(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

MULTILOCUS GENOTYPING FOR DIVERSITY ANALYSIS IN MARATHWADA AND SURTI BUFFALOES (BUBALUS BUBALIS)

Vanshika Goel and Sainyam Goel
Indraprastha Convent Senior Secondary School
Begumpura, Rohini, New Delhi

1. INTRODUCTION

Conservation of indigenous animal genetic resources tops the list of prime livestock development activities for India. The conservation programs, especially followed by farmers are complementary and supplementary to the agricultural systems and have helped in agricultural production and food security in the country. The nomadic and semi-nomadic production system at zero input levels have helped in conservation of many of the indigenous breeds. The situation is far from satisfactory in terms of collection, computerization, collation and utilization of information on AnGR for their conservation. India sub-continent is a rich source of diverse live germplasm and only very few countries have such a large number of breeds with wide genetic diversity. These vast genetic resources are comprised of 10 breeds of buffaloes, 30 of cattle, 43 of sheep, 20 of goats, 7 of camel, 8 of horse, 19 of poultry and few types of pigs, yolks and other distributed in various agro climatic zone of the country.

Among all the animal genetic resources buffalo is a unique bovine which despite a highly contribution to the agrarian economy and providing food security remains by and large neglected. The buffaloes are discontinuously distributed throughout globe. Buffaloes are the richest genetic resource that is found in abundance in India and form important component of the livelihood of rural masses. Total population of buffaloes was about 88 million as per live stock census of 1998. In India the buffalo have emerged as the main dairy animal, and from where it has spread to the other parts of the world. The share of the buffalo of total milk production was around 56 percent (Sethi, 2004). Buffaloes have been classified into two major groups viz. Swamp (*Babulus carabanesis*) and Riverine (*Bubalus bubalis*). They belong to the same genus and have different habitats. Swamp buffaloes are found in South and Southeast Asia. In several East-African countries water buffaloes were introduced from Sri Lanka and India. However, these relatively small buffalo herds were exterminated on account of epidemics and inadequate care. Swamp buffalo is more or less permanent denizen of marshy lands. Swamp buffaloes are stocky animals with marshy land habitats and have 48 (2n) number of chromosomes. Earlier on the basis of phenotypes Nagpuri, Pandharpuri, Toda and Marathwada were classified as swamp type. The swamp buffaloes have been further classified

into breeds. The swamp buffalo is characterized by massive horns growing outwards from the skull in a semi-circle and is mainly used for work. It has great potentiality for meat production and draught production (primarily used), which is now being exploited at an increasing rate while milk production is very limited. The river type buffaloes are found in Indo-Pakistan subcontinent and are primarily bred for milk production while meat production and draught are secondary traits. River buffaloes are generally large in size mostly with curled horns, prefer to enter clear water and have 50 (2n) number of chromosomes. The river buffaloes are found throughout India where clear water of rivers, irrigation canals and ponds is available for wallowring. The river buffaloes are mainly dairy type and Murrah, Nilli-Ravi, Surti, Mehsana, Jaffrabadi and Bhadawari breeds belong to this category. India alone has now the privilege of having about 45 percent of the world's buffalo population. According to an FAO survey, buffaloes contribute about 90% of draft power for agricultural sector in the Southeast Asian countries.

The animal genetic resources have encouraged studies to determine genetic variations within and among breeds. For this purpose various genetic markers are increasingly being utilized. Among the various molecular genetic markers, the highly polymorphic microsatellites have been extensively utilized for various population genetic studies owing to exceptional variability and relative ease of scoring. Microsatellites or short tandem repeat (STR) loci consisting of tandem repeat units of very short (1-6 base pairs) nucleotide motif. In case the nucleotide sequences in the flanking regions of the microsatellite are known, specific primers (generally 20-25 bp) can be designed to amplify the microsatellite by the Polymerase Chain Reaction (PCR). Microsatellite DNA sequences are important in many types of genetic analysis, such as loss of heterozygosity testing, linkage analysis, gene mapping and discovery, paternity and forensic identity testing, evolutionary studies, and clinical diagnostics. These loci are highly polymorphic and are also ubiquitously distributed throughout the genomes (Dib et al. 1996; Dietrich et al. 1996). The microsatellite have been rapidly replaced RFLP and RAPDs in most applications in population biology, from identifying relatives to inferring demographic parameters (Bowcock et al. 1994; Goldstein et al. 1996; Jarne and Lagoda, 1996). STR analysis usually consists of locus-specific PCR amplification followed by electrophoretic separation of the products. The major advantage of microsatellite markers is that it is relatively easy to identify numerous polymorphic markers even between closely related populations. Many STR based genetic studies, such as linkage analysis, gene mapping, and loss of heterozygosity testing, require the screening of large numbers of microsatellite loci in thousands of individuals. To address the needs of these and future STR-based studies, new generations of genetic analysis devices and methods that rapidly perform sensitive, specific, and high-throughput STR analyses are required and are gaining prominence.

