1. INTRODUCTION

The fields of applied cytogenetics and molecular genetics have helped in providing sufficient scope for understanding the characteristics of animal genomes. With the knowledge of the chromosomes, the subject of cytogenetics developed as an offshoot of genetics, therefore cytogeneticists remained devoted to the principles of heredity in relation to the chromosomes. This subject proved quite useful for differentiating species among the sympatric taxa and chromosomal forms in some of the important groups of insects including mosquitoes. Mosquitoes of the subfamilies Culicinae and Anophelinae of the family Culicidae include some of the major vectors of diseases like malaria, filariasis, yellow fever, dengue and dengue hemorrhagic fever in human populations all over the world. Many species of anopheline mosquitoes are the exclusive vectors of human malaria in which some of the species dominate as the serious urban and rural vectors (Coetzee et al., 2000; Van Bortel et al., 2001). In relevance to this, a new paradigm to manage vector transmitted diseases was offered by the developments in 80’s and early 90’s involving biochemical and cytogenetic techniques so as to generate information about the genetic basis of disease transmission. (Morris et al., 1989; Kidwell and Rebeiro, 1992) Developments in molecular cytogenetics apparently offered promising possibilities for extension and future applications of genetic engineering based
control programmes (Besansky et al., 1992; Carlson, 1996; Rai, 1996, 1999). In fact, the biological world which is governed by constant evolutionary processes, there are many observable variations in the same animal which often occur in a range of variations at the extremes of which two distinct forms may be recognized whether in shape, size, colour, measurement of body parts and sometimes in behaviour. It is interesting to explore as to how one species differs from the other or even the same species in different sets of environments. In large number of animals including mosquitoes, which reproduce sexually, a species is a genetically distinct but reproductively isolated entity. In reference to these facts of population genetics and the past achievements in chromosomal cytology, studies on mosquitoes revealed that many species groups had subspecies with indistinct or poorly defined morphotaxonomic characters but very clear differences in their polytenic chromosomal banding pattern. In other words, there were complexes of several species, which were morphologically similar but chromosomally different. In fact, within the genus *Anopheles*, identification is sometimes difficult due to the prevalence of biotypes, ecotypes and cytotypes etc., which are difficult to differentiate on the basis of morphological features alone. The phenomenon of species complexes further complicated their identification because a single species often exhibit significant heterogeneity across a broad geographic range (Anyanwu et al., 1997). This additional phenotypic and genotypic plasticity exacerbates the difficulties in identification of vector populations and implementation of effective surveillance and control strategies aimed at the target species (Foley et al., 1996). For example, in mosquitoes, cytogenetic studies remained one of the most important means of differentiating all members of the *Anopheles gambiae* complex in which the intra-specific variants could be identified as different chromosomal forms (Coluzzi et al., 1979,
1985). In the same way, one could reliably differentiate all the nine members of the *Anopheles funestus* group (Green and Hunt, 1980). With the discovery of giant pollytene chromosomes in mosquitoes, much of the success was achieved by studying the pollytene chromosomes which proved to be ideal tools for understanding a number of genetic principles operating in them.

Since, well organized and well banded pollytene chromosomes can only be obtained from the salivary gland cells of fourth instar larvae and ovarian nurse cells of adults, the identification of adults usually involved the laborious process of rearing progeny broods from wild caught females and immediate study and interpretation of results from the temporary preparations. Inspite of these difficulties, the detailed study of the species specific banding pattern of these chromosomes was successfully utilized in solving the problems of speciation and taxonomy in the genus *Anopheles*. This was achieved by constructing the pollytene chromosome standard maps of a large number of vector and non-vector species. These maps have now become more useful in the pollytene chromosome based molecular level researches because they act as the reference standards for locating various genetic loci on the chromosomes (Kitzmiller, 1967; Coluzzi and Sabatini, 1967; Coluzzi et al., 1970; Kreutzer et al., 1970; Chaudhry, 1971, 1979, 1981, 1999, 2003; Rabbani and Kitzmiller, 1974; Green and Miles, 1980; Green et al., 1985a, b; Mehmood and Sakai, 1985). Although these procedures were time consuming and requiring considerable expertise, yet they also helped in identifying as many as 30 species complexes or groups of more than one cryptic species (Coluzzi and Sabatini, 1967; Collins et al., 1987; Wesson et al., 1992; Scott et al., 1993; Cornel et al., 1996; Porter and Collins, 1996). These investigations lead to the understanding of population genetics, comparative cytogenetics and karyosystematics of several members belonging to family
Introduction

