater than that of tabulated value, we failed to accept null hypothesis. i.e. the distribution of sickle cell allele is cast dependent at 5% level of significance at 7 d.f. There is a significant difference between observed and expected frequencies, $\chi^2(7, N = 386) = 58.3823, p < 0)$. The distribution of sickle cell allele is cast dependent.

5. DISCUSSION

5.1 Prevalence of sickle cell disease in Tharu community

This study was carried out in indigenous Tharu tribes of Rupandehi district of Mid-Western Terai of Nepal. A total of 386 people participated from different sub-ethnic groups of Tharu tribe. All the collected samples were undertaken through sickling test and sickle cell solubility test. There we found 67 positive from sickling test and 68 from solubility test. Out of total 69 positive samples two were positive only for solubility test and one was for only sickling test and remaining 66 were positive for both tests. When those 69 samples were undertaken through RFLP test, 57 found to be of HbAS genotype, 10 of HBSS and two of HbAA genotype. Two samples showed false positive and one false negative from solubility test while sickling test showed all true results. The prevalence of sickle haemoglobin was found to be 17.35%. Out of them Kanwar Tharu were found to have highest prevalence(53.88%) and Marchaha Tharu have lowest prevalence(0%). Dadaha Tharu had 46.87%, Kathariya 18%, Purbiya 12%, Khausiya 9.52%, Dangoriya 6.94% and Baatar 6.09% prevalence of sickle haemoglobin.

A retrospective study was performed Shrestha et al. (2020) at five different sites of Nepal Government that uses Capillary electrophoresis for screening of hemoglobin disorders and they were National Public Health Laboratory, Kathmandu, Bharatpur Hospital, Chitwan, Lumbini Zonal Hospital, Butwal, Bheri Zonal Hospital, Nepalgunj and Seti Zonal Hospital, Dhangadhi. Data were collected from laboratory records of those sites from the beginning of hemoglobinopathy testing (July 2016) till December 2018. The samples were analyzed using SEBIA Minicap Flex piercing that uses the principle of capillary electrophoresis to separate different hemoglobin variants. Among the 4018 cases recorded from all the five sickle cell and thalassemia testing sites, 1470 cases were found diagnosed of haemoglobinopathy related disorders. Out of total 1470 cases recorded positive for hemoglobinopathies, address of 325 cases were missing while 1145 cases had valid address. Those 1145 cases were segregated based on their permanent address as mentioned in the laboratory records which shows that the highest number of cases (64%) were from province 5 followed by province 7(10%) that comprises of majority of indigenous Tharu population living in western Terai of Nepal. The least number of cases were recorded from province 4 (2%).

Shrestha and Karki (2013) conducted a study in Tribhuvan University Teaching Hospital, Kathmandu, Nepal from January 2011 to January 2013. Haemoglobin electrophoresis was performed by cellulose acetate electrophoresis at alkaline pH method. Sodium dithionite was used for sickling test. A total of 35 cases were diagnosed as sickle cell disorder. The male: female ratio was 2.5:1with the commonest age group 11-20 years (42.8%). Tharu, Chaudhary and Tharu, Rana (91.3%) were the commonest ethnic group with both sickle cell anemia and trait. It was found that Tharu (Chaudhary; 82.8%) was the most common ethnic group with this disorder followed by Tharu (Rana; 8.5).

Pande et al. (2019) conducted a retrospective study in sickle cell disease between 2012 to 2018 at Bheri Provincial Hospital, Nepalgunj, Nepal by using haemoglobin electrophoresis and HPLC method for diagnosis. A total of 1459 individuals of haemoglobinopathies were seen, out of which 1250 had sickle cell disease and carrier. Out of 1250patients and carriers, 608 (48.6%) were male and 642 (51.4%) were female. The mean age of patients and carriers was 24.5 ±12 yrs. Maximum number of patients 381 (30.5%) were in the age group 21-30 years. Only 156 (12.5%) patients and carriers were under the age of 11 years. Around 1221 (97.7%) of the patients and carriers belonged to the Tharu ethnic group of Nepal. Rest were non-Tharus.

Marchand et al. (2017) screened a total of 2899 Tharu individuals aged 6 months to 40 years in the rural district of Dang in Western Nepal by using a sickling test, of whom, 271 screened positive for HbS. Those who screened positive were offered diagnostic gel electrophoresis testing. Of the 133 individuals who underwent diagnostic testing, 75.9% (n = 101) were confirmed to be Hb AS heterozygotes, 4.5% (n = 6) were confirmed to be Hb SS homozygotes and 19.5% (n = 26) were false positives.

Adhikari (2017) tested 116 individuals of Tharu tribal population of Far-Western Terai of Nepal found 25 to be carrier and none were diagnosed to have sickle cell anaemia. Out of them, Rana Tharu showed 24% prevalence and Dangoriya Tharu showed 20.08% prevalence of sickle cell allele in their gene pool.

5.2 Gene and genotype frequency of SCD in Tharu community

In this study conducted on Tharu tribal population of Rupandehi district, Nepal, 386 individuals were tested to find gene and genotype frequency of sickle cell disease. Out of 386 samples, 319 were found homozygous dominant, 57 heterozygous and 10 were homozygous recessive for sickle cell disease. Thus, genotype frequency of homozygous dominant (HbAA) was 82.64%, heterozygous (HbAS) 14.77% and homozygous recessive (HbSS) was 2.59%. The allele frequency of normal adult beta haemoglobin gen, HbA was 0.9 and allele frequency of mutated beta haemoglobin gene, HbS was 0.1. The distribution of alleles supported the Hardy-Weinberg Theory.

Adhikari (2017) tested 116 individuals to find allelic and genotypic frequency of sickle cell anaemia in Tharu community of Kailali and Kanchanpur district of Nepal . he found genotypic frequency of homozygous HbAA to be 78.44%, heterozygous HbAS 21.55% and mutant HbSS to be 0%. The allele frequency of HbA and HbS were found to be 0.892 and0.107 respectively. These allelic frequencies do not accommodate with Hardy-Weinberg theory. A survey conducted by Ghai et al (2015) in tribal population of Chhattisgarh, India tested 1056 samples for screening of sickle cell anaemia and sickle cell trait. Out of them 78.6% were with HbA/HbA, 20.2% HbA/HbS and 1.3% with HbS/HbS genotype. The allelic frequency for HbA and HbS alleles were 0.89 and 0.11 respectively and is totally compatible with Hardy-Weinberg equilibrium theorem.

A total of 2899 Tharu individuals aged 6 months to 40 years in the rural district of Dang in Western Nepal were screened by Marchand et al. (2017) using sickling test, of whom, 271 (9.3%) screened positive for Hb S. Those who screened positive were offered diagnostic gel electrophoresis testing. Of the 133 individuals who underwent diagnostic testing, 75.9% (n = 101) were confirmed to be Hb AS heterozygotes, 4.5% (n = 6) were confirmed to be Hb SS homozygotes and 19.5% (n = 26) were false positives. These findings support a large burden of sickle cell disease among the Tharu population and highlight the importance of appropriate resource allocation and management

5.3 Test preference between sickle cell solubility test and sickling test

In this study conducted on 386 samples, sickling test showed 67 positive and solubility test showed 68 positive results. Two samples were positive for solubility test only, one sample was positive for sickling test only and 66 were positive both. But upon confirmatory test by molecular method 67 out of 69 positive samples showed positive for sickle cell allele. Hence two samples were found to be false positive and one false negative from solubility test. Sickling test gave no false positive/negative result. On the other hand, sickling test consumed more time than that of solubility test on average but was found to be cheaper and reliable. Both methods were found to be feasible for field-based inference of result. So the sickling test method is preferred over solubility test for the screening of sickle cell allele in the blood.

Adhikari (2017) tested 116 samples from Tharu individuals of Kailali and Kanchanpur district of Nepal and found 26 samples positive from solubility test. But upon molecular test, only 25 found to be positive. The solubility test showed three false positive and two false negative result. Ghai (2017) studied 1056 samples from Chattisgarh, India and found 231 to be positive by sickling test. But upon HPLC, 213 were positive for AS genotype, 13 for SS genotype and three for AA genotype. Two samples were degenerated. It means there were at least three false positive samples from sickling test.

6. CONCLUSION AND RECOMMENDATION

This research is aimed to find prevalence of sickle cell disease in Tharu communities of Rupandehi district of Mid-Western Nepal. Blood samples were collected from different eight sub-ethnic groups of Tharu tribe. Out of total 386 blood samples, 67 were confirmed to be carrying sickle haemoglobin. Thus prevalence of sickle cell allele in Tharu community of Rupandehi district was found to be 17.36%. the prevalence of sickle cell allele was not found to be common in Tharu population of Rupandehi, Nepal, χ^2 (2, N = 386) = 686.389, p < .00001). Out of eight different sub-ethnic groups, the prevalence of sickle cell allele was found to be highest in Kanwar Tharu(53.84%) and lowest prevalence in Marchaha Tharu(0%). Dadaha Tharu have 46.87%, Kathariya Tharu 18.4%, Purbiya Tharu 12%, Khausiya Tharu 9.52%, Dangoriya Tharu 6.94% and Baatar Tharu 6.09% prevalence of

sickle cell allele. The distribution of sickle cell allele in Tharu population is cast-dependent, χ^2 (7, N = 386) = 58.3823, p < 0).

