CHAPTER - 1

INTRODUCTION
INTRODUCTION

Every individual surviving on this Earth is incredibly unique in every respect. Man though have descended from a common ancestor had gradually accumulated over a long period of time different changes. These changes are a gradual process leading to change in gene frequencies of a population and in course of time results in different genetic makeup. One individual differs from other not only in terms of physical appearance like height, weight, skin colour, nose pattern, facial appearance etc. but also has different eating habits, behaviour and culture. The physical variation observed among human populations surviving in different geographical regions of the world or even those of the same region is a clear evidence of ever going human evolution.

Different human populations live in clusters over wide geographical regions and show wide differences in socio-cultural attributes. The biology of a population is governed by a complex interaction of environmental and genetical factors. The genetic constitution of a population in interaction with environment of ecological niches produces variation among different populations inhabiting the same or diverse habitats. The habitats of environmental factors in association with the micro-evolutionary forces ultimately influence the biological structure of the human populations (Pandey et al., 2012). Genetical instructions which one receives from his or her parents together with the effect of different environmental factors also guide the development of an individual. The small variation slowly gets accumulated resulting in due course of time into a new type different from their predecessor and contemporaries of that region.

Evolution is a process of change which gradually results into different gene frequencies of the population. The evolutionary process results into diversity at every level of biological organisation, including species, individual organisms and molecules such as DNA and proteins. According to Underwood (1979), evolution is a non-cyclical change in the allele frequencies of a population’s gene pool from one generation to another. The factors which contribute to evolution are mutation, migration, selection, adaptation and genetic drift. Boas (1911) has first shown the effects of migration by making investigation on the differences between the migrants and their American born children.
Subsequent studies in Hawaiian born Japanese, Chinese born in United States, migrant Mexican-American and their children, and among various other groups, confirmed the existence of differences in their first generation offspring in certain traits such as head shape, body height, body weight and pigmentation. According to Chai (1972), slow and small rate of change essential for speciation is known as microevolution. Das (1981) gives the opinion that the development of dissimilarities between the ancestral and descendant population is the result of evolution. The dissimilarities can be in gene frequencies as well as morphological characters. Every slight change in gene frequencies that occur generation after generation in population of various sizes has been called as microevolution by Kraus and White (1956). Dobzhansky (1972) opines that microevolution and macro evolution are only relative terms with no difference in the underlying causal agencies. The cumulative effect of microevolution through ages finally leads to macroevolution.

Socio-cultural variables such as economic activities, linguistics, religions, marriage patterns, dietary habits, social instructions etc. are equally important factors in bringing about human evolution. Moreover, the socio-cultural differences among the populations, to a considerable extent determine the rate of fertility and mortality. One classical example which illustrates the link between Natural selection and socio-cultural activities is the spread of haemoglobin S gene in association with the spread of agriculture from Middle East to Peninsular India and Africa (Weiss and Mann, 1978; Singh, 1991).

The molecular level of evolutionary process involves substitution, insertion, addition or deletion, of nucleotide in the DNA. If the DNA encodes a polypeptide, this process brings about changes in the amino acid sequence. These changes gradually accumulate and in course of time finally results in a molecule that bears little resemblances to its progenitor. Advances in molecular biology have made it possible to determine the nucleotide sequences of DNA and the amino acid sequences of polypeptides. By comparing related sequences, the molecular details of evolution can be analysed (Gardner, 2006). The term ‘breeding population’, also referred to as a ‘Mendelian population’ is the unit of study of genetic variation in man (Bhasin and Walter, 2007). Harrison et al. (1988) quotes that “the collective unit of evolution is the population and it is in populations that all the forces we have considered operate”. Thus according to Bhasin and Walter (2007), various factors like selection, gene flow,
genetic drift, founder effects etc. are always operating on populations and in course of time result into shaping of the specific genetic profiles of populations. The minimal integrated unit of evolutionary changes is the ‘breeding population’. As far as delineating evolutionary factors are concerned, the ‘breeding populations’ as a unit of study meet almost every logical requirement of the unit and any change in its genetic profile from one generation to the next will constitute an evolutionary change.