Culicidae (Coluzzi and Kitzmiller, 1975; Kitzmiller, 1976; Green, 1982; Steiner et al., 1982; Rai, 1999; Walton et al., 1999; Chaudhry et al., 2005). During the last four decades, the work on mosquito cytogenetics has grown to such an extent that it has actually made valuable contributions in understanding their basic biology, evolution, speciation and genetics of vectorial capacity and insecticide resistance (Bullini and Coluzzi, 1982; WHO, 1984; Subbarao, 1996; WHO/UNDP/World Bank, 2003; Chaudhry et al., 2005). As a result of comparative cytogenetic studies, polytene chromosome information is now available for as many as 100 species of the genus *Anopheles* alone (Kitzmiller, 1973, 1976; Coluzzi and Kitzmiller, 1975; Steiner et al., 1982; Subbarao and Sharma, 1984; Pape, 1992; Chaudhry, 1999). Although chromosomal investigations formed the initial subject of research, yet many more parameters of study were also used to supplement the data which included isozyme studies and amino acid estimations (Van Bortel et al., 1999; Sukowati et al., 1999; Krzywinski and Besansky, 2003). Alternatively, allozymes could also be used to differentiate the species (Green et al., 1992), through electrophoresis of fresh specimens or those stored in liquid nitrogen. However, non-morphological features such as cytological or isoenzyme characteristics have practical disadvantages that preclude their more frequent applications. As compared to these limitations, the DNA based methods have the advantage that all developmental stages can be used and the DNA samples can be prepared from individuals preserved in ethanol. For example, DNA probes have been used to distinguish all the four species in *An. dirus* complex in Thailand (Panyim et al., 1988; Audtho et al., 1995). In order to identify cryptic species, many techniques have been used which evolved as a result of the technological advancement in the instruments and tissue processing. In fact, once the genetic relationship of mosquitoes in
the transmission of diseases and vectorial capacity was established, taxonomists sought to identify, describe, classify and determine the diversity and phylogenetic relationships among the various members of the species belonging to different genera. It was realized that species complexes are comprised of two or more isomorphic sibling species, which may exhibit different ecological, behavioural and physiological characteristics, which ultimately determine their vector potential. Therefore, modern systematics need the ecological, morphological, physiological and biochemical data to characterize such species. Some of these include studies of egg morphology (Hinton, 1968; Linley et al., 1993; Chaudhry and Gupta, 2003, 2004), larval chaetotaxy (Coetzee, 1984), salivary and ovarian polytene chromosome banding patterns (Kitzmiller et al., 1967; Subbarao et al., 1988; Subbarao, 1996; Coetzee et al., 1999, Ramirez and Dessen, 2000a, b), zymotaxonomy (Korvenkontio et al., 1979; Bullini and Coluzzi, 1982), cuticular hydrocarbon estimations (Phillips et al., 1990) and most recent approaches of DNA characterization and gene sequencing through the technique of polymerase chain reaction (PCR) (Marinucci et al., 1999; Proft et al., 1999; Wilkerson et al., 1995; Chaudhry et al., 2004). The techniques involving microsatellite DNA and mitochondrial DNA provided some of the most accurate results of the sequence data of interest (O’Brein et al., 1991; Bernatchez and Danzmann, 1993). In fact, the reason for adopting PCR based procedures for genomic characterization lies in the fact that only tiny amount of starting material is necessary, making it possible to sequence the DNA of very small organisms and even the degraded samples.