Out of 386 individuals 319 were found to be homozygous dominant(HbAA), 57 heterozygous(HbAS) and remaining 10 recessive(HbSS). The genotypic frequency of wild, hetero and mutant genotype were 82.64%, 14.77% and 2.59% respectively. Likewise the allelic frequency of normal adult haemoglobin HbA and mutated sickle haemoglobin HbS were 0.9 and 0.1 respectively and was fitted in Hardy-Weinberg model.

Sickling test, sickle cell solubility test and PCR-RFLP test were three methods used in this study. Out of them sickling and solubility test are field-based methods. From the comparative analysis it was found that sickling test is more reliable and cost effective method for detection of sickle haemoglobin in the blood sample. Thus sickling test can be used for mass screening of sickle cell disease in the field. PCR-RFLP is a confirmatory molecular test that segregates different genotypes of sickle cell disease. Although RFLP is a time-consuming, costly and difficult method for confirmation of sickle cell allele in the genome.

There is high prevalence of sickle cell anaemia in Tharu belt of Nepal. The invasive studies are advised to screen out all suspected populations and to establish gene banks for sickle cell anaemia for reference. The programs for awareness, treatment and prevention of this disorder are under current necessity. Advanced curative treatment methods such as stem cell transplantation and molecular medicines should be applied to control the burden of SCD in Tharu communities of Nepal and elsewhere in the globe.

The major recommendations from this research are:

- More study is recommended to detect the exact gene frequency of HbS in this community.
- There is a chance of outcome of false negative or false positive results in solubility test and to reduce such false results detailed family history of patient should be understood and other vital tests should also be performed upon the parameters that interfere the result.
- Some mid-costly methods such as HPLC, haemoglobin electrophoresis and capillary electrophoresis should be given priority to avoid false negative/positive results and also to detect other haemoglobinopathies such as HbC and β-thalassemia .

7. REFERENCES

Adamkiewicz, T.V., Mehta, P.S., Boyer, M.W. et al. 2004. Transplantation of unrelated placental blood cells in children with high-risk sickle cell disease. Bone Marrow Transplantation **34**(5): 405–11.

Adhikari, K. 2017. Detection of the prevalence of sickle cell disease in Tharu population of Far-western Nepal by RFLP method. M.Sc. thesis. Central Department of Biotechnology, Tribhuvan University, Kirtipur, Nepal.

Allison, A.C. 2009. genetic control of resistance to human malaria. Current Opinion in Immunology **21**(5):499-505.

Anderson, N. 2006. Hydroxyurea therapy: improving the lives of patients with sickle cell disease. Pediatric Nursing ***32***(6):541- 43.

Apanah, S. and Rizzolo, D. 2013. Sickle cell disease: taking a multidisciplinary approach. Journal of the American Academy of Physician Assistants **26**(8):28-33.

Ashley-koch, A., Yang, Q. and Olney, R.S. 2000. Sickle hemoglobin (HbS) allele and sickle cell disease: a HuGE review. American Journal of Epidemiology ***1*;151**(9):939-45.

Auer, P.L., Johnsen, J.M., Johnson, A.D. et al. 2012. Imputation of exome sequence variants into population- based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project. The American Journal of Human Genetics **91**(5):794–808.

Awasthy, N., Aggarawal, K.C., Goyal, P.C., Prasad, M.S., Saluja, S. and Sharma, M. 2008. Sickle cell disease: experience to the tertiary health care in a nonendemic area. Annals of Tropical Medicine and Public Health **1**(1):1- 4.

Awotua-Efebo, O., Alikor, E.A. and Nkanginieme, K.E. 2004. Malaria parasite density and splenic status by ultrasonography in stable sickle-cell anaemia (HbSS) children. Nigerian Journal of Medicine **13**(1):40–43.

Azar, S. and Wong, T.E. 2017. Sickle cell disease: a brief update. The Medical Clinics of North America **101**(2):375-393.

Balgir, R.S. 2017. Do tribal communities show an inverse relationship between sickle cell disorders and glucose-6-phosphate dehydrogenase deficiency in malaria endemic areas of Central-Eastern India? Hemoglobin **41**(4-6):278-282.

Benson, J.M., Therrell, B.L.Jr. 2010. History and current status of newborn screening for hemoglobinopathies. Seminars in Perinatology **34**(2):134–144.

Bertram, J.S. 2000. The molecular biology of cancer. Molecular Aspects of Medicine **21** (6): 167–223.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M. C., Tassani, S. 2013. An estimation of the number of cells in the human body. Annals of Human Biology **40** (6): 463–71.

Bista, D. B**.** 1971. "Tharu". People of Nepal (Second ed.). Kathmandu: Ratna Pustak Bhandar. pp. 108–118.

Bonham, V.L., Dove,r G.J., Brod,y L.C. 2010. Screening student athletes for sickle cell trait: a social and clinical experiment. New England Journal of Medicine **363**(11):997–99.

Brawley, O.W., Cornelius, L.J., Edwards, L.R. et al. 2008. National Institutes of Health Consensus Development Conference statement: hydroxyurea treatment for sickle cell disease. Annals of Internal Medicine **148**:932–38.

Bruce, S. and O'Brien K. 2006. when bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. The Journal of Experimental Biology **209**(10):1791-1802.

Brydon, H. W., Joshi, G. and Pradhan, S. P. 1961. Anopheles distribution and its relation to the malaria eradication programme in central Nepal. Mosquito News **21** (2): 102–105.

Buchanan-Hamilton, F. 1838. History, Antiquities, Topography and Statistics of Eastern India Martin, M. (Ed.). W. H. Allen & Co.: London **2**: 705–706.

Caetano, C.P., Kraaijenbrink, T., Tuladhar, N.M., Driem, G.L., Knijff, P.,Tyler-Smith, C., and Carvalho-Silva, D.R. 2006. Nepalese populations show no association between the distribution of malaria and protective alleles. Journal of Molecular Genetic Medicine **2**(1): 101–106.

CBS,2014. National Population and Housing Census, 2011. Government of Nepal. National Planning Commission Secretariat. Central Bureau of Statistics, Kathmandu, Nepal. Volume 06.

CDC. Sickle Cell Disease. Centers for Disease Control and Prevention. url: https://www.cdc.gov/ncbddd/sicklecell/articles.html. Assessed on 19 January 2021.

Central Bureau of Statistics. 2009. Chapter 1: Area and Population; Table 1.7: Population Distribution by Caste/Ethnic Groups and Sex for Nepal, 2001. Statistical Year Book of Nepal 2009. Kathmandu: Government of Nepal.

Central Bureau of Statistics. 2012. National Population and Housing Census 2011. Kathmandu: Government of Nepal.

Chaubey, G., Singh, M., Thangaraj, K. et al. 2014. Unravelling the distinct strains of Tharu ancestry. European Journal of Human Genetics **22**:1404–12.

Chiabrando, D., Mercurio, S., Tolosano, E. 2014. Heme and erythropoiesis: more than a structural role. Haematologica **99**(6):973-83.

Colah, R.B., Mukherjee, M.B., Martin, S. and Ghosh, K. 2015. Sickle cell disease in tribal populations in India. Indian Journal of Medical Research **141**(5): 509–515.

Costa, F.F. and Conran, N. 2016. Sickle cell anemia: from basic science to clinical practice. Springer. P. 35.

D'Alessandro, A. 2017. Red blood cell proteomics update: is there more to discover? Blood Transfusion **15** (2): 182–87.

Dahm, R. 2008. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. Human Genetics **122** (6): 565–81.

D'Alessandro, Angelo.2017. Red blood cells proteomics update: is there mort to discover? Blood Transfusion **15** (2): 182-87.

Dominguez de Villota ED, Ruiz Carmona MT, Rubio JJ, de Andrés S. 1981. Equality of the in vivo and in vitro oxygen-binding capacity of hemoglobin in patients with severe respiratory disease. British Journal of Anaesthesia **53** (12): 1325–28.

Elzouki, A.Y. 2012. Textbook of clinical pediatrics. 2nd edition. Berlin: springer.p.2950.

Epstein, F. H. and Hsia, C. C. W. 1998. Respiratory Function of Hemoglobin. New England Journal of Medicine **338** (4): 239–47.

Erazo, B.M., Ramirez, G.A., Gerrato, L.E., Pinto, L.G., Castro, E.J., Yanej, N.J. et al. 2015. Prevalence of HbS (HHB: c.20 A›T) in Honduran population of African descent. Hemoglobin **39**(2):134-37.

Estcourt,L.J., Fortin, P.M., Trivella, and Hopewell, S. 2016. Preoperative blood transfusions for sickle cell disease. Cochrane Database of Systemic Reviews **4**:CD003149.