The study of genetic variation in man has been used to understand the ‘breeding populations’ as biological or evolutionary units in man and to describe them in terms of gene frequencies or if this is not possible (anthropometric, morphological, dermatoglyphic, biochemical, molecular etc. traits) in terms of phenotype frequencies and mean values, respectively. Such exact and comprehensive descriptions are the basic requirements for the understanding of genetic variation in man and thus are used for the analysis of the various evolutionary factors, which caused this variation in course of time (Bhasin and Walter, 2007). The association between gene frequencies and environmental factors has been made for a number of polymorphic loci (Flint et al., 1986; Flatz, 1987). It seems that this type of associative study is an effective way to make inference about selective mechanisms, which could be done by having data on the distribution of gene frequencies among various populations, provided the population is endogamous.

The study of human population genetics deals with the genetic variation and the processes that influence these variations within and among the populations. The changes in the genetic makeup of a population over time, is usually measured in terms of allele frequencies, which is taken as equivalent to evolutionary change. Because of this reason, population genetics provides the groundwork for scientists understanding the evolutionary process, specially microevolution, or changes occurring within one or different populations over a time span. The existing of sufficient variation even at a small number of genetic loci can be helpful to study population variations existing among or within populations. However, the study of distribution of a large number of polymorphic alleles is more useful to study the various aspects of population dynamics (Kumari et al., 2013).

Gene frequency data are very useful for studying the genetic relationship and evolution of human population. Comparison of gene frequencies for one or two loci is
not reliable since each locus has a different distribution. Only when a large number of loci are examined, the genetic relationship becomes clear (Cavalli-Sforza and Edwards, 1964). This is particularly because the inter-populational genetic variation is very small compared with the intra-populational variation at the gene level (Lewontin, 1972; Nei and Roychoudhry, 1974; 1982). However, if a large number of loci are examined, even small differences can be detected with sufficient accuracy. If there are gene frequencies for a number of loci in a population, the heterozygosity for individual locus can be calculated and average heterozygosity per locus for a population is obtained. The average heterozygosity indicates the magnitude of genetic variation that exists within a population. Once a set of population have been investigated for a number of genetic characters, the populations can be grouped in some way on the basis of their inter-relationships. The similarities and differences between the populations in different groups can then be used to get some idea about the nature of the forces which have prevailed in the evolutionary history of the population. The inter-relationships may be studied in terms of the variations in the gene frequencies at individual loci (Arzoo, 2005). The genetic differentiation between a pair of population is usually measured by a quantity called the genetic distance which is a function of gene frequency. There are several different measures of genetic distances (Sanghvi, 1953; Cavalli-Sforza and Edwards, 1964; Latter, 1973; Nei, 1973). Once genetic distance is estimated for a group of populations, their genetic relationships can be studied by using dendrogram, principal component analysis etc. (Cavalli-Sforza and Bodmer, 1971; Nei, 1973; Sneath and Sokal, 1973).

India has complex ethnic history and different climatic and ecological zones; it offers a rich field for the study of genetic differentiation process in man, and also for the analysis of factors responsible for considerable genetic variability seen among its various populations (Bhasin, 2002). It has almost all major racial groups of mankind. Present-day India is represented by a complex socio-cultural mosaic comprised of 20 major languages and 750 dialects (Kosambi, 1991) partitioned into 2,000 castes and tribal groups (Puppala and Crawford, 1996). During prehistoric and historic times, multiple waves of migration have taken place in India. The subsequent cultural differentiation have resulted in strict rules of governing mating practices, which is the major cause of the genomic diversity observed among contemporary ethnic Indian populations.
The contemporary people of India are culturally stratified as tribals or the aborigins and non-tribals or the Aryans. The tribals can be classified into one of the three language families: Austro-Asiatic, Dravidians and Tibeto-Burman. Contemporary non-tribal populations of India tend to belong to the overall Hindu religious fold and are hierarchically arranged into four main caste classes, viz. Brahmin (priestly class), Kshatriya (warrior class), Vasya (business class) and Sudra (menial labour class). In addition, there are several religious communities, who practice different religions, like Islam, Jainism, Judaism, Christianity, Sikhism and Buddhism. The language spoken by non-tribals predominately belongs to the Indo-Aryan and Dravidian families (Basu et al., 2003). Caste system in India has its origin in verna system with its language, state and religious base, hence study of caste differentiation can be undertaken with these three points of views (Karve, 1961).