In the light of the past considerations related to mosquitoes, a unique system of identification of species has been recommended by various workers (Narang et al., 1993; Beebe and Saul, 1995; Reinert et al., 1997; Chaudhry et al., 2004). This involves more
than one parameters for differentiation of species and their populations (Fig. 1). The PCR technique requires the utilization of highly stable DNA polymerase in a thermal cycler for an *in vitro* enzymatic amplification of a DNA template. As conceived by Mullis *et al.* (1986) and later developed by Williams *et al.* (1990), it is a hightech method for *in vitro* production of unlimited copies of the desired fragments of DNA. It involves the multiplication of DNA sequences using forward and reverse oligonucleotide primers that anneal to opposite strands and flank the desired region in the target area. Thus the applications of this technique have been extended to molecular diagnostics, forensic investigations, clinical diagnosis, familial relationships and environmental surveillance. In addition to the use of genomic DNA, the amplification of microsatellite markers, mitochondrial DNA (mt. DNA) and rDNA spacers have proved quite reliable for detecting micro and macro geographic variations in species and for drawing phylogenetic relationships among sympatric and allopatric species. (Munstermann, 1995; Cornel *et al.*, 1996; Wang *et al.*, 1999; Walton *et al.*, 1999; Chaudhry *et al.*, 2005). The process of choosing a region of DNA that is likely to be appropriate for the recognition of species is the most critical step in any phylogenetic analysis. For a successful phylogenetic reconstruction, it is necessary that a gene chosen must be the one which allows appropriate balance between molecular conservation and variability because the rate of evolution of a gene and its functional role are closely related. However, this relationship is variable because a single gene evolves at different rates in different species. Therefore, there is a close relationship between gene component and rate of nucleotide substitution which results in a single gene being useful at different phylogenetic levels. Compared to nuclear rDNA, it is more difficult to design universal
Fig. 1: Multiple Technique Identification Approach (Ecological, Morphological, Cytological, Genetic, Biochemical and Molecular) (Narang et al., 1993; Reinert et al., 1997; Chaudhry et al., 2006). An overview of progress in genomics of mosquitoes from 1910 till date.
primers for amplifying specific regions in mt. DNA due to high variability in its nucleotide component. Primers can either be specific to a particular DNA nucleotide sequence or they can be "universal." Universal primers are complementary to nucleotide sequences which are very common in a particular set of DNA molecules. Thus, they are able to bind to a wide variety of DNA templates. The high degree of conservation of a particular gene region makes it possible to construct universal primers for that region. In addition, universal primers for amplifying certain genes were also designed by Lunt et al. (1996) and Zhang and Hewitt, (1996). As universal primers or primers specific to a certain taxon are continuously developed, DNA sequence data from the mt. DNA are being used with increasing frequency to estimate phylogenetic relationships among animal taxa.

Out of the different sources of DNA, the mitochondrial DNA is frequently utilized in phylogenetic and population genetic studies. In fact, analysis of the parts of mt. DNA and direct sequencing of specific regions of the mitochondrial genome are currently the methods of choice for majority of the population level studies in the genera Anopheles, Culex, Aedes and Armigeres (Hwang and Kim, 1999; Chaudhry et al., 2006; Kohli and Chaudhry, 2007). The small size of mitochondrial genome, lack of introns and recombinations, and maternal inheritance, are some of its features for which it is preferred for DNA characterization. Further, it is known to evolve much faster than the nuclear genome. Consequently, most of the mitochondrial protein coding genes have been used to examine phylogenetic relationships. The coding regions in mt. DNA have been extensively used for displaying phylogenetic relationships as they are single copy, clonally inherited, nonrecombining and easy to manipulate (Moritz et al., 1987). In fact, in actual practice mt. DNA is used to examine
DNA from those samples which cannot be analysed by the techniques of restriction fragment length polymorphism (RFLP) and analysis of short tandem repeats (STR). This is because the nuclear DNA must be extracted from samples for use in PCR, RFLP and STR. To the contrary mt. DNA analysis uses DNA extracted from the mitochondria. Sometimes the materials like hair bones and teeth from the older specimens lack nucleated cellular material, therefore they cannot be analyzed by STR and RFLP but can only be analyzed with the amplification and sequencing of their mt. DNA. The logic of using both mt. DNA and nuclear DNA in genome diagnostics of insects in the sexually reproducing individuals such as mosquitoes lies in the fact that the females have the mt. DNA inherited from female parents only while nuclear DNA is a combination of the two at the time of fertilization. Therefore, comparison of mt. DNA and nuclear DNA profiles have been considered more useful as it provides an opportunity to trace the phylogenetic history of the species. In the earlier chromosomal based studies, Y-chromosome analysis was always considered ideal genetic marker for tracing genetic relationships among the males of the species; therefore, nuclear DNA has its own advantages in genome analysis. For example, nuclear genes have the potential of recombination due to natural modes of deletions and insertions of sequences resulting in the removal of linearity from the phylogenetic trees. Lately, it has been suggested to use more than one cytogenetic marker in establishing the extent of genetic kinship among the species. For example, Besansky et al. (1994) compared four different DNA based markers to establish the phylogeny of the members of *gambiae* complex. These markers included the chromosome inversions, mitochondrial DNA, NADH dehydrogenase IV and NADH dehydrogenase V (ND4 and ND5) genes, rDNA Intergeneric
Spacers (IGS) and an esterase gene. They found that chromosome inversions and DNA sequences when read together were of considerable utility in phylogeny construction. In addition, mt. DNA genes being AT rich, increase the potential for multiple site substitutions. For the studies on relationships among closely related species, cytochrome b (cyt. b), cytochrome oxidase I (CO I), cytochrome oxidase II (CO II), NADH dehydrogenase I (ND I), NADH dehydrogenase II (ND II), NADH dehydrogenase V (ND V) and nuclear internal transcribed spacers 1 and 2 ITS1, ITS2 are some of the convenient sources of genomic diagnosis. Unlike, the nuclear genome, even a small segment of mt. DNA can provide accurate pictures of sequence variations for use in the construction of phylogenetic trees or dendrograms of species and for understanding the genetic structure of their molecular evolution (Avise et al., 1987).