Ford, J. 2013. Red blood cell morphology. International Journal of Laboratory Hematology. **35**(3): 351–357.

GBD, 2016. Global, regional and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet **388** (10053):1545-1602.

Genetics Home.2016. Sickle cell disease. Genetics Home Reference.

Genome Dictionary. 2007. Somatic cell genetic mutation. Genome Dictionary. Athens, Greece: Information Technology Associates. 30 June 2007.

Ghai, N., Singh, G. and Rais, N. 2015. Genotype and sickle cell allele frequency distribution in tribal population of India. Frontiers of Biological and Life Sciences **3**(1): 25-28.

Gillies, M.T. 1988. Anopheline mosquitos: vector behaviour and bionomics. In: Wernsdorfer,W.H .M.I. (Ed.). Malaria: principles and practice of malariology. Edinburgh: Churchill Livingstone. pp. 453–85.

Gladwin, M.T., Sachdev, V., Jison, M.L., Shizukuda, Y., Plehn, J.F.,Minter, K. et al. 2004. Pulmonary hypertension as a risk factor for death in patients with sickle cell disease. New England Journal of Medicine **350**(9):886-95.

Griffiths, A.J.F. 2020. Mutation. Encyclopaedia Britanica. https://www.britannica.com/science/mutation-genetics. Assessed on 18 January 2021.

Guneratne, A. 2002. Many tongues, one people: The making of Tharu identity in Nepal. Ithaca, New York; Cornell University press.

Hafen, B.B., Sharma, S. 2020. Oxygen Saturation. StatPearls [Internet]. StatPearls Publishing; Treasure Island (FL).

Harewood, J., Azevedo, A.M. 2020. Alpha Thalassemia. StatPearls [Internet]. StatPearls Publishing; Treasure Island (FL).

Hassell, K.L. 2010. Population estimates of sickle cell disease in the U.S. American Journal of Preventive Medicine. **38**(4):512-21.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M. and Ira, G. 2009. Mechanisms of change in gene copy number. Nature Reviews. Genetics **10** (8): 551–64.

Higgins, J. 2015. Red blood cell population dynamics. Clinics in Laboratory Medicine. **35**(1):43–57.

Howard, J. 2016. Sickle cell disease: when and how to transfuse. Hematology. American Society of Hematology. Education Program **2016**(1): 625–631.

Hsia, C.C.W. 1998. Respiratory function of hemoglobin. New England Journal of Medicine **338**(4):239-47.

https://en.wikipedia.org/wiki/Restriction_fragment_length_polymorphism. Assessed on 21 November 2020.

https://international.neb.com/products/restriction-endonucleases/restriction-endonucleases/types-of-restriction-endonucleases. Assessed on 21 November 2020.

Huisjes, R., Bogdanova, A., van Solinge, W., Schiffelers, R., Kaestner, L., and van Wijk, R. 2018. Squeezing for life – properties of red blood cell deformability. Frontiers in Physiology **9**:656.

Information Center for Sickle Cell and Thalassemic Disorder,2002. How does sickle cel cause disease? http.//sickle.bwh.harward.edu/scd_background.html.

Jha, A. 2012. Older fathers pass on more genetic mutations, study shows. The Guardian. Published on 22 August 2012.

Johnson, F.L., Look, A.T., and Gockerman, J. 1984. Bone-marrow transplantation in a patient with sickle-cell anemia. The New England Journal of Medicine **311**(12) :780–83.

Jr, Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; with Gregory J. Gatto (2012). Biochemistry 7th ed. New York: W.H. Freeman. pp. 455- 609.

Kassim, A.A. and DeBaun, M.R.2014. The case for and against initiating either hydroxyurea therapy, blood transfusion therapy or hematopoietic stem cell transplant in asymptomatic children with sickle cell disease. Expert Opinion on Pharmacotherapy **15**(3):325-36.

Kato, G.J., Piel, F.B., Reid, C.D., Gaston, M.H., Ohene-Frempong, K., Krishnamurti, L., Smith, W.R., Panepinto, J.A., Weatherall, D.J., Costa, F.F., Vichinsky, E.P. 2018. Sickle cell disease. Nature Reviews. Disease Primers **4**:18010.

Kondrashin, A.V. 1992. Malaria in the WHO Southeast Asia region. Indian Journal of Malariology **29**(3):129-60.

Kotila, R., Okesola, A. and Makanjuola, O. 2007. Asymptomatic malaria parasitaemia in sickle-cell patients: how effective is chemoprophylaxis? Journal of Vector Borne Diseases **44**(1): 52–55.

Krauskopff, G. 1995. The anthropology of the Tharus: an annoted bibliography. Kailash **17** (3/4): 185–213.

Krishnamurti, L., Abel, S., Maiers, M., and Flesch, S. 2003. Availability of unrelated donors for hematopoietic stem cell transplantation for hemoglobinopathies. Bone Marrow Transplantation **31** (7): 547–550.

Kryndushkin, D.S., Alexandrov, I.M., Ter-Avanesyan, M.D. and Kushnirov, V.V. 2003. Yeast(PSI+) prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. Journal of Biological Chemistry **278**(49):49636-43.

Kumar, V., Abbas, A.K., Fausto, N. and Aster, J. 2009. Robins and Cotran pathologic basis of disease, professional edition. Elsevier Health. Kindle edition.8th edition.

Kwiatkowski, D.P. 2005. How malaria has affected the human genome and what human genetics can teach us about malaria. American Journal of Human Genetics **77**(2):171-92.

Lewis, M. P., Simons, G. F., Fennig, C. D. (Eds). 2014. Tharu, Rana: a language of Nepal. Ethnologue: Languages of the World (Seventeenth ed.). Dallas, Texas: SIL International.

Lewis, M. P., Simons, G. F., Fennig, C. D. (Eds.). 2014. Tharu, Kochila: a language of Nepal. Ethnologue: Languages of the World (Seventeenth ed.). Dallas, Texas: SIL International.

Lewis, M. P., Simons, G. F., Fennig, C. D.(Eds.). 2014. Sonha, a language of Nepal. Ethnologue: Languages of the World (Seventeenth ed.). Dallas, Texas: SIL International.

Lewis, M. P., Simons, G. F., Fennig, C. D., eds. (2014). Tharu, Chitwania: a language of Nepal. Ethnologue: Languages of the World (Seventeenth ed.). Dallas, Texas: SIL International.

Lewis, M.P., Simons, G.F. and Fennig, C.D. 2014. Tharu, Kathoriya: a language of Nepal. Ethnologue: Languages of the World 7th edition.

Lodish, H., Bwerk, A., Matsudaira, P., Kaiser, C.A., Krieger, M., Scott, M.P. et al.2004. Molecular cell biology. WH Freeman: New York, NY. 5th edition. pp 366-418.

Loewe, L. and Hill, W. G. 2010. The population genetics of mutations: good, bad and indifferent. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **365**(1544): 1153–67.

López, L., Saravia, C., Gomez, A., Hoebeke, J., Patarroyo , M.A. 2010. Mechanisms of genetically-based resistance to malaria. Gene **467**(1-2):1-12.

Luzzatto, L. 2012. Genetic Factors in Malaria Resistance. In Boulyjenkov, V., Berg, K., Christen, Y. (Eds.). Genes and Resistance to Disease. Heidelberg: Springer. pp. 105–111.

Makani, J., Ofori-Acquah, S.F., Nnodu, O., Wonkam, A. and Ohene-Frempong, K. 2013. Sickle Cell Disease: New Opportunities and Challenges in Africa. The Scientific World Journal, vol. 2013, Article ID 193252, 16 pages, https://doi.org/10.1155/2013/193252.

Makani, J., Komba, A.N., Cox, S.E. et al. 2010. Malaria in patients with sickle cell anemia: burden, risk factors, and outcome at the outpatient clinic and during hospitalization. Blood **115** (2):215–220.

Marchand, M., Gill, C., Malhotra, A.K., Bell, C., Busto, E., McKeown, M.D. et al.2015. The assessment and sustainable management of sickle cell disease in the indigenous Tharu people of Nepal. Hemoglobin **39**(2):134-37.

Marchand, M., Gill, C., Malhotra, A.K., Bell, C., Busto, E., McKeown, M.D. et al.2017. The assessment and sustainable management of sickle cell disease in the indigenous Tharu population of Nepal. Hemoglobin **41**(4 - 6):278-82.

Marengo-Rowe, A.J. 2006. Structure-function relations of human hemoglobins. Proceedings (Baylor University Medical Center) **19**(3): 239–245.

McGann, P.T. and Ware, R.E. 2011. Hydroxyurea for sickle cell anemia: What have we learned and what questions still remain ? Current Opinion in Hematology **18**(3): 158–165.

McLean, J. 1999. Conservation and the impact of relocation on the Tharus of Chitwan, Nepal**.** Himalayan Research Bulletin **XIX** (2): 38–44.

Meier, E.R. 2018. Treatment options for sickle cell disease. Pediatric Clinic North America **65**(3):427-43.