Generally complementary approach for estimating adaptive genetic diversity is to measure genetic variation using molecular genetic markers. There are many different sources of data relevant to genetic variation between breeds. However, more preference is being given to

microsatellites in population genetic studies. They appear to be abundant, averaging between 50,000 to 100,000 in mammals, and are ubiquitously distributed throughout the genome. Microsatellites display high degree of polymorphism with mean polymorphism information content (PIC) of 0.6 (Vaiman et al., 1995). They also improve the resolution of the measurement of variation and population in rare and endangered species, permitting the study of bottlenecks in these populations, with a high degree of consanguineous mating (Gotelli et al., 1994).

Domestication of the buffalo took place at an early age during Indus Valley civilization (3250 to 2750 BC). The buffalo is the first foremost a beast of labour, employed in the crop fields, pulling of carts, and working of mills and wells. Especially in India and other Southeast Asian countries it is stronger and more consistent than domestic cattle. In India ten important breeds of buffaloes have been described. The present study was under taken in two of the recognized buffalo breed viz; Marathwada buffalo of Maharashtra and Surti buffalo of Gujarat with following objectives:

1. To study the heterologus microsatellite loci primers for amplification in buffaloes.
2. To study the diversity of Marathwada and Surti buffaloes.
3. To study To the Hardy-Weinberg equilibrium, neutrality and linkage disequilibrium of the microsatellite loci in the buffalo populations.
4. To study the genetic distance among the two buffalo populations.
5. To assess the assignment and dispersion of the buffalo populations.
6. To analyze whether the two populations are in mutation drift equilibrium.
7. To analyze and study the various population genetic parameters of Marathwada and Surti buffaloes.

2. REVIEW OF LITERATURE

2.1 BIODIVERSITY AND CONSERVATION OF BUFFALOES

Biological diversity, or biodiversity, is the collection of species of plants, animals, bacteria, etc. (all forms of living organisms) that inhabit a particular ecosystem. The importance of indigenous livestock breeds lies in their adaptation to local biotic and abiotic stresses and to traditional husbandry systems. However, most of these animal genetic resources are still not characterized and boundaries between distinct populations are unclear. In such cases breeds have been defined on the basis of subjective data and information obtained from local communities. Moreover, the historical evidence is not always accurate; relying on it is subjective. Evolutionary research can reveal much about the original type of a breed/population or strain and the molecular genetics provides evidence that is factual and more precise. It is in this sphere the biotechnology has an important role to play for both the characterization and conservation of animal genetic resources.

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

Genetic uniqueness of populations is measured by the relative genetic distances of such populations is measured from each other. The first DNA polymorphism to be used widely for genome characterization and analysis were the restriction fragment length polymorphism (RFLP), which detects variations ranging from gross rearrangements to single base changes. Minisatellites sequences of 60 or so bases repeated many hundred or thousands of times at one unique locus within the genome have been used to genetic DNA fingerprinting typical of individuals within species. Microsatellites repeats of simple sequences, the most common being dinuceotide repeats are abundant in genomes of all higher organisms, including livestock species.

The change in environment results into the changes in the biological diversity. The change in genetic diversity due to loss of a particular species breeds or sudden changes in genotype because of large-scale migration through interbreeding or mutations, may result in loss of biological diversity. Genetic diversity refers to the variation at the level of individual genes (polymorphism), and provides a mechanism for populations to adapt to their ever-changing environment. Higher the variation, the better will be the chance that at least some of the individuals will have an allelic variant that is suited for the new environment, and will in produce offspring with the variant that will turn reproduce and shall maintain the population through subsequent generations.

