CHAPTER II
REVIEW OF LITERATURE

In the present chapter, we shall make a review of related literature with special reference to those works carried out in populations of Northeast India. It may be made it clear that the following review is far from being complete and exhaustive, but its main purpose is to have a glance at those related literature with a view to understanding of the genetic structure of the study population.

Phylogenetic Affinity

Establishing the genetic relationship between and within human populations is one of the major focuses of anthropological population genetics. Along with genetic markers of the blood, morphological (including dermatoglyphic, somatometric, somatoscopic and behavioural traits) characters have been widely used for this purpose. Of course, it is generally believed that morphological and behavioural traits like arm-folding, hand-clasping, earlobe attachment, tongue rolling, mid-phalangeal hair, etc., are not as valuable as genetic markers from the genetic point of view. In the present study, however, we agree with Salzano (1961) that such traits may be useful in population genetic studies for several reasons. One of such reasons is that the findings on the genetic affinity of human populations in respect of known loci are still not as clearly understood as were in the case of these traits. In several cases, the findings on genetic markers corroborate those on anthropometric and other morphological and behavioural traits (Harrison, 1977). Moreover, it is also realized that even the so-called non-adaptive traits like blood groups are completely non-adaptive. According to Boyd (1952 from Crawford, 1973), 'it is doubtful if any hereditary characters are completely non-adaptive, and . . . probably we can deal only with different degrees of adaptive value'.

As mentioned earlier, anthropologists before 1950s and even in the late 1960s took great interest in the racial classification of populations. However, after 1950s a new development in anthropological research had emerged because of the advancement in
population genetic research. It has been realized that the use of genetic markers is very helpful in understanding the genetic variation and affinity between populations. Also, the genetic variation does exist not only between races but also within a race. For this reason, genetic relationship of populations within a particular racial group in respect of various genetic markers has attracted a great deal of attention among the physical anthropologists (see review, Dobzhansky et al., 1973; Crawford, 1973; Majumder, 1991). The main purpose of this short review is not to mention all those important works in the field of anthropological genetics, but to make a glance at the overall scenario that has developed in this field. In fact, we would like to point out that the population genetic researches in Northeast India have also reflected the overall interest in the field of population genetics all over the world. Therefore, it may be necessary to have a glance at the development in this field with regards to populations of Northeast India.

Like in other parts of the globe, scholars in Northeast India have used the various genetic markers with a view to understanding the evolutionary relationship between populations of the region. Perhaps, Mitra (1938) for the first time reported data on the ABO blood groups for the Plains Assamese, Angami and Lushai populations of this part of the country. It was then followed by two sets of data among the Khasi (Basu, 1938; Macfarlane, 1941), and those data reported by the British Research Association Committee (1939) for the Angami, Konyak and Ao tribes of Nagaland. But a large number of data on the ABO blood groups were published after India’s Independence, though most of the works have been carried out in Assam only (see reviews Das, 1974; Phookan, 1975; Chakravarty, 1990).

As just mentioned above, for the purpose of this presentation, an attempt has been made to give a general picture rather than giving a detail of all the works carried out by different scholars. One of the important general characteristics, relating to the gene frequencies of the ABO blood groups in the population of Northeast India, is that the B allele occurs more frequently than the A allele in the tribal populations. While among the Hindu caste populations, the frequency of A allele is higher than that of B. In fact, this is one of the general differences between the Mongoloid (tribal) and Caucasoid (non-tribal) populations in Northeast India. In the case of O allele, there is not much difference between tribal and non-tribal populations. It may be noted that the Hindu and
Muslim populations of Assam are similar in respect of the ABO blood group system, occurring in the pattern of \( O > B > A > AB \) (Ahmed Das, 1980), though the Muria Muslims show more resemblance to the tribal population (Danker-Hopfe et al., 1988). According to Das (1984), "No doubt that among the tribal population of Northeast India O gene predominates. The other two genes, i.e. A and B also occur in considerably higher numbers. But among most of the groups like the Khasi, Naga and Lushai-Kuki A (allele) is much more frequent than B. On other hand, among the Boro B (allele) is slightly more frequent than O. The reverse is true in the case of certain Arunachal tribes."