Modiano, G., Morpurgo, G., Terrenato, L., Novelletto, A., Di Rienzo, A., Colombo, B., Purpura, M. et al. 1991. Protection against malaria morbidity- near fixation of alpha thalassemia gene in a Nepalese population. American Journal of Human Genetics **48**(2): 390-397.

Nalbandian, R.M. et al. 1971. Dithionite tube test: a rapid inexpensive technique for the detection of hemoglobin S and non-S sickling hemoglobin. Clinical Chemistry **17**(10): 1028-1032.

Nation Nepali Times. 2014. Life with sickle. Issue 20-26 June, 2014.

National Library of Medicine. URL= ghr.nlm.nih.gov/condition/sickle-cell-disease. accessed on 10/11/2020.

National Human Genome Research Institute. National Institute of Health. https://www.genome.gov/genetics-glossary/Mutation. Assessed on 18 January 2021.

National Population Census 2016. National Planning Commission Secretariat, Central Bureau of Statistics (CBS), Government of Nepal.

Nature. 2018. Mutation / learn science at citable. Nature Education. Retrieved at 24 September 2018.

NHLBI,2015. National Heart, Lung and Blood Institute, Bethesda, Maryland, United States.

NHLBI.2016. What is sickle cell disease? National Heart, Lung and Blood Institute, Bethesda, Maryland, United States.

NHLBI.2017. learning about sickle cell disease. National Heart, Lung and Blood Institute, Bethesda, Maryland, United States.

NHRC. 2018. Population based screening of sickle cell disorder in Tharu community of Bardia district, Nepal. Government of Nepal. Nepal Health Research Council, Kathmandu, Nepal.

Noronha, S.A., Sadreameli, S.C. and Strouse, J.J. 2016. Management of sickle cell disease in children. Southern Medical Journal **109**(9):495-502.

Office of the Registrar General. 2001. Uttaranchal. Data Highlights: The Scheduled Tribes. Census of India 2001. Retrieved 16 March 2008.

Old, J., Harteveld, C.L., Traeger-Synodinos, J., Petrou, M., Angastiniotis, M. and Galanello, R. 2012. Prevention of Thalassaemias and Other Haemoglobin Disorders, 2nd edition. Volume 2: Laboratory Protocols. Thalassaemia International Federation, Nicosia,Cyprus.

Olujohungbe, A. and Howard, J. 2008. The clinical care of adult patients with sickle cell disease. British Journal of Hospital Medicine(London, England:2005) **69**(11):616-19.

Pande, R., Ghimire, P.G., Chand, P.B. and Gupta, S. 2019. Sickle cell disease in Western Nepal. Nepal Journal of Medical Sciences **4**(1):15-19.

Park, B., Yoo, K.H. and Kim, C. 2015. Hematopoietic stem cell expansion and generation: the ways to make a breakthrough. Blood Research **50** (4): 194–203.

Perutz, M.F. 1978. Hemoglobin structure and respiratory transport. Scientific America **239**(6):92-125.

Perutz, M.F., Fermi, G., Luisi, B., Shanan, B. and Liddington, R.C. 1987. Stereochemistry of cooperative mechanisms in hemoglobin. Accounts of Chemical Research **20** (9): 309–321.

Perutz, M.F., Fermi, G., Luisi, B., Shaanan, B. and Liddington, R.C. 1987. Stereochemistry of cooperative mechanisms in hemoglobin. Cold Spring Harbor Symposia on Quantitative Biology **52**:555-65

Piel FB.2016. The present and future global burden of the inherited disorders of hemoglobin. Hematology/ Oncology Clinics of North America **30**(2):327-41.

Pingoud,A. and Jeltsch,A.2001. Structure and function of type II restriction endonucleases. Nucleic Acids Research. **29**(18):3705–27.

Powars, D.R., Chan, L.S., Hiti, A., Ramicone, E. and Johnson, C. 2005.outcome of sickle cell anemia: a four-decade observational study of 1056 patients. Medicine(Baltimore) **84**(6):363-76.

Rajaure, D. P. 1981. Tharus of Dang: the people and the social context. Kailash **8** (3/4): 155–185.

Raleigh, E.A. and Brooks,J.E.1998. Bacterial Genomes. Chapman & Hall, New York, pp.78–92.

Rees, D.C., Williams, T.N. and Gladwin, M.T. 2010. Sickle cell disease. Lancet **376**(9757): 2018-31.

Rich, S.M., Licht, M.C., Hudson, R.R. and Ayala, F.J. 1998. Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America **95**(8): 4425-30.

Rodgers, K. and McVey, M. 2016. Error-prone repair of DNA double-strand breaks. Journal of Cellular Physiology **231**(1): 15-24.

Sambrook, J. and Russel, D.W.2001. molecular cloning: a laboratory manual 3rd edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

SANCHHIPTA PARICHAYA: RUPANDEHI JILLA SAMANWAY SAMITI. www.ddcrupandehi.gov.np. Assessed on 04 December 2020.

Sawyer, S.A., Parsch, J., Zhang, Z. and Hartl, D.L. 2007. Prevalence of positive selection among nearly neutral amino acid replacements in Drosophila. Proceedings of the National Academy of Sciences of the United States of America **104** (16): 6504–10.

Schong, J.B., Duits, A.J., Muskiet, F.A., ten Cate, H., Rojer, R.A., Brandjes, D.P. et al. 2004. Sickle cell disease; a general overview. The Netherlands Journal of Medicine **62**(10):364-74.

Serjeant, G.R. 2001. The emerging understanding of sickle cell disease. British Journal of Haematology **112**(1):3-18.

Serjeant, G.R.2010. one hundred years of sickle cell disease. British Journal of Haematology **151**(5):425-29.

Shafer FE, Lorey F, Cunningham GC, et al. 1996. Newborn screening for sickle cell disease: 4 years of experience from California's newborn screening program. Journal of Pediatric Hematology/Oncology **18**(1): 36-41.

Sherchand ,J.B., Shrestha, M.P., Shreshta, B.L., Banerjee , M.K., Shakya, S. 1995. A preliminary study on field trials with insecticide-treated mosquito nets for malaria control in a rural endemic community of Nepal. Journal of the Nepal Medical Association **33**:195–203.

Shrestha, A. and Karki, S. 2013. Analysis of sickle hemoglobin. Journal of Pathology of Nepal **3**(6): 437 - 40.

Shrestha, R.M., Pandit, R., Yadav, U.K., Das, R., Yadav, B.K. and Upreti, H.C. 2020. Distribution of hemoglobinopathy in Nepalese population. Journal of Nepal Health Research Council **18**(46): 52-58.

Skar, H. O. 1995. Myths of origin: the Janajati Movement, local traditions, nationalism and identities in Nepal. Contributions to Nepalese Studies **22** (1): 31–42.

Steinberg, M.H. 1999. Management of sickle cell disease. The New England Journal of Medicine **340**(13):1021–30.

Steinberg, M.H., McCarthy, W.F., Castro, O., Ballas, S.K., Armstrong, F.D., Smith, W., Ataga, K., Swerdlow P., Kutlar, A., DeCastro, L. and Waclawiw, M.A. 2010. The Risks and Benefits of Long-term Use of Hydroxyurea in Sickle Cell Anemia: A 17.5 Year Follow-Up. American Journal of Hematology **85**(6): 403–8.

Strose, J. 2016. Sickle cell disease. Handbook of Clinical Neurology **138**:311-24.

Templeton, A.R. 2006. Population genetics and micro evolutionary theory. John Wiley and Sons, Inc., Hoboken, New Jersey. Pp 21-28.

Terrenato, L., Shrestha, S., Dixit, K.A., Luzzatto, L., Modiano, G., Morpurgo, G., Arese, P. 1988. Decreased malaria morbidity in the Tharu people compared to sympatric populations in Nepal. Annals of Tropical Medicine and Parasitology **82**(1):1-11.

Thom, C.S., Dickson, C.F., Gell, D.A. and Neiss, M.J. 2013. Haemoglobin variants: biochemical properties and clinical correlates. Cold Spring Harbor Perspectives in Medicine **3**(3): a011858.

Thomas, A.N., Pattison, C. and Serjeant, G.R. 1982. Causes of death in sickle-cell disease in Jamaica. British Medical Journal **285** (6342):633–35.

Thomas, C. and Lumb, A.B. 2012. Physiology of haemoglobin**.** Continuing Education in Anaesthesia Critical Care & Pain **12**(5):251-56.

Verma, I.C. 2004. Hemoglobinopathies in India - An overview. Kochi: Proc. Indo-French Symposium on Recent Trends in Clinical, Diagnostic and Research Aspects of hemoglobinopathies. pp. 2–4.

Viallat, A., and Abkarian, M. 2014. Red blood cell: from its mechanics to its motion in shear flow. International Journal of Laboratory Hematology **36**(3) :237–243.

Vichinsky, E. 2002. New therapies in sickle cell disease. Lancet **360**(9333):629-31.

Vichinsky, E.P. 2004. Pulmonary hypertension in sickle cell disease. The New England Journal of Medicine **350**(9):857-59.Walters, M.C., Storb, R., Patience, M. et al. 2000. Impact of bone marrow transplantation for symptomatic sickle cell disease: an interim report. Blood **95** (6) :1918–24.