The buffalo biodiversity in India constitutes 10 well-defined breeds of buffalo apart from non-descript types and some lesser-known breeds, which constitute around 60% in buffaloes. Nagpuri, Pandharpuri, Toda and Marathwada buffaloes are classified as swamp type. The river buffaloes are mainly dairy type and Murrah, Nili-Ravi, Surti, Mehsana, Jaffarabadi and Bhadawari breeds belong to this category. Besides the ten defined breeds, India possesses some population groups or strains also termed as Lesser Known Breeds. Some of these population groups have sufficient numbers and sufficiently distinctive traits to be conserved as distinct populations. The various indigenous breeds of buffalo in the country are the result of thousands of years of selection, evolution and development of the wild species in the process of domestication suiting to the local agro-climatic conditions. These breeds are now losing ground due to stiffer competition from other breeds and due to their poor economic viability under the present system of management. Cytogenetic studies have revealed that Nagpuri, Pandharpuri, Toda and Marathwada breeds also belong to riverine group (Nair et al. 1986).

The buffalo, aptly termed as the 'Black Goldmine of India', due to its contribution in terms of milk production occupies a crucial place in the economy of the country by being a source of plethora of items, utilize and services. The buffalo (*Bubalus bubalis*) belongs to:

Phylum : *Chordata,*
Class : *Mammalia*
Order : *Artiodactyla*
Sub order : *Ruminata*
Family : *Bovidae*

Sub family : *Bovinae*
Group : *Bubalina*
Genus : *Bubalus*
Species : *bubalis*

Although a vast buffalo population of Indian sub-continent and South East Asia is non-descript unimproved, yet India and Pakistan have numerous well-characterized breeds. Imperial Council of Agricultural Research in 1939 for the first time published a description of breed characteristic of Murrah. Brief descriptions of the two breeds undertaken in this present study are:

2.1.1 Marathwada Buffaloes

Marathwada buffaloes are found in the Marathwada region of Maharashtra state are entirely different from that of Western and Northern types and clearly represent a very ancient indigenous type characterized with medium-sized, lighter built and long flat horns. The breed is mainly found in Parbhani, Nanded, Bid, and Latur district of Maharashtra. The climate of Marathwada region is dry to semi-arid and sub-tropical. There are two agricultural divisions, viz. Aurangabad and Latur, in the region. The region is predominantly agrarian.

Marathwada buffaloes are of light to medium built with compact stature, and have adult weight of 300 and 370 kg. Averages of body length, height and heart girth of males are 132, 130 and 191 cm respectively (fig. 1). Averages of body length, height and heart girth of females are 130, 128 and 190 cm respectively (fig. 2). Coat colour varies from greyish-black to jet black, although white markings on forehead and on lower parts of the limbs with white switch of tail is not uncommon. Horns are medium in length, parallel to neck, reaching up to shoulder but never beyond shoulder blade like those commonly seen in Pandharpuri buffaloes and usually not flat. Fore head is moderately broad, and eyes are generally red tinged. Neck is short. Legs and feet are properly set which in males suit for draught and transportation in hilly tract. Tail is of moderate length reaching up to hock.

2.1.2 Surti Buffaloes

This breed is widely distributed in Surat district of Gujarat province of India. The home tract lies in Kheda, Bharuch, and Vadodara and Surat district of Gujarat covering a vast area between Mahi and Sabarmati rivers. Several other names are prevalent in other parts of Gujarat, which include Decani, Deshi, Nadiad, Charotar, Gujarati, Surati and Talabda.

The Surti buffaloes are brown or black in appearance. The hair in several specimens appears greyish or rustily brown. A characteristic feature is presence of two white chevrons. One around jowl from ear to ear and the other lower down the brisket. There may be a streak of white hair above the eyes. The region below the knees and hocks has a whitish or grey tuft of hair. The horns are medium-sized, sickle shaped, curving downward and backward and then inward in terminal portion. The tip of the horn forms a characteristic hock. These animals are of the

medium size with straight back and low legs. The hindquarters are wide, deep and straight unlike heavy buffalo breeds. The hips of Surti females are broad and flat. The udder is well-developed pinkish in appearance. The teats are medium sized, squarely placed with prominent milk vein. The average height of the adult male is 130 cm (fig. 3) and adult female is 125 cm (fig. 4). The average birth weight of male is 26 kg and female is 24 kg. The mature female's weight varies between 550 to 650 kg and while in males the range is between 640 to 730 kg (ICAR, 1941). The average lactation yield varies from 1200 to 2200 kg per lactation. The Surti animals are of short stature compared to other recognized buffalo breeds. The average height of animals is 1.3 meters and 1.25 meters in male and female respectively. The Surti females are fair milkers with average lactation milk yield 1285 kg per lactation. Gujarat farmers undertook extensive breeding of Surti buffaloes with Murrah. Mehsana breed is said to have been evolved from above crossing.