With respect to the phylogenetic position of population, various studies have made a comparison of their findings with those reported for other populations. Some studies try to find out the differences and similarities between the so-called Mongoloid and Caucasoid populations (Das et al., 1987; Sengupta, 1987; Danker-Hopfe et al., 1988); others deal with the genetic variation within these populations (Phookan, 1975; Das, 1978; Das et al., 1985a, 1985b, 1986a, 1986b; Singh et al., 1986), and some others made a comparison of these populations with other populations living outside the region (Flatz et al., 1972; Das, 1978). For comparison within and between populations, some studies have used only ABO blood groups, and others have considered with anthropometric traits and other genetic markers like PTC, colour blindness, and other blood group polymorphisms.

Danker-Hopfe et al. (1988) have made an attempt to show the phylogenetic relationship of 13 endogamous Assamese-populations on the basis of certain genetic markers, finger ridge patterns, anthropometric and behavioural traits. It was assumed that the two major population groups, namely, Mongoloids and Caucasoid, would form two distinct clusters as revealed by the distribution of five ABO blood group polymorphisms (Das et al., 1987). The distance analyses of Sanghvi (1953), Hériaux (1965) and Nei (1972) were taken into consideration. "It appears that distance analyses of data of different natures produce different results. The population exhibits differences and similarities among themselves in different manners with regard to different traits.... With regard to genetic traits the populations present a dendrogram which is difficult to explain" (Danker-Hopfe et al., 1988). In fact, many studies have revealed the existence
of such a difficult situation. For example, Walter et al. (1986) have reported the
distribution of haptoglobin, transferin (Tf) and Gc sub-types among the Brahmin, Kalita,
Kaibarta, Rajbanshi, Muslim, Ahom, Chutia, Kachari and Sonowals. It is evident from
this study that the Sonowals belong to one sub-cluster with the Ahom and Chutia.
Accordingly, the authors have suggested that the Chutia and Ahom are from the same
racial stock and thereby these populations “show a close genetic relationship”. It may,
however, be noted that these two populations are distant from each other with respect to
the distribution of GM and Km allotypes (Walter et al., 1987). Walter et al. (1986) have
also shown that the Brahmin, Kaibarta and Rajbanshi form another sub-cluster with the
Brahmin showing a ‘somewhat different position’. They have explained that the
Brahmins are different from the Kaibarta and Rajbanshi because of the absence of
enough gene flow due to caste marriage system. It is, however, surprising to find that
the Brahmin show a close genetic relationship to the Sheik Muslims and the Kalita
(Danker-Hopfe et al., 1988), which is difficult to explain. Similarly, according to the
Nei’s distance analysis of genetic traits carried out by Walter et al. (1987) and Danker-
Hopfe et al. (1988), the Kaibarta and Kalita are very close to each other. But these two
populations are quite different from each other in respect of the distribution of
haptoglobin, transferin and Gc polymorphisms (Walter et al., 1986).