The vast spectrum of genetic diversity in Indian populations has been a subject of great interest to population geneticists, anthropologists and evolutionary biologists. The studies conducted among different Indian populations based on morphological and classical genetic markers show greater diversity (Bhasin and Walter 2001), when compared with existing global populations. The enormous genetic diversity among various sub-castes and castes is guided primarily by geographical proximity, ethno-historical factors, and socio-cultural rules of marriage practices (Malhotra 1979; Malhotra and Vasulu, 1993; Cavalli-Sforza et al., 1994; Bhasin and Khanna, 1994). Tandon et al., (2004) opined that some studies indicate that in some regional populations factors like ethno-historical background play a significant role than other factors while in certain other cases geographical proximity constitute a strong force than socio-cultural factors. In India, most of the Hindu and Muslim communities though have been following strict endogamy but the gene flow mostly takes place among the people with similar occupation, tradition and those inhabiting same geographical territories. But in general, among tribal populations, isolation and endogamy plays a greater role in maintaining the genetic diversity whereas in some regional populations, especially among castes and sub-castes in northern parts of the country, geographical proximity is observed as major influencing factor than socio-cultural variables.

Muslims in India comprise more than 12% (Shariff, 1998) of the population and they exhibit diversity of linguistic and ethnic groups belonging to different
biradaris (so-called castes), besides few tribes. At present, Islam is the second most practiced religion in India after Hinduism, encompassing 13.4% (138 million) of the total Indian population (Census of India, 2001). The presence of Islamic settlements in India is believed to have resulted into at least three distinct campaigns (Farah, 2003) initiated from different geographic regions. In 711 AD, an Arab military invasion precipitated the formation of the Sind Indo-Muslim state in the Indus delta region (Keay, 2000). A few centuries later, between 997 and 1027 AD, Muslim converts from the Central Asian Turkish tribe staged multiple raids into the northwest province of Punjab. Finally, during the 13th and early 14th centuries AD, Afghan and Persian Muslims arrived from the northwest, reached New Delhi and from there, penetrated into points east, west, and south (Wolpert, 1991). These Muslim immigrants, who were mostly males, reportedly married local Hindu females and generated a new admixed genetic pool, perhaps with sex-specific differences (Robb, 2002; Naqvi, 2003). In addition to different possible source populations, the Indo-Muslim groups may have subsequently evolved through several distinct cultural modes: cultural diffusion, elite dominance via military expansions, and colonization, which may have involved varying levels of genetic admixture with the indigenous Indian groups. It has been stated that today most of the Muslims in the Indian subcontinent represent the descendants of converts and are the offspring of Hindu mothers (Wolpert, 1991). According to Roychaudhary et al. (2000) the Muslims originated in one of the two ways (a) groups came and settled during the historic migrations and invasions (b) another group is believed to be formed through process of proselytization of the indigenous Hindu population of comparatively lower and middle order in the prevailing caste hierarchy, besides sections of a few tribes living in the fringe area of caste dominated regions.