In addition to these properties, the mt. DNA is also unique in having circular DNA molecule which is made up of only 37 genes coding for twenty two molecules of tRNA, two of rRNA and thirteen of mRNA. It consists of the following major coding regions for which the genes lie in a sequence as: ND 2, CO I, CO II, ATP, C3, N3, ND5, ND4, ND6, cyt b, 16S, 12S and A+T regions (Clary and Wolstenholme, 1985). Cytochrome c oxidase is an important respiratory enzyme that is involved in the terminal oxidation steps of the mitochondrial electron transport chain. Mt. DNA encodes three subunits which are synthesized in the mitochondrial ribosomes of vertebrates and invertebrates (Capaldi et al., 1983; Clary and Wolstenholme, 1983, 1985). Out of the three subunits of mt. DNA, subunit II has been studied extensively in vertebrates (Busse et al., 1978; Bisson et al., 1980; Brown and Simpson, 1982; Millet et al., 1983; Liu and Beckenback, 1992). Subunit II is not only a high affinity binding
Introduction

site for cytochrome c, but also contains ligands for one copper atom in cytochrome c oxidase. In addition to this, COII shows highly divergent sequence in which the proportion of silent nucleotide substitutions in closely related species is extraordinarily high (Brown and Simpson, 1982). However, higher incidence of divergence of both amino acid and nucleotide sequences has been shown in COII gene in orthopterous and hymenopterous insects (Liu and Beckenback, 1992). Foley et al. (1998) found 684bp mt. COII sequence to be useful for resolving Australasian anophelines at the species level only and not at the levels of subgenera and series. Therefore, the evolutionary relationships based on the sequences of COII gene were found to be quite useful.

Depending upon the specific motives of research workers, the use of intergenic spacers (IGS) and internal transcribed spacers (ITS1 and 2) within the nuclear ribosomal genome have also been found to be quite convenient in generating molecular data of taxonomic value. The nucleotide sequence of these spacer regions are often much more polymorphic between species than within species (Morgen and Blair, 1998; Navajas et al., 1998; Marrelli et al., 1999; Hackett et al., 2000).

The limitations of morphological taxonomy for resolving species boundaries advocated the multidisciplinary approaches to mosquito systematics because it is still not known whether there are any constraints in vectorial capacity between different Anopheles species and between different members of sibling species (Faran, 1979a). For example, Walton et al. (1999) utilized allele specific PCR amplification of rDNA ITS2 to detect differences in the base sequence of this region in An. dirus complex. According to them, it was not possible to find differences between all members of the An. dirus complex on the basis of their
morphology alone, therefore sequence analysis of ITS2 was successfully used to differentiate them.