Without going into detail of all the contradictory results, it may be pointed out
here that different results are shown according to different studies. Consequently, the
interpretation of the phylogenetic affinity of the populations in Northeast India has
become more complicated. In most cases, the results of the analyses of genetic traits are
more or less similar to those of traditional anthropometry. Harrison (1977) has also made
such an observation on the over-all picture of population genetic study. In other parts of
the world there has been an increasing interest in DNA polymorphisms so as to have a
better understanding of the origin of modern humans and the phylogenetic relationship
between various human populations. It may be noted that biologists, especially the
molecular anthropologists, have recently given more attention to mitochondrial DNA
analysis with a view to understanding the origin and divergence of human populations
(Johnson et al., 1983; Cann et al., 1987; Stoneking, 1993; Hagelberg, 1996). Human
mitochondrial DNA (mtDNA) is a self replicating circular molecule of approximately
16,569 base pairs in length, which is located in the cellular cytoplasm (Anderson et al., 1981). It has only 37 genes, no introns, and codes for 13 polypeptides that are essential to the energy metabolism of the cell (Hagelberg, 1996; Bertranpetit et al., 1996). The mother to her offspring always transmits it. As such, there is no recombination during meiosis and thereby every individual (male or female) receives identical copies of the mother’s DNA genome. This is important in constructing the phylogenetic relationship of human populations. Being inherited in maternal fashion, any variation in mtDNA through generations should be due to mutation. Thus, evolution in this sense takes place by the accumulation of mutations from generation to generation. Since mtDNA evolves about 10 times faster than nuclear DNA, it implies that deleterious mtDNAs are important in understanding not only genetic differentiation between and within populations, but human diseases and ageing as well (Wallace, 1995;). Similarly, there are numerous neutral or harmless mutations, which can provide useful genetic information of the human populations. Since each human cell consists of thousands of mtDNA copies, it is possible that mtDNA sequences should persist for a long time in biological samples of archaeological sites, thereby providing useful information concerning paleoanthropological interests (Hagelberg, 1996).

Since the early 1980s, mtDNA analysis has occupied an important position in the study of the evolutionary relationship of human populations. Perhaps, the most notable example of such works is that carried out by the late Allan Wilson and colleagues at the University of California in Berkeley (Cann et al., 1987). Using a high resolution mapping of mtDNA of 147 women from Africa, Asia, Europe, Australia and New Guinea; these scholars (Cann et al., 1987) have found that there is little variation in the mtDNA types of all the women, despite the differences in their geographical origins. Moreover, women with an African ancestry show more variation in the mtDNA types, suggesting that the African origin was the oldest and had more time to accumulate mutations. Therefore, it is suggested that all modern mtDNA types could be traced to a single female ancestor, known as the ‘African Eve’, who lived in Africa about 200,000 years age, or between 290,000 and 140,000 years ago. Interestingly, the concept of an African Eve seems to be consistent with the fossil record, which suggests that all modern human populations trace their origin to Africa (Poirer et al.,1994). In other words, it has
been interpreted that the finding of Cann et al. (1987) is in confirmation with the SingleReplacement Hypothesis, which states that all modern humans originated in Africa only, and then replaced the archaic humans throughout the Old world in a recent expansion from Africa. This hypothesis is in contrast to the Multi-regional Transition Hypothesis, or Regional Continuity Model, which states that although Homo erectus originated in Africa, the archaic Homo sapiens evolved independently into modern humans in different parts of the world with enough hybridization to produce a single biological species of modern Homo sapiens (Stringer and Andrews, 1988; Hagelberg, 1996). Although the present work are not concerned with DNA analysis, we just mention to acknowledge and appreciate the latest development in population genetics, which in future, we hope that such studies would be carried out in this part of the country with a view to having a better understanding the phylogenetic position of populations (Khongsdier, 2001).

DEMOGRAPHIC-GENETIC STRUCTURE

According to Roberts (1973), “No is it satisfied with counting genes in populations. Instead, being aware of differences in gene frequency between peoples, human population genetics has endeavoured to understand the processes responsible for these differences and the mechanisms by which the observed frequencies are maintained and regulated”. Thus, genetic markers are used not only to understand the genetic composition and population affinity but also to quantify the evolutionary processes of various evolutionary forces. “Following these applications of the concept of genetic constitution, the frequencies of genes in the array that characterises a given population, there came a new concept, that of genetic structure. Where as genetic constitution is concerned essentially with individual loci, genetic structure concerns the way in which genes are distributed and combined within populations. As such it is concerned not with gene frequencies but with measures of gene relationships (linkage disequilibrium coefficients of inbreeding, coefficients of kinship, parameters of the decline of kinship with distance). For all these, factors are of relevance that do not enter the simple concept of genetic constitution – the effective population size, population distribution, population
density, assortative mating, migration. These all affect the evolution and differentiation of populations, and are themselves affected by social, cultural, as well as natural environmental, factors” (Roberts, 1991). Thus, the aim of population genetics is now not only to understand the genetic constitution of a population but also to make out the genetic structure of such a population.