Ware, R.E., Steinberg, M.H. and Kinney, T.R. 1995. Hydroxyurea: an alternative to transfusion therapy for stroke in sickle cell anemia. American Journal of Hematology **50**(2): 140–143.

Wattavidanage, J., Carter, R., Perera, K.L., Munasingha, A., Bandara, S., McGuinness, D., Wickramasinghe, A.R., Alles, H.K., Mendis, K.N., Premawansa, S. 1999. TNFalpha*2 marks high risk of severe disease during Plasmodium falciparum malaria and other infections in Sri Lankans. Clinical and Experimental Immunology **115**(2):350-55.

Weed, R.I., Reed, C.F. and Berg, G. 1963. Is hemoglobin an essential structural component of human erythrocyte membrane? The Journal of Clinical Investigation **42** (4): 581-88.

Wellstein, A., Pitschner, H.F. 1988. Complex dose- response curves of atropine in man explained by different functions of M1-and M2-cholinoreceptors. Naunyn-Schmiedeberg's Archives of Pharmacology **338**(1):19-27.

WHO. 2010. sickle cell anemia- report by Secretariat World Health Organization, Geneva, Switzerland.

Wierenga, K.J., Hambleton, I.R. and Lewis, N.A. 2001. Survival estimates for patients with homozygous sickle cell disease in Jamaica: a clinic-based population study. Lancet **3 57**(9257):680-83.

Zhang, J., Wang, X. and Podlaha, O. 2004. Testing the chromosomal speciation hypothesis for humans and chimpanzees. Genome Research **14** (5): 845–51.

Appendix 1

Amplified β-globin sequence

chr11:5226559-5227097 539bp
AGTCAGGGCAGAGCCATCTA AGGGTCCCATAGACTCACCC

AGTCAGGGCAGAGCCATCTAttgcttacatttgcttctgacacaactgtg
ttcactagcaacctcaaacagacaccatggtgcatctgactcc/tgaggag
aagtctgccgttactgccctgtggggcaaggtgaacgtggatgaagttgg
tggtgaggccctgggcaggttggtatcaaggttacaagacaggtttaagg
agaccaatagaaactgggcatgtggagacagagaagactcttgggtttct
gataggcactgactctctctgcctattggtctattttcccaccc/ttaggc
tgctggtggtctacccttggacccagaggttctttgagtcctttggggat
ctgtccactcctgatgctgttatgggcaaccc/taaggtgaaggctcatgg
caagaaagtgctcggtgcctttagtgatggcctggctcacctggacaacc
tcaagggcacctttgccacac/tgagtgagctgcactgtgacaagctgcac
gtggatcc/tgagaacttcaGGGTGAGTCTATGGGACgCT

APPENDIX 2

Restriction digestion

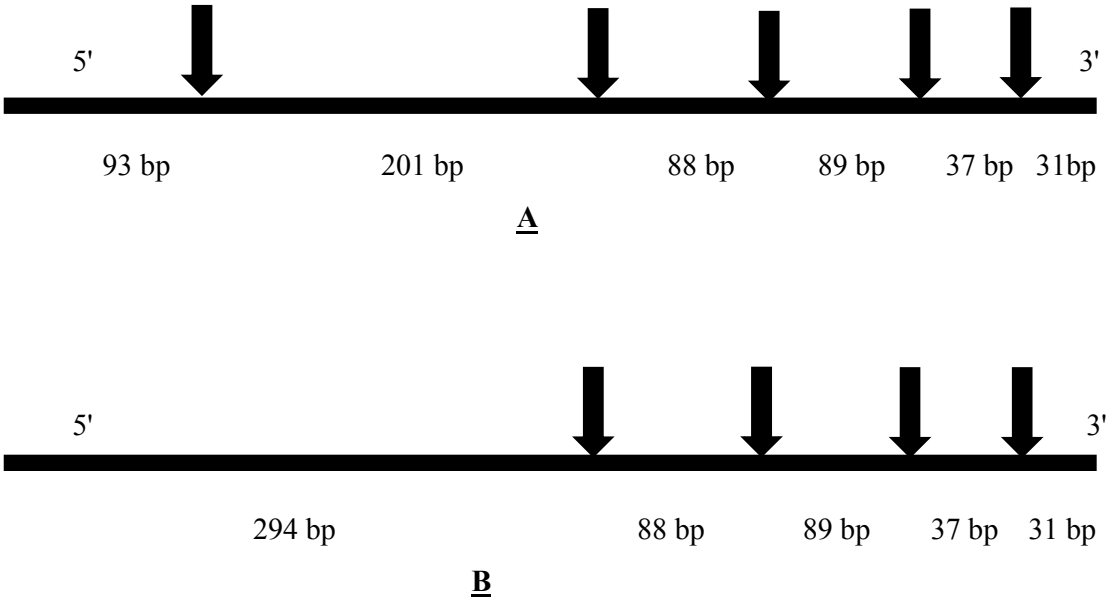


Fig.: An illustration showing recognition site of restriction enzyme in the given template DNA. Figure A is a normal template with five restriction sites producing six fragments and figure B is a mutated template DNA which has lost a restriction site due to point mutation thus producing only five fragments upon digestion by DdeI restriction endonuclease.

APPENDIX 3

Genomic DNA sequence of HBB

NCBI Reference Sequence: NC_000011.10 Homo sapiens chromosome 11, GRCh38.p13 Primary Assembly

chr11: 5,225,464-5,227,071 1608 bp

5' UTR

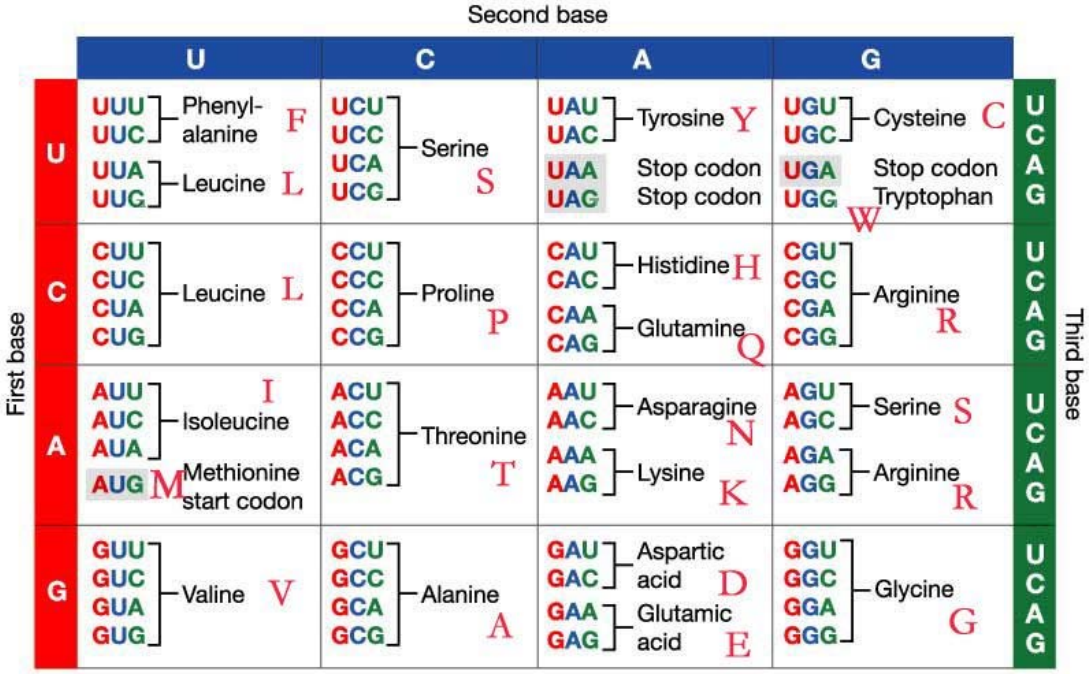
```
AGTCAGGGCA GAGCCATCTA TTGCTTACAT TTGCTTCTGA CACAACTGTG 50
TTCACTAGCA ACCTCAAACA GACACCATGG TGCATCTGAC TCCTGAGGAG
AAGTCTGCCG TTACTGCCCT GTGGGGCAAG GTGAACGTGG ATGAAGTTGG 92
TGGTGAGGCC CTGGGCAGgt tggtatcaag gttacaagac aggtttaagg
agaccaatag aaactgggca tgtggagaca gagaagactc ttgggtttct 130
gataggcact gactctctct gcctattggt ctattttccc acccttagGC
TGCTGGTGGT CTACCCTTGG ACCCAGAGGT TCTTTGAGTC CTTTGGGGAT
CTGTCCACTC CTGATGCTGT TATGGGCAAC CCTAAGGTGA AGGCTCATGG 223
CAAGAAAGTG CTCGGTGCCT TTAGTGATGG CCTGGCTCAC CTGGACAACC
TCAAGGGCAC CTTTGCCACA CTGAGTGAGC TGCACTGTGA CAAGCTGCAC
GTGGATCCTG AGAACTTCAG Ggtgagtctat gggacgctt gatgttttct
ttccccttct tttctatggt taagttcatg tcataggaag gggataagta
acagggtaca gtttagaatg ggaaacagac gaatgattgc atcagtgtgg
aagtctcagg atcgttttag tttcttttat ttgctgttca taacaattgt
tttcttttgt ttaattcttg ctttcttttt ttttcttctc cgcaattttt
actattatac ttaatgcctt aacattgtgt ataacaaaag gaaatatctc
tgagatacat taagtaactt aaaaaaaaac tttacacagt ctgcctagta
cattactatt tggaatatat gtgtgcttat ttgcatattc ataatctccc
tactttattt tcttttattt ttaattgata cataatcatt atacatattt 850
atgggttaaa gtgtaatgtt ttaatatgtg tacacatatt gaccaaatca
gggtaatttt gcatttgtaa ttttaaaaaa tgctttcttc ttttaatata
ctttttgtt tatcttattt ctaatacttt ccctaatctc tttctttcag
ggcaataatg atacaatgta tcatgcctct ttgcaccatt ctaaagaata
acagtgataa tttctgggtt aaggcaatag caatatctct gcatataaat
atttctgcat ataaattgta actgatgtaa gaggtttcat attgctaata
gcagctacaa tccagctacc attctgcttt tattttatgg ttgggataag
gctggattat tctgagtcca agctaggccc ttttgctaat catgttcata
cctcttatct tcctcccaca GCTCCTGGGC AACGTGCTGG TCTGTGTGCT
GGCCCATCAC TTTGGCAAAG AATTCACCCC ACCAGTGCAG GCTGCCTATC 129
AGAAAGTGGT GGCTGGTGTG GCTAATGCCC TGGCCCACAA GTATCACTAA
GCTCGCTTTC TTGCTGTCCA ATTTCTATTA AAGGTTCCTT TGTTCCCTAA
GTCCAACTAC TAAACTGGGG GATATTATGA AGGGCCTTGA GCATCTGGAT 134
TCTGCCTAAT AAAAAACATT TATTTTCATT GCAA
```