India is privileged in having a valuable diversity of buffalo types spread out in the length and breadth of the country- each type having its own special characters/features which make it efficient in its own habitat. A number of breeds have been undertaken for survey/characterization under different schemes viz. Network Project on Animal Genetic Resources (NPAGR), Network Project on Buffaloes, National Agricultural Technology Project (NATP), Adhoc schemes of ICAR and Research Project. It may be emphasized that India has not only the best breed of buffalo in the world (i.e., Murrah) but also the prized genetic diversity in the form of breeds and population groups spread out in different parts of the country. Microsatellite DNA markers for buffalo biodiversity studies have been identified and molecular genetic characterization has been completed for 5 breeds, while DNA analysis is underway for more. Genetic characterization will help to identify genetic relationship and uniqueness of Indian buffaloes.

2.2 MICROSATELLITE- A UNIQUE GENETIC MARKER

Awareness of the value of livestock genetic resources has stimulated the study of genetic diversity of native breeds- a first and most essential step for planning domestic animal diversity conservation plans. Arrays of molecular marker have been used for genetic variation studies at DNA level: Mitochondrial DNA, restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPDs) and variable number of tandem repeats (VNTRs). VNTR markers are characterized by a core sequence, which consists of a number of identical repeated sequences. They can be divided into two categories based on the repeat length. These are minisatellite, 15-70 base pairs, and microsatellites, 2-6 base pairs. Among these microsatellites are now considered as most suitable marker systems for genetic evolution studies (Paul, 2000).

Minisatellites are molecular marker loci consisting of tandem repeat units of a 10-50 base motif, flanked by conserved endonuclease restriction sites. They are detected by gel electrophoresis of restricted DNA and subsequent Southern blot hybridization to a radiolabeled DNA probe containing multiple copies of the minisatellite core sequence. A minisatellite profile

consisting of many bands (within the 4-20 kb range) is generated by using common multilocus probes that are able to hybridize to minisatellite sequences in different species. Molecular cloning of DNA restriction fragments and subsequent screening with a multilocus minisatellite probe can develop locus specific probes. Variation in the number of repeat units is considered to be the main cause of length polymorphisms. Due to the high mutation rate of minisatellites, the level of polymorphism is substantial, generally resulting in unique multilocus profiles. Therefore, minisatellites are particularly useful in studies involving genetic identity, parentage, clonal growth and structure, and identification of varieties and cultivars (Jeffreys et al., 1985a; Jeffreys et al., 1985b).

Microsatellites are markers; contain repetitive sequence composed of di-, tri-, or tetra-nucleotides (Litt and Luty, 1989). Microsatellites are also known as simple sequence repeats (SSRs) (Tautz, 1989), short tandem repeats (STRs) (Edward et al., 1991) and sequenced tagged microsatellite repeats (STMRs). The number of repeats may vary from one person to another, i.e. polymorphic, and follow Mendelian inheritance. Due to their high informativeness and easy accessibility, these markers have been widely applied in various field of genetic research. For testing the genetic bottleneck the microsatellites are probably the best marker available at present.

Microsatellite in different individual are not just the same or a different, they can differ by just a few repeat units or by many repeat units. Because pedigree experiments have shown that most mutations are a change in one repeat unit, some researchers have used microsatellites as molecular clocks. By studying the average number of repeat differences in many microsatellite loci, one can infer the time to the most recent common ancestor of two individuals. Microsatellite have been used to estimate the age of non-microsatellite mutations. Microsatellites mutate by changing the number of their repeats, for example the $(AT)_{25}$ might become on $(AT)_{24}$ or $(AT)_{26}$ in that individual's offspring. These length-changing mutations occur at rates several orders of magnitude higher than point mutations. The reason for microsatellites popularity as genetic markers is that their high mutation rates make them highly polymorphic, and they are relatively dense in the genomes of many organisms. The vast majority of microsatellites in higher organisms are believed to evolve naturally, i.e., there is no selection pressure on the number of repeats.