Population structure is characterised by a colossal number of interrelated components or characters that may be arranged in terms of genetic, taxonomic, demographic, social and ecological hierarchical orders of relatedness for the expediency of a given study at a given point of time. Basu (1995) writes, “While (1) in genetics we have the hierarchy of endogamous groups..., (2) in demography we have the hierarchy of segmentation as well as the hierarchy of age groups, (3) in social science we have the hierarchy of social groups arranged in ascending/descending order of social status/economic condition/power (authority) as in the case of caste, class and community; (4) in ecology we have the hierarchy of populations inhabiting niches, subniches, and so forth within the broad range of distribution of the group...; (5) in taxonomy we have the hierarchy of categories, i.e. phylum, class, order, etc. whether or not we accept the existence of infra-specific taxa in the case of humans”. Each of these structures is closely interrelated (Harrison and Boyce, 1972; Yablokov, 1986). The relationship between demography and genetic structure, i.e. demographic-genetic structure, has been a focus of attention in population genetics (Neel and Salzano, 1967; Basu, 1969; Roberts, 1968; Salzano, 1972).

From the genetic point of view, endogamous groups are known as Mendelian populations. “A Mendelian population is a community of individuals of a sexually reproducing species within which matings take place. There is a hierarchy of Mendelian populations. The most inclusive Mendelian population is the species. The lowermost member of the hierarchy is a panmictic unit, within which matings take place” (Dobzhansky, 1970). Despite a number of problems, it is believed that an understanding of the concept of the hierarchy of Mendelian populations is a vital requisite to understanding the genetic structure of human population (Harrison and Boyce, 1972; Basu, 1995). Let us have a glance on some of the works carried out in this respect
among the populations of Northeast India, taking into consideration the demographic-genetic structure of a population.

In his book entitled *Microevolution*, Das (1981) has described the micro-variation in the Boro, Chutia and Khasi populations. He has observed that each of these populations is divided into different sub-populations/subgroups, which are different from one another in respect of anthropometric, somatoscopic, dermatoglyphic and genetic traits. Such differences within a given population have also been observed in other populations like the Brahmin (Das *et al.*, 1986a) and Kalita (Das *et al.*, 1986b). All these studies have revealed that scholars in this part of the country have also made an attempt to understand how the genes are maintained and regulated within a population. Of course most of the works have been carried out in the populations of Assam with stray researches in other populations of the different states in the region. It may also be worthwhile to mention that the classification of these populations has been based mostly on the frequencies of certain genes and anthropometric traits in view of the geographical location, linguistic affinity, or ethnohistoric background of the population concerned. For example, the Garo, Rabha and Kachari are known as Boro mainly because of their linguistic affinity (i.e. since they speak the Boro language of the Tibeto-Burman group). On the other hand, the Khasi population consists of five major sub-divisions, namely, Khynriam, Pnar, Bhoi, War and Lyngngam. The question of how the Khasi population is known by different names is not fully understood, though it is likely that these five groups are known according to the names of their geographical locations (Khongsdier, 1996). Das (1981, 1984) has suggested the importance of both hybridization and geographical isolation in bringing about the differences between the Khasi sub-groups. He is of the opinion that the Bhoi, who show the greatest deviation from the other Khasi sub-groups, inhabit in a lower attitude area in the northern part of Meghalaya towards Assam. Therefore, intermarriage with the other groups like Khynriam, who are living in the higher altitude, is infrequent. Instead, there is a possibility of gene flow to the Bhoi from other neighbouring populations in Assam. This sort of speculation has also been given in connection with the micro-variation in anthropological traits within the Boro, Chutia, Brahmin and Kalita (Das, 1981, 1984; Das *et al.*, 1986a, 1986b). So it appears that studies in Northeast India have taken anthropometric, dermatoglyphic and genetic
traits along with geographical and socio-cultural (including linguistic) factors with a view to understanding the hierarchy of Mendelian populations.