3'UTR

Fig.: Genomic DNA sequence of HBB gene of human. CDS in upper case while introns in lower case.5'UTR and 3'UTR are highlighted in yellow, exons in red and introns not highlighted. Amplified region given in bold.(source: Genome Browser, UCSC. http://genome-asia.ucsc.edu/index.html)

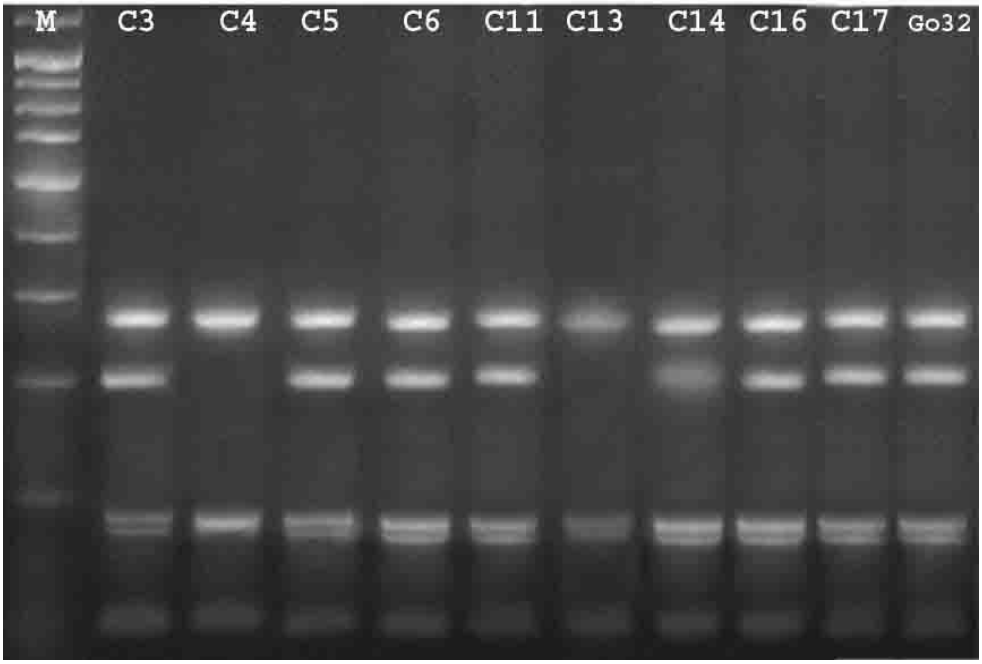
APPENDIX 4

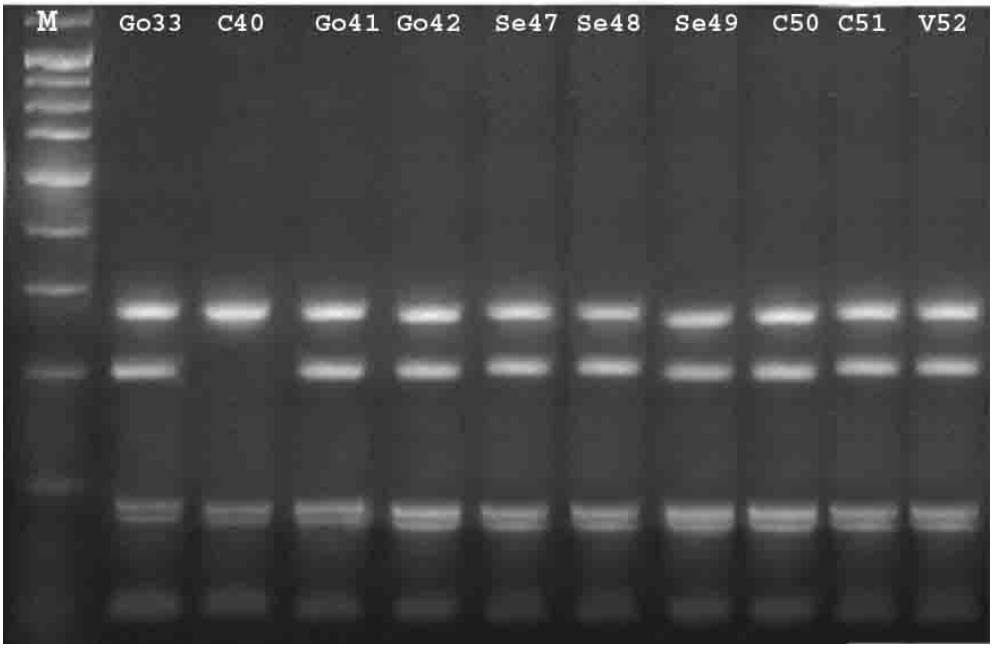
Genetic code table

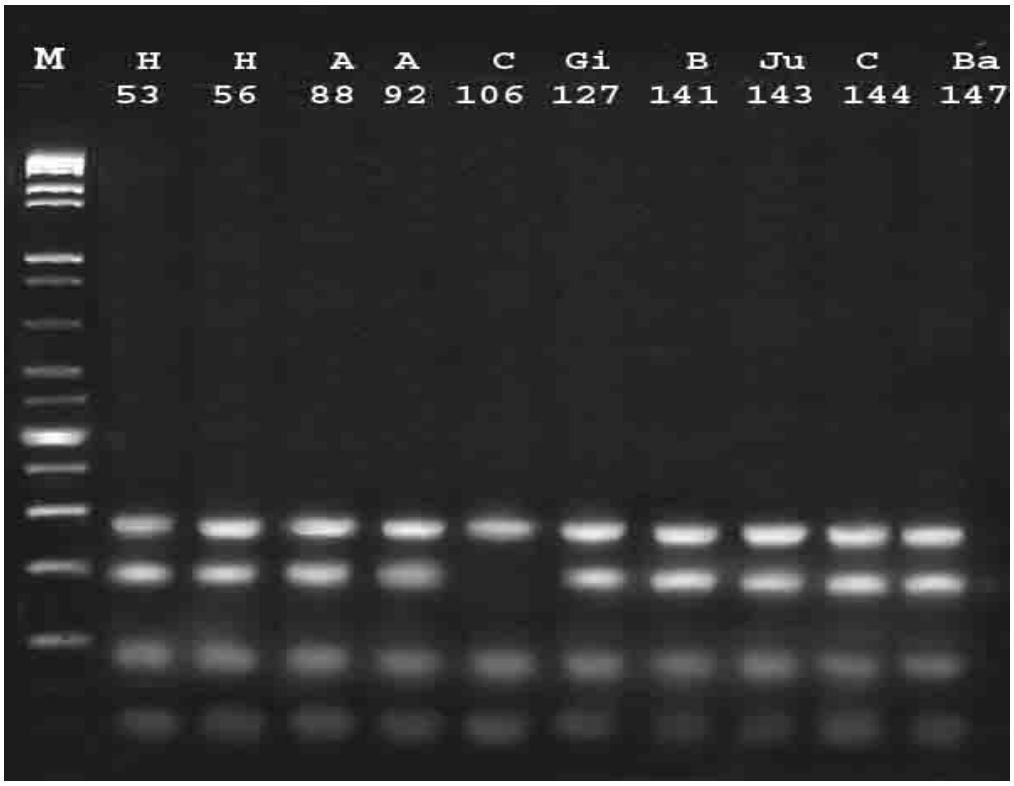


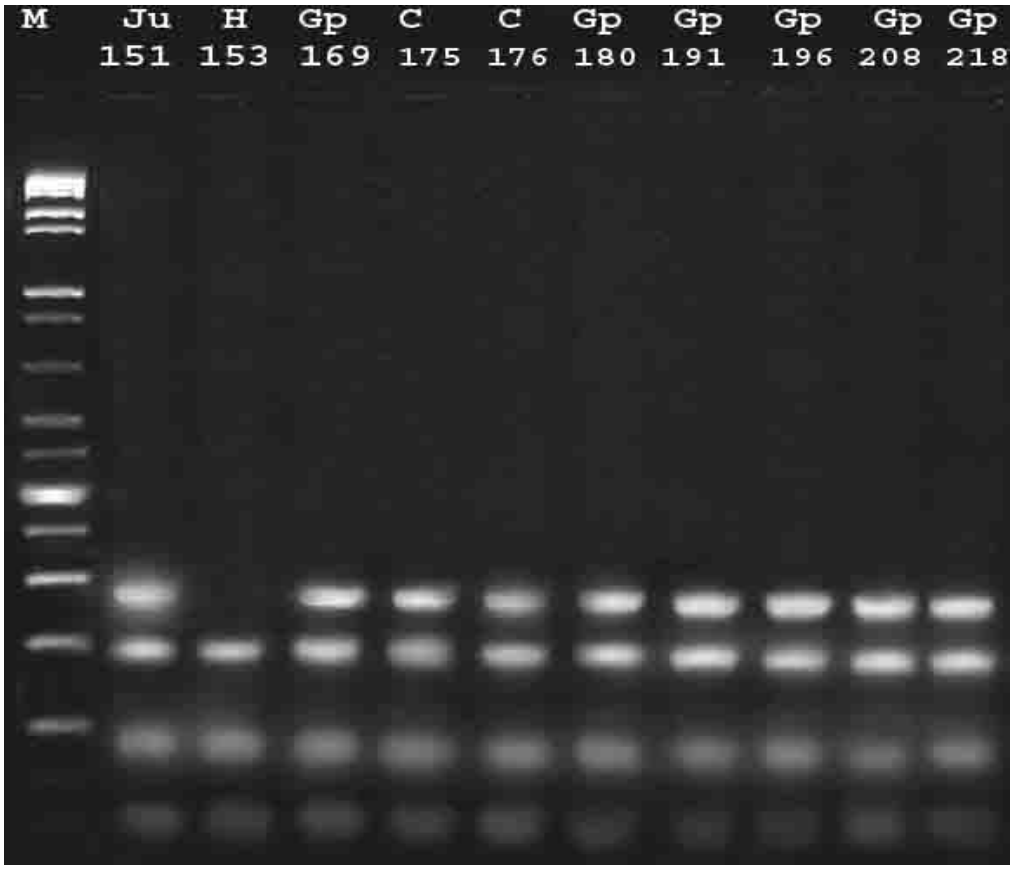
APPENDIX 5

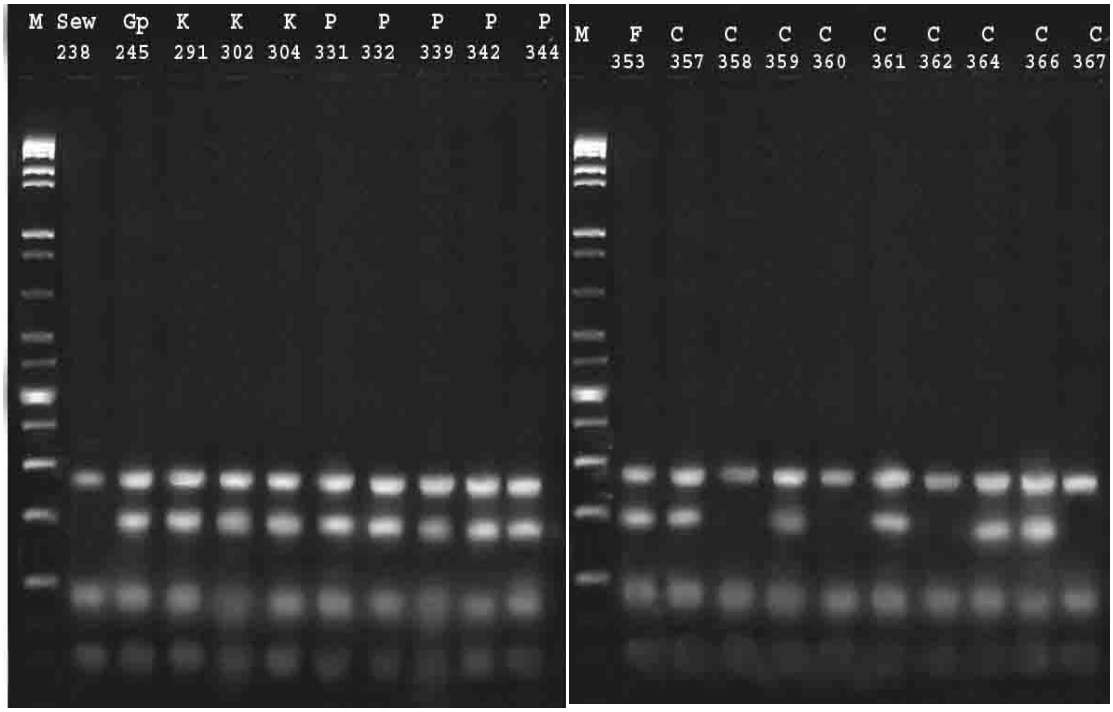
Agarose gel electrophoresis photographs

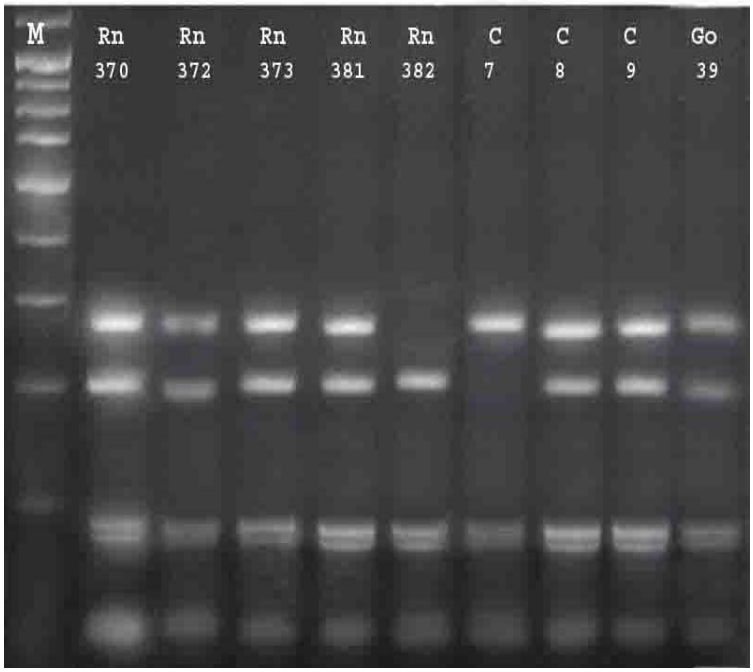





M
H 53
H 56
A 88
A 92
C 106
Gi 127
B 141
Ju 143
C 144
Ba 147


M
Ju 151
H 153
Gp 169
C 175
C 176
Gp 180
Gp 191
Gp 196
Gp 208
Gp 218


M Sew Gp K K K P P P P P
238 245 291 302 304 331 332 339 342 344