From the foregoing introductory information on the progress of genomic studies in mosquitoes, it is evident that a variety of molecular parameters have been used for describing the variations and homologies in the genomic qualities of the species. The variety and number of genetic markers available for the study of malaria vectors has literally revolutionized the genomic analysis of mosquitoes. These molecular markers have proved useful in a wide variety of applications including molecular taxonomy, evolutionary systematics, population genetics, genetic mapping and related topics in molecular diagnostics. There is a rapid growth in the database of insect sequences, therefore most of the molecular level systematic studies consider DNA sequencing studies the best method of choice. This is due to the fact that by using suitable DNA probes, the sequence analysis has been found to be most appropriate for studies aimed at intra and interspecific levels of genetic relatedness as it provides the desired results with amazing accuracy (Avise, 1994; Hillis et al., 1996). The knowledge of particular genetic loci with the help of sequencing has actually helped in developing phylogenetic models that make optimal use of observed variations.

For analyzing the molecular variations in the various regions of the genome, rDNA sequences have long been recognized as suitable markers for phylogenetic analysis and identification of species (Hillis and Dixon, 1991). In addition to the techniques mentioned above, internal transcribed spacers (ITS) are also being applied to detect micro and macrogeographic variations in different species and their populations in the mosquitoes belonging to genera Anopheles, Culex and Aedes for drawing phylogenetic relationships among their sympatric and allopatric
species (Lehmann et al., 1996; Wang et al., 1999). The ITS of rDNA are often treated as typical non-functional spacer sequences in which ITS1 is required for the processing of 3' end of 18S and the 5' end of the 5.8S rRNA coding regions, while ITS2 is required for the processing of the 3'end of the 5.8S and the 5'end of the 28S rRNA molecules (Fig. 2). They are also important sequences for the processing of rRNA secondary structure. As mentioned earlier in the text, in addition to ITS1 and 2 sequences, the intergenic spacers (IGS) have also proved valuable as they contain species-specific sequences which help in species differentiating by both targeted PCR and DNA probes. The size of IGS (4-5kb) is far longer than those of ITS regions (1kb). However, due to the large size of IGS, ITS regions have been preferred in phylogenetic analysis. However, as a part of feasibility studies, variations in the length and copy number of intergenic spacers (IGS) of nuclear ribosomal DNA were examined for genetic differentiation of Anopheles albimanus populations in which the analysis of IGS patterns in individual mosquitoes indicated that populations on the Pacific side of Central America were genetically distinct from those on the Atlantic side or from South America.

<table>
<thead>
<tr>
<th>5' ETS</th>
<th>IGS</th>
<th>18 S</th>
<th>ITS1</th>
<th>5.8S</th>
<th>ITS2</th>
<th>28 S</th>
<th>IGS</th>
<th>ETS 3'</th>
</tr>
</thead>
</table>

Fig. 2: rDNA unit

ETS = External transcribed spacer
IGS = Intergenic spacer
ITS1 = Internal transcribed spacer 1
ITS2 = Internal transcribed spacer 2
DNA probes are single stranded DNA sequences, tagged with a detectable agent such as radiolabelled bases that recognize specific complementary single stranded target DNA sequences in the species of interest and demonstrate their presence by annealing or hybridization to the complementary sequences so as to form a part of the DNA strands in the in vitro amplification of target DNA. The first DNA probes for anophelines were developed for the species of *Anopheles gambiae* complex which is a chief vector of malaria in the entire African continent. As a result of such studies these probes are now available for the identification of all the five cryptic species (A, B, C, D, E) of this complex (Gale and Crampton, 1987), *Anopheles dirus* complex (Panyim *et al.*, 1988), *Anopheles quadriannulatus* complex (Cockburn and Seawright, 1988) and *Anopheles punctulatus* complex (Beebe *et al.*, 1994, 1996). It has also been noticed that higher the DNA sequence repeat copy number in the genome, the more sensitive the probe and less amount of mosquito DNA required for the identification of the DNA bands generated through PCR.