Recently, demographic data have also been taken into consideration to define the hierarchy of Mendelian populations. Demographic data on marital distance, i.e. the distance between the birth places of spouses, and village endogamy are believed be very important in making out the boundary of endogamous groups and the extent of gene flow into the local populations. A bio-demographic study among the War Khasi has revealed that the Khasi population as a whole is not only divided into four or five sub-groups. Instead, each sub-group, like the War Khasi, is again subdivided into several endogamous units comprising a village, or a group of few villages (Khongsdier, 1994, Khongsdier and Ghosh, 1994, 1996). It is observed that among the War Khasi there is a very high tendency to village endogamy with low admixture rate and marital distance. Accordingly, it is suggested that each village, or a number of few villages, is likely to form a separate deme, which is different from one another in respect of anthropometric and genetic traits. The findings on anthropometric characters seem to confirm such a hypothesis (Khongsdier, 1997). It is also suggested that village endogamy is largely responsible for the active operation of natural selection and genetic drift. It may be noted that the rate of village endogamy seems to vary from one population to another in this part of the country (Barua, 1986, 1993). Interestingly, among the Semsa, a sub-group of the Dimasa in North Cachar hills of Assam, the rate of village endogamy is 100 % (Limbu and Khongsdier, 2000).

Besides the study of the hierarchy of Mendelian populations, an attempt has also been made to show how demographic variables like fertility, mortality, population size, mating, etc. are indispensable for understanding the evolutionary mechanisms that are operating in human populations of Northeast India. An overview of the works done in this respect may be summarised under the following headings:

**Natural Selection:** Natural selection is one of the important evolutionary forces, which brings about changes in gene frequencies of a population from generation to generation. It occurs when individuals of the different genotypes in a population are different from one another in their fitness known as *Darwinian fitness, or genetic fitness.* Darwinian
fitness is defined as the “reproductive capability of an individual or class of individuals, in terms of the number of offspring they contributed to the next generation” (Johnston, 1973). Thus differential fertility and mortality are the fundamental events of natural selection. From the demographic-genetic point of view, “differences in rates of reaching maturity, mating, fecundity, fertility, mortality and emigration are the raw materials of natural selections” (Spuhler, 1973).

Natural selection is believed to operate at four different levels: (i) Total or individual selection, which is measured through differential fertility and mortality (Crow, 1958), assuming that some phenotypic variation in reproduction has a genetic basis and fitness is heritable; (ii) Phenotypic selection, which is concerned with the selective differential of the optimum set of phenotypes in relation to the overall fitness, e.g. birth weight and survival to 28 days after birth (Haldane, 1954); (iii) Genotypic selection, which is concerned with the selective differentials of certain genetic markers, e.g. selective advantage of HbS over HbA; and (iv) Genic selection, which is concerned with the selective differential at molecular level (Tanaka and Nei, 1989).

The total selection, which is believed to measure the maximum opportunity for the changes in the genetic composition of a population, has been widely studied in several populations of Northeast India. Many scholars have followed the Crow’s (1958) formula, and some others have also taken into consideration its modified version (Johnston and Kensinger, 1971). According to reviews (Sengupta and Gogoi, 1995; Sengupta and Kalita, 1996), the Index of opportunity (I), according to Crow’s formula varies between 0.1070 for the Punjabi Sonar of Shillong and 1.0700 for the Gallong of Arunachal Pradesh. It is observed that, in many populations of Northeast India, the mortality component due to selection contributes more towards the Index of opportunity for selection.

Reddy and Chopra (1990) have reported that the mean value of the ‘Index of opportunity for selection’, according to Crow’s formula, for 96 Indian populations is 0.665 with a standard deviation of 0.316. Considering these figures, the population mean was estimated as lying between 0.600 and 0.730, following the 95% confidence interval suggested by Snedecor and Cochran (1967). Accordingly, Khongsdier (2000) has suggested that the different degrees of the total opportunity for selection for the Indian
populations may be arbitrarily classified as shown in Table 2.1. It may be noted that when there is no change in the genetic composition of the population, the value of I is zero (Livingstone and Spuhler, 1965).