Muslims belong to two major sects: Sunnis and Shias, each sect has different groups, which are grouped under Ashraf and Ajlaf (Ansari, 1959). The Ashrafs comprise of higher rank Muslims like Syeds, Sheikhs, Pathans and Moghuls while the Ajlafs comprise of Qureishis, Ansaris, and other groups of lower occupation (Ahmad, 1978). A large number of the latter may also have been converted from local indigenous populations of other faiths (Afzal, 1984). While the Muslim conquerers from the North-West, though numerically a minority group is almost left untouched because of some unexplained reasons.
Manipur was an ancient Kingdom where the first known king ruled in 33-154 CE according to the *Cheitharol Kumbaba*, the royal chronicle (Singh and Singh, 1989). Curiously, in tracing the early history of the Manipuri Muslims waiting for the period of Muslim conquest of Bengal unto Assam and so on was not needed, as they had already advanced in Manipur from unlikely directions which was facilitated by the nature of location as the meeting point of the Mongoloid and Indian worlds (Yule, 1968; Sanajaoba, 1988), at the foot of the eastern Himalayan range, serving as 'silk route' connecting the Middle East and Indo-China (Phayre, 1983; Sharma, 1994). Thus Manipur was historically a convenient and strategic trans-Asiatic route that attracted the curiosity of Alberuni (973-1048 CE). He described the kingdoms of Kamrup and Udayagiri (Manipur) in his *Kitab al Hind*. “The Muslims joined the Meitei society before the people (Meitei) professing Hindu religion (Vedic religion) migrated to Manipur and became part of Meitei community (Mangang, 1988; Ahmed, 2011). Manipuri Muslims are the descendants of Turks (Moghuls), Mongols, Pathans, Afghans, Bengali Muslims as well as local converts to Islam (Irene, 2010). Some Muslims arrived in Manipur in early seventh century CE, though main settlement in Manipur is traced to 1606 CE (Irene, 2010). “They spoke languages like Urdu, Farsi, Turki, Arabic etc. before they started adopting Meiteilon (Manipuri) but they for long continued to use many words and phrases of these languages (Irene, 2010).

Several studies have been made on different Muslim populations or they have been studied as a part of local Hindu and Muslim populations using morphological and behavioural traits; serological and biochemical markers (Rizvi, 1984; Hakim, 1971; Chahal et al., 1989; Reddy and Reddy, 1989; Das et al., 1985; Das, 1980; Vijayakumar et al., 1987; Chahal et al., 2004; Arzoo, 2005; Pandey et al., 2013; Chadha and Sandhu, 2013). The prevalence of cousin marriages among Muslims has attracted many scholars who are engaged in the study of biological consequences of inbreeding (Afzal and Sinha, 1983; Badaruddoza, 1992). Several authors were also trying to reconstruct the origin of Muslim populations in India using molecular markers. Terreros et al. (2007) studied the mtDNA composition of two Muslim sects from the northern Indian province of Uttar Pradesh, the Sunni and Shia, using sequence information from hypervariable regions 1 and 2 (HVI and HVII, respectively) as well as coding region polymorphisms. Their study revealed that Y chromosome and the mtDNA haplogroup composition of the Indo-Sunni align closely
to the neighboring Indian populations. But the mtDNA haplogroup differences between the Indo-Shia and Indo-Sunni do not signal a differential contribution of foreign mtDNA to either sect. Yet, they may be indicative of unique population dynamics with native castes and tribes.

To estimate the contribution of West Asian and Arabian admixture to Indian Muslims, Easwarkhanth et al. (2010) assessed genetic variation in mtDNA, Y-chromosomal and LCT/MCM6 markers in six Muslim communities from different geographical regions of India. They found that most of the Indian Muslim populations received their major genetic input from geographically close non-Muslim populations. However, low levels of likely sub-Saharan African, Arabian and West Asian admixture were also observed among Indian Muslims in the form of L0a2a2 mtDNA and E1b1b1a and J*(xJ2) Y chromosomal lineages. Between the Iranian and Arabian sources it was difficult to make the distinction with mtDNA and the Y chromosome, because of similar gene pool compositions in the sources, the estimates were highly correlated. However contrast could be observed in the LCT/MCM6 locus, which shows a clear distinction between the two sources and enabled the researchers to rule out significant gene flow from Arabia. Overall, their results support a model according to which the spread of Islam in India was predominantly cultural conversion associated with minor but still detectable levels of gene flow from outside, primarily from Iran and Central Asia, rather than directly from the Arabian Peninsula.