M F C C C C C C C C C
353 357 358 359 360 361 362 364 366 367


M Rn Rn Rn Rn Rn C C C Go
370 372 373 381 382 7 8 9 39

APPENDIX 6

List of screening test positive samples

S.N.	Sample code	sickling test	solubility test	RFLP test
1.	C3	+VE	+VE	AS
2.	C4	+VE	+VE	SS
3.	C5	+VE	+VE	AS
4.	C6	+VE	+VE	AS
5.	C11	+VE	+VE	AS
6.	C13	+VE	+VE	SS
7.	C14	+VE	+VE	AS
8.	Rj16	+VE	+VE	AS
9.	Rj17	+VE	+VE	AS
10.	Go32	+VE	+VE	AS
11.	Go33	+VE	+VE	AS
12.	C40	+VE	+VE	SS
13.	C41	+VE	+VE	AS
14.	Go42	+VE	+VE	AS
15.	Se47	+VE	+VE	AS
16.	Se48	+VE	+VE	AS
17.	Se49	+VE	+VE	AS
18.	C50	+VE	+VE	AS
19.	C51	+VE	+VE	AS
20.	V52	+VE	+VE	AS
21.	H53	+VE	+VE	AS
22.	H56	+VE	+VE	AS
23.	A88	+VE	+VE	AS
24.	A92	+VE	+VE	AS
25.	C106	+VE	+VE	SS
26.	Gi127	+VE	-VE	AS
27.	B141	+VE	+VE	AS
28.	Ju143	+VE	+VE	AS
29.	C144	+VE	+VE	AS
30.	Ba147	+VE	+VE	AS
31.	Ju151	+VE	+VE	AS
32.	H153	-VE	+VE	AA