**Table 2.1. Degree of opportunity for selection.**

<table>
<thead>
<tr>
<th>Degree of intensity</th>
<th>Crow’s index of opportunity for selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Below 0.340</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.340 - 0.470</td>
</tr>
<tr>
<td>Mild</td>
<td>0.470 - 0.600</td>
</tr>
<tr>
<td>Average</td>
<td>0.600 - 0.730</td>
</tr>
<tr>
<td>High</td>
<td>0.730 - 0.860</td>
</tr>
<tr>
<td>Very high</td>
<td>Above 0.860</td>
</tr>
</tbody>
</table>

Source: Khongsdier (2000)

Following the above classification, the intensity of natural selection is observed to be very low in populations like the Ahom (0.2180), Kachari (0.2500) and Khamti (0.3120) of Assam (see review Sengupta and Gogoi, 1995). Natural selection operates with moderate to average intensity in Sonowal (0.3640) and Kaibarta (0.3360) of Assam, Pnar (0.4012), Christian War Khasi (0.3592), Non-Christian War Khasi (0.4463), Semsa (0.6165) and Hajong (0.6310) of Meghalaya. It is likely that natural selection plays an important role in regulating the genetic composition of the populations of Arunachal Pradesh like the Apatani (0.8890), Gallong (0.1.070) and Khamti (0.9340).

**Genetic Drift:** Another important factor, which brings about changes in the genetic composition of a population, is genetic drift. Brooks (1899) first gave the idea about genetic drift. Then, it was systematically developed mostly by Wright since 1921. That’s why, it is often referred to as the *Sewall Wright effect*. Genetic drift is a random
fluctuation in gene frequencies in a population from one generation to another. It is very effective in small populations because of the greater random sampling error in such populations. Accordingly, one of the basic assumptions of the Hardy-Weinberg law is that a population should be large. In fact, genetic drift largely depends on the effective population size, which is a measure of the actual numbers of breeding individuals in a population (Wright, 1938, 1940; Nei, 1965; Crow and Kimura, 1970; Cavalli-Sforza and Bodmer, 1971). Since the gene pool of each generation represents a sample drawn from the previous generation, the smaller the population the greater the fluctuations will be. Thus, the allele frequencies of the new generation may not be totally representative of the parental population in a small population. For example, if the frequency of allele d is q in a parental population, the probability that q should take a particular value in the next generation is given by

\[
\frac{2N}{K} (q)^k (1-q)^{2N-k} \]

Where \( N \) = total number of individuals, \( k \) = expected number of alleles, and \( q \) = allele frequency in the parental population. Suppose, the frequency of q in a population of 5 diploid individuals is 0.5, the probability that the same frequency of q (0.5) will occur in the next generation is 24.61 %, whereas in the case of a population with only 2 individuals, it is about 37.50 %.

For simplicity, let us consider the mating between two heterozygotes, i.e. individuals who carry 50 % of allele D and 50 % of allele d. This type of mating would produce three types of genotypes with probabilities: \( \frac{1}{4} \) DD, \( \frac{1}{2} \) Dd, \( \frac{1}{4} \) dd. Substituting the above formula, we get,

\[
\frac{N}{D} (\frac{1}{4})^D (\frac{1}{4})^d (\frac{1}{2})^x \]

D / d / x
Where \( N \) = total number of individuals, \( D \) = expected number of individuals with DD genotypes, \( d \) = expected number of individuals with dd genotypes, and \( x = N - (D + d) \), i.e. expected number of individuals with Dd genotypes. Assuming these two parents have only two children, the genetic constitution, or gene pool of the next generation would be either one of the six combinations of genotypes \((3 \times 2)\), that is, if there are three children, it would be \(3 \times 3\) and so on. It is seen from Table 2.2 that the probability that both the children would be DD genotypes is 6.25 %. Similarly, the probability that the two children of dd genotypes would form the genetic composition of the next generation is 6.25 %. Consequently, in the absence of mutation, selection and migration, either one of the alleles would be lost or fixed in small population due to random sampling process. As a result, the fate of small population is either extinction or fixation of the advantageous allele.