Several workers have studied the genetic diversity among some localized groups of Manipur using classical genetic polymorphisms to understand the genetic structure and microevolution process (Singh, 1978; Chakravartti et al., 1979; Shah, 1990; Kapaiwo, 1995; Shanta and Shyamacharan, 2001; Soram et al., 2014). Singh et al. (1986) studied different serological and biochemical markers in Brahmin and Meitei populations of Manipur. But very few studies have been made on Manipuri Muslims. Molecular genetic studies based on autosomal, mtDNA and Y-chromosomal markers have been helping to understand the population stratification in addition to validating the mythological history of origin, route of migration and affinity with other groups. Few studies have been reported based on biochemical (Asghar et al., 2009; Achoubi et al., 2010; Singh et al., 1986) and molecular (Saraswathy et al., 2009a, b, 2012; Meitei et al., 2010; Achoubi et al., 2012; Asghar et al., 2013) markers on Manipuri Muslims but none of them attempted to understand
the population sub-structuring within the population (Asghar et al., 2013). No systematic attempt has ever been made to study the genetic variation, ethnic affiliations and phylogenetic relationships of different castes of Manipuri Muslims which are endogamous, isolated and culturally divergent groups. Besides, no work has ever been conducted on Manipuri Muslims using enzyme markers like Esterase D, Acid Phosphatase I and Sickle cell trait etc. which has been attempted in our present study.

Present study attempts to study the genetic polymorphism using various morphological and behavioural traits (Colour blindness, PTC tasting ability, tongue folding, tongue rolling, hypertrichosis, hitchiker thumb, cleft chin, ear lobe pattern, widows’s peak, ear wax, Darwin tubercle and dimple); serological and biochemical markers (ABO, Rh, Bombay Phenotype, sickle cell trait, G6PD, Esterase D and Acid Phosphatase I) and a microsatellite marker (TPOX) among different Muslim populations of Manipur.

According to Bhasin et al. (1992) the human genetic variations play a significant role in bringing about the diversity in population structure and contribute to the dynamics of human evolution. The genetic similarities which exist between populations show the common origin or admixture of gene pools, while the genetic heterogeneity between the populations indicates the diversity or isolation by some unknown barriers. A population is characterized by a set of gene frequencies. Because of this, the gene frequency data are essential prerequisite for studying the genetics of any population. The population diversity could provide an opportunity to study the morpho-genetic, sero-biochemical, molecular level variation and similarities among different populations inhabiting the region since they have cohabited for a long time and presumably there might have been gene flow between them, since the existence of genetic variation in man is caused by interaction of many factors like selection, migration, gene flow and genetic drift. So the present study could help to understand the above mentioned factors in six different populations of Manipur which includes Muslims with four different castes i.e. Sheikh, Syed, Pathan, Moghul and the other group includes Meitei (Hindu) and Naga, the latter a tribal population of Manipur.

The present work is divided into six chapters. Chapter one is the Introduction and chapter two deals with the review of literature. The relevant literature related with
the different genetic markers which includes morphological and behavioural markers, serological and biochemical markers and the study of genomic marker i.e. STR marker (TPOX), has been incorporated in this chapter. The aim and objective of the present study is also discussed in this chapter. The third chapter deals with the description of the land and the people of Manipur. This chapter also deals with the distribution of different major ethnic groups in the state and also the ethno history of the populations taken in the present study. The fourth chapter outlines the methodology that has been used to meet the research objectives. A brief description of the sampling procedure, different techniques used for the study of morphological and behavioural markers, serological and biochemical markers, STR marker (TPOX) has been thoroughly discussed in this chapter. The results obtained are given in the fifth chapter. And finally, discussion and conclusion is given in the last chapter.