33.	Gp169	+VE	+VE	AS
34.	C175	+VE	+VE	AS
35.	C176	+VE	+VE	AS
36.	Gp180	+VE	+VE	AS
37.	Gp191	+VE	+VE	AS
38.	Gp196	+VE	+VE	AS
39.	Gp208	+VE	+VE	AS
40.	Gp218	+VE	+VE	AS
41.	Sew238	+VE	+VE	SS
42.	Gp245	+VE	+VE	AS
43.	K291	+VE	+VE	AS
44.	K302	+VE	+VE	AS
45.	K304	+VE	+VE	AS
46.	P331	+VE	+VE	AS
47.	P332	+VE	+VE	AS
48.	P339	+VE	+VE	AS
49.	P342	+VE	+VE	AS
50.	P344	+VE	+VE	AS
51.	F353	+VE	+VE	AS
52.	C357	+VE	+VE	AS
53.	C358	+VE	+VE	SS
54.	C359	+VE	+VE	AS
55.	C360	+VE	+VE	SS
56.	C361	+VE	+VE	AS
57.	C362	+VE	+VE	SS
58.	C364	+VE	+VE	AS
59.	C366	+VE	+VE	AS
60.	C367	+VE	+VE	SS
61.	Rn370	+VE	+VE	AS
62.	Rn372	+VE	+VE	AS
63.	Rn373	+VE	+VE	AS
64.	Rn381	+VE	+VE	AS
65.	Rn382	-VE	+VE	AA
66.	C7	+VE	+VE	SS
67.	C8	+VE	+VE	AS
68.	C9	+VE	+VE	AS
69.	C39	+VE	+VE	AS

बिषय जानकारी तथा मन्जूरीनामा फारम

रुपन्देही जिल्लामा थारु समुदायमा हँसिया रक्तकोष रोगको प्रकोपको अध्ययन ।

प्रमुख अनुसन्धानकर्ता : मनोज कुमार थारु 9808013230/chaudharymanoj754@gmail.com **फारम** सँ.:

एम. एस. सी. चौथो सेमेस्टर, प्राणीशासस्त्र केन्द्रिय बिभाग

त्रिभुवन विश्वबिद्यालय, कीतिपुर, काठमाँडौ ।

प्रायोजक : विश्वबिद्यालय अनुदान आयोग, सानोठिमी, भक्तपुर

परिचय :

हँसिया रक्तकोष नेपालका थारु समुदायमा देखिने एउटा वंशाणुगत रोग हो । यो रोग लागेमा रगतको राता रक्तकोषहरु हँसिया आकारको हुन गई मसिना रक्तनलीहरुमा अड्किन्छन र फलस्वरुप हातखुट्टाहरु सुन्निने, दुख्ने, किटाणूको संक्रमण, रक्तअल्पता, कमजोरी, अल्पआयूमा मृत्यु आदिजस्ता लक्षणहरु देखिन्छन् । यो सरुवा रोग होइन । तपाइलाई यस अनुसन्धानमा भाग लिन आमन्त्रित गरिन्छ । यो अनुसन्धानमा भाग लिने व नलिने अधिकार केवल तपाइसँग छ । यो अध्ययनमा तपाइलाई हँसिया रक्तकोष छ व छैन भन्ने मात्र पत्ता लगाइनेछ उपचार गरिने छैन । यो जाँच तपाइको लागि निशुल्क छ । यसमा परिक्षणको लागि तपाइको रगतको नमुना लिइनेछ र यस क्रममा डीएनए परीक्षणसमेत गरिनेछ । यो अनुसन्धानमा तपाइको व्यक्तिगत परिचय गोप्य राखिनेछ ।

यो अनुसन्धानको उद्देश्य के हो ?

वर्तमान समयमा यो रोगको प्रकोप रुपन्देही जिल्लाको थारु समुदायहरुमा कति छ भन्ने पत्ता लगाउनु र यसको रोकथामको लागि सम्बन्धित सबैमा जनचेतना जगाउनु यसको प्रमुख उद्देश्य हो । साथै यो मेरो स्नातकोत्तर तहको थेसिसको बिषयशीर्षक पनि हो ।

यो अनुसन्धानमा क-कसले भाग लिन सक्छ ?

थारु समुदायमा पर्ने दश वर्ष उमेर पुगेका र हालमा रगत नचढाएको व रक्तदान नगरेको महिला र पुरुष जोसुकैले यसमा भाग लिन सक्छन्।

यो अनुसन्धानबाट आउन सक्ने जोखिमको जिम्मा कसले लिने ?

अनुसन्धानकर्ता स्वयं ।

माथि उललेखित सम्पूर्ण व्यहोरा मैले ठीकसँग बुझेँ । यो मन्जूरीनामाको एक प्रति मैले पाउने गरी स्वेच्छाले म रगतको नमुना दिन तयार छूँ ।

नाम : ..... उमेर : ..... वर्ष

बुवाको नाम : .....

लिंग : महिला / पुरुष

ठेगाना : .....

फोन नं : .....

विशिष्ट जाति समूह : बातर / कठरिया / डडहा / पूर्विया / कवँर अन्य : .....

हस्ताक्षर : ..... मिती : .....

APPENDIX 8

Hypothesis test

(case 1)

The genotypic frequency of homozygous dominant, heterozygous and homozygous recessive should be distributed in the ratio of 1:2:1 according to theory.

Genotype	AA	AS	SS
frequency	319	57	10

Null hypothesis (Ho): the distribution of HbS allele is common in Tharu community.
Alternative hypothesis (H_1): the distribution of HbS allele is not common in Tharu community.

Genotype	**Observed frequency (O)**	**Expected Frequency (E)**	**O-E**	**$(O-E)^2$**	**$\frac{(O-E)^2}{E}$**
AA	319	96.5	222.5	49506.25	513.02
AS	57	193	-136	18496.00	95.83
SS	10	96.5	-86.5	7482.25	77.54
	$\Sigma O = 386$	$\Sigma E = 386$			$\Sigma\frac{(O-E)^2}{E} = 686.389$

Chi square test of goodness of fit
Under null hypothesis, the test statistics is:

$$\chi^2 = \Sigma \left[\frac{(O-E)^2}{E} \right] = 686.389$$

i.e. χ^2 cal. = 686.389

(case 2)

There were total eight ethnic sub-groups of Tharu participated in the study and all show presence of sickle cell allele in in their gene pool except Marchaha Tharu.

Cast	**Observed frequency of negative sample (O)**	**Observed frequency of positive sample (O')**	**Total (N)**
Baatar	77	5	82
Kathariya	102	23	125
Dadaha	17	15	32
Kanwar	12	14	26
Khausiya	19	2	21
Purbiya	22	3	25
Dangoriya	67	5	72
Marchaha	3	0	3
Total(N)	319	67	386

Null hypothesis (Ho): the distribution of HbS allele is cast independent.
Alternative hypothesis (H_1) : the distribution of HbS allele is cast dependent.

The expected frequency of each cell was calculated by using relation,

$$Eij = (Ci\ X\ Rj)/N$$

The first cell is:

$$E11 = (C1\ X\ R1)/\ N=\ 319\ X82/\ 386 = 67.7668$$

Filling in the rest of the cells in the same way for each expected value, Eij, using the same equation but by using frequencies from the corresponding row Ri and column Cj for each cell, we obtained the following expected values:

Cast	Expected frequency of negative sample (E)	Expected frequency of negative sample (E')	Total (N)	$\frac{(Oij - Eij)^2}{Eij}$	$\frac{(O'ij - E'ij)^2}{E'ij}$
Baatar	67.7668	14.2332	82	1.25801	5.98962
Kathariya	103.303	21.6969	125	0.016438	0.0782643
Dadaha	26.4456	5.5544	32	3.37369	16.0628
Kanwar	21.487	4.51295	26	4.18876	19.9435
Khausiya	17.3549	3.64508	21	0.155937	0.742448
Purbiya	20.6606	4.33938	25	0.0868287	0.413408
Dangoriya	59.5026	12.4974	72	0.944684	4.49782
Marchaha	2.47927	0.520725	3	0.109369	0.520725
Total	319	67	386	58.3823	

Chi square test of independence

Under null hypothesis, the test statistics is $\chi^2 = \Sigma \{\frac{(Oij - Eij)^2}{Eij} + \frac{(O'ij - E'ij)^2}{E'ij}\}$

$= 58.3823$

i.e. χ^2 cal. = 58.3823