**Table 2.2.** Probabilities of the two offspring genotypes in the mating between heterozygotes.

<table>
<thead>
<tr>
<th>Offspring genotype</th>
<th>Allele frequency</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( D )</td>
<td>( d )</td>
</tr>
<tr>
<td>DD, DD</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>DD, Dd</td>
<td>0.75</td>
<td>0.25</td>
</tr>
<tr>
<td>DD, dd</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Dd, Dd</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Dd, dd</td>
<td>0.25</td>
<td>0.75</td>
</tr>
<tr>
<td>dd, dd</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Thus, it is believed that in small population the role of genetic drift is more important than that of mutation and selection, i.e. an allele may be lost or fixed with little reference to selection and mutation pressures. In the case of the population shown in Table 2.2, there is a chance of 6.25 % that an allele d will be lost or fixed, though such an allele is favoured by the natural selection. So it means that the intensity of natural
selection will become either zero or 100% in such a population. However, selection is more effective in large population than the genetic drift. Suppose, for example, the selection coefficient (s) against the allele d is 0.001. In a population with the effective population (N_e) of 100 individual, the product of N_e and s is 100 x 0.001 = 0.1, whereas in the population with the N_e of 10,000, the product is 10. Therefore, the changes in allele frequencies are largely due to genetic drift in small isolated populations, but the situation is just reverse in large populations where selection is more important. Similar phenomenon is in the case of mutation. According to Dobzhansky (1970), the mutation and selection rates may be regarded as small if the product 4Nμ and 4Ns is less than unity. In fact, the role of genetic drift in regulating the genetic composition of a population is enhanced by the neutral theory of protein evolution postulated by Kimura (1968, 1983) and others. The exponents of this theory have proposed that most of the polymorphisms found in natural populations are neither useful nor harmful to their carriers, but simply neutral so that natural selection has little role to play. The frequency of such neutral mutants in populations largely depends on chance and random sampling.

In Northeast India, there are hardly studies, which are concerned with the effect of genetic drift on the genetic composition of the population. The findings among the War Khasi (Khongsdier and Ghosh, 1994, 1996) and the Semsa (Limbu and Khongsdier, 2000) indicate that genetic drift plays a very important role in regulating the genetic make up of these populations. It is likely that there are still several small and isolated populations in different states of Northeast India where genetic drift plays an active role. Thus, it may be necessary to carry out thorough studies in this field with a view to having a better understanding of the evolutionary effectiveness of genetic drift in natural population.

**Gene Flow:** According to Johnston (1973), “The process by which genetic variation is introduced into a population is called gene flow, migration, or admixture”. It occurs when genes from outside are introduced in the gene pool of a native population, or when a hybrid population is formed owing to the admixture of the gene pools of two or more populations. It is expressed as $m = (q_a - q_n)/(q_a - q_b)$, where $m$ stands for the admixture rate, $q$ is the allele frequency in population $n$, and $q_a$ and $q_b$ are the frequencies of the
same allele in the populations \( a \) and \( b \), respectively. In Northeast India, no report has been published in this respect, though it is always mentioned that gene flow is very important factor for regulating the gene frequencies in the populations. An attempt has, however, been made in some studies (Barua, 1986; Khongsdier, 1994, Khongsdier and Ghosh, 1994, 1996) to estimate the admixture rate on the basis of the number of gamates introduced from outside into the native population. These studies have revealed that the admixture rate in populations of Northeast India varies from zero per cent onwards.

Thus it is obvious from the present review that many studies have been carried out in Northeast India to find out the genetic variation between and within populations, but there are hardly any studies, which deal with the causes of such genetic variation. The speculation that the micro-genetic variation within a given population is due to either mating patterns or geographical isolation is always possible, but hardly meaningful without empirical evidence. Therefore, it warrants further in-depth studies with a view to understanding the causes of genetic variation between and within populations in Northeast India.