The cytogenetic studies carried out so far have established the fact that the size of the total genome varies among different groups of organisms. This aspect of the nuclear genome can provide considerable information of cytogenetic interest. Due to their abundance within a genome and the convenience of amplification and sequencing, nuclear genes have been widely used by various workers. For example, the nuclear ribosomal sequence repeats comprising internal transcribed spacers are some of the highly conserved and variable regions found valuable for resolving relationships at higher taxonomic levels. These genes are considered to be highly informative because the functional regions that produce the ribosomes are highly conserved (Hillis and Dixon, 1991; Miller *et al.*, 1997). The results achieved so far in
mosquitoes have shown that ITS1 and 2 and non-transcribed spacers (IGS) have high interspecific and low intraspecific variability which make them useful for detecting relationships of recently diverged species. In addition to these qualities, the non-coding and intergenic regions show more variation than coding sequences, therefore sequences of introns have proved to be quite useful for phylogenetic studies. This is because of the fact that the regions flanking intron sequences are evolutionary more conserved and hence ideal markers for annealing of PCR primers. For example, if the sequence of 5' or 3' non-coding region of a gene is known then specific primers can be conveniently designed for the amplification of these regions.

In addition, the nuclear rDNA spacer regions have evolved much faster than nuclear rDNA coding regions because the substitutions occurring in spacer regions do not show lethal effects on the organisms. These regions have been found to undergo neutral mutations without any genetic damage. This resulted in a relatively rapid rate of genomic evolution of spacer regions. As compared to them, the rDNA coding regions yield mutations of encoded rRNA sequences. Any harmful mutations in them can prevent the successful ribosome construction which can have negative affect on the protein synthesis machinery. Therefore, rRNA coding regions are more conserved as compared to the spacer regions. Because of the different rates of evolution among different regions of nuclear rDNA, the nucleotide sequences of nuclear rDNA have been used to infer a broad spectrum phylogenetic relationship among closely related species of mosquitoes and their populations. Due to the high variability, rDNA spacer regions have also been utilized for resolving phylogenetic problems in a number of other insect genera (Morgen and Blair, 1998; Navajas et al., 1998; Pereira et al., 1998). In the application of PCR based species discrimination random amplified
polymorphic DNA (RAPD) was the first technique to be used for molecular level systematics and genetic studies of populations as it required no prior sequence information and could reveal accurate nucleotide polymorphism. This procedure involves the randomly selected primer sequences obtained through trial and error (Sucharit and Komalamisra, 1997). As compared to RAPD, the technique of arbitrary fragment length polymorphism (AFLP) is considered as a dominant marker as it generates a number of desired fingerprints of the genome. This technique utilizes PCR amplification of restriction enzyme generated fragments of DNA by digestion with a suitable restriction enzyme (Vos et al., 1995). The AFLP markers offer several advantages over other currently used DNA markers such as the detection of simple sequence repeats (SSR) and single nucleotide polymorphisms (SNP) in populations. Just like RAPD-PCR, AFLP technique requires no prior sequence information and hence, has a relatively low start-up cost.

Depending upon the motive of the genetic analysis of populations, the microsatellite DNA has also become a popular marker for genetic studies of anophelines as it contains simple tandem repetitive (STR) sequences. Microsatellites have proved to be useful for analysing population genetics and genetic mapping of species. They show high mutation rate which ensures high level of polymorphism useful for examining relationships among individuals of freely interbreeding populations (Schlotterer et al., 1998). Mosquito microsatellites have been studied in the members of the *An. gambiae* complex, in which both micro and macrogeographic variations were found which revealed extensive gene flow between neighbouring populations in slightly different habitats and restricted gene flow between distantly placed geographic populations (Lehmann et al., 1996).
The PCR products of ITS combined with restriction site analysis (RFLP) provide an ideal protocol for species diagnostics. In the same way, ITS2 sequences of *Anopheles* species are also preferred as ideal molecular markers because of their two important features. First, they are relatively short, less than 1 kilo base pairs in length, making the amplification convenient using primers from highly conserved domains of flanking coding regions. Second, the levels of intraspecific variations in them are lower than the interspecific variations. For example, Hackett *et al.* (2000) distinguished *An. (Cellia) funestus* Giles from *An. (Cellia) rivulorum* Leeson and discovered a cryptic taxon within the *funestus* group on the basis of sequence divergence in the ITS2. Similarly, Manonmani *et al.* (2001) sequenced the ITS2 fragment of *An. fluviatilis* collected from the Koraput and Malkangiri districts of Orissa (India) while Naddaf *et al.* (2003) determined the composition and distribution of the members of same species complex from the Islamic Republic of Iran using ITS sequence alignment.

In the present research, Spectral Repeat Finder (SRF) method is used to locate and identify repetitive DNA and its constituent units. Spectral repeat finder is a program to find repeats through an analysis of the power spectrum of a given DNA sequence which means the repeated occurrence of a segment of N nucleotides within a DNA sequence. The repeats can be contiguous, in which case they are termed tandem repeats, or not, in which case they are dispersed. SRF is an *ab initio* technique as no prior assumptions need to be made regarding either the repeat length, its fidelity, or whether the repeats are in tandem or not. In order to use the program, a sequence is submitted in various specified formats. There are two major groups of repeats in eukaryotic genomes: tandem repeats that are usually confined to...
specific chromosomal regions and repeats interspersed with genomic DNA mainly represented by inactive (pseudogenes) copies of historically or contemporarily active transposable elements (Strachan and Read, 1999). Tandem repeats are grouped into three major subclasses: satellites, minisatellites and microsatellites (Strachan and Read, 1999). Satellite repeats are composed of very long tandem arrays of 5-171 bp repeat units usually present at centromeres. Minisatellites consist of tandem repeats of short units of about 6-64 bp located near the telomeres, while microsatellite repeats are highly repetitive sequences consisting of 1-4 bp repeated up to 50 times as tandem arrays dispersed throughout all the chromosomes. Likewise, interspersed repeats can also be subgrouped into five types: Short Interspersed Nuclear Elements (SINEs), Long Interspersed Nuclear Elements (LINEs), Long Terminal Repeats (LTRs), DNA transposons and others (Smit, 1996).

In reference to the advances made in the chromosome cytogenetics and molecular cytogenetics of mosquitoes, the present topic of research entitled "Molecular Cytogenetics of Some Anopheles Mosquitoes (Diptera: Culicidae)" was undertaken to carry out the DNA diagnostics of six epidemiologically important species of the genus Anopheles viz: An. stephensi, An. culicifacies, An. maculatus, An. subpictus, An. annularis and An. splendidus. In order to meet these objectives, the sequence variations of three specific regions of the genome of these species were characterized as they have been identified as potential molecular markers for studying the phylogeny of mosquitoes. These molecular markers include internal transcribed spacer regions 1 and 2 of the ribosomal DNA of the nuclear genome component and COII gene of the mt. DNA. These specific sequences of nuclear and mt. DNA were amplified by the technique of polymerase chain reaction.
(PCR) for which specific primers were used. The PCR amplified products in the form of DNA bands were analyzed for the number of nucleotides after which they were sequenced for reading the variations in the base pair composition from the sequence alignment data. From these sequences, it was possible to calculate the GC:AT content, incidence of dimers, trimers, tetramers and polymers etc., ratio of transitions and transversions and tandem repeats by making use of suitable software programs such as CLUSTAL W, PHYLIP, GENERUNR and SPECTRAL REPEAT FINDER (SRF). In case of mitochondrial gene sequences analysis of gene (subunit) COII, the multiple sequence alignment of nucleotide sequence based amino acid (AA) sequences were also deduced by making use of the insect mitochondrial codons in GENERUNR software after referring to these AA sequences in those species of genus *Anopheles* in which such an exercise has been carried out. The motive of this parameter of study was to identify the relationship between a particular set of nucleotides coding for an amino acid. Because it is ultimately the synthesis of proteins which influence the physiology of an insect acting as a vector or a non-vector in the spread of mosquito borne diseases. At the same time, such amino acid coding variations also help in identifying the sequence variations in different species of mosquitoes in relevance to their vectorial behaviour. In the final assessment of the sequence variations, phylogenetic relationships among the selected six species were drawn by constructing the dendrograms of genetic relatedness. With this, the phylogenetic position of these six species in relation to other species of the subgenus *Cellia* were also inferred.