Microsatellites and their flanking sequences can be detected by constructing a small-insert genomic library, screening of the library with a synthetic labeled oligonucleotide repeat and sequencing of the positive clones. Alternatively, primers designed for closely related species may be used. Polymerase slippage during DNA replication (or slipped strand mispairing) is considered to be the main cause of variation in the number of repeat units, resulting in length polymorphisms that can be detected by gel-electrophoresis. Due to their high level of polymorphism, microsatellites are informative markers that can be used for many population genetic purposes, ranging from the individual level (e.g. clone and strain identification) to closely related species. In addition, microsatellites are considered ideal markers in gene

mapping studies (Hearne et al., 1992; Morgante & Olivieri et al., 1993; Queller et al., 1993; Jarne & Lagoda, 1996). Based on the nature of the repeats, microsatellites are broadly classified into three families- pure, compound, and interrupted repeats (Weber, 1990).

- Pure - CACACACACACACA
- Compound - CACACACACACACAAACGCGCGCG
- Interrupted - CACACATTCACATTACACACACACA

Any kind of combination between these families is possible. Microsatellite is highly polymorphic in natural population but compound and interrupted microsatellites tends to be less Polymorphic than pure families. Dinucleotide, trinucleotide, and tetranucleotide repeats are most commonly found.

Dinucleotide repeats: These are most frequently used loci with the CA repeat being a ubiquitous repetitive element in eukaryotic genomes. The density of distribution varied widely with about one locus per 5 kb. At each locus the number of dinucleotide repeat is generally less than 30 (Di Rienzo et al., 1994).

Trinucleotide repeats: They occur in both plants and animals are largely studied in connection with some human diseases and cancer (Charlesworth et al., 1994) and are often found within exons where they do disrupt the reading frame (Jurka and Petiyagoda, 1995). This makes them unlikely candidates for population genetic studies. Their mean number of repeats and polymorphic is similar to that of dinucleotides. The rate of mutation of trinucleotide repeat is substantially less than dinucleotide.

Tetranucleotide repeats: GATA and GACA are almost the only tetranucleotide studied. GAT/GCA occurs in many higher organisms but they vary in distribution among species. In the human genome, they are clustered near centromeres. They are often found as compound or interrupted stretches and are highly polymorphic. They are useful for population genetic studies especially in evolution of populations for bottlenecks.

Pentanucleotide repeat: In all mammalians taxa pentanucleotide repeats as abundant as triplet repeats both in introns and intergenic regions A+T rich are generally more abundant.

Since their discovery, microsatellite have been used in gene mapping programmes and by population biologist for studies of population genetic structure and kinship investigations. Microsatellite can be easily characterized in those organisms for which genomic mapping or sequence data is available. But for those organisms for which such information is not yet available, the process of microsatellite characterization tends to become laborious as it involves preparation of genomic library, cloning detection of microsatellites and sequencing in order to determine the flanking sequences which can then be used as PCR primers for that specific locus.

Microsatellite loci are found in large numbers and are relatively evenly spaced throughout the genome. Technically microsatellites are more desirable than the larger VNTR loci because they can be analyzed via the polymerase chain reaction (PCR) (although improvements to the

thermal stable polymerases are enabling the amplification of larger fragments of DNA) and the alleles can be unambiguously sized on polyacrylamide gels. The lack of exact size resolution of larger VNTR loci has lead to procedures such as binning which reduce the statistical power of the analysis (Budowle et al., 1991). PCR analysis of small fragments also allows the analysis of degraded samples in which the mean fragment size of the genomic DNA has been severely reduced through environmental insult (e.g. Paabo et al., 1989). Finally microsatellites have been found to be variable even in populations, which have low levels of allozyme and mitochondrial variation (Estoup et al., 1995a, 1996; Paetkau & Strobeck, 1994).

2.3 ISOLATION OF MICROSATELLITE MARKERS

In order to construct linkage maps, polymorphic markers need to be identified, isolated and characterized. Tandem repeat sequences (microsatellites) are first detected from the entire genome and their unique flanking sequences are used to develop primers for amplification of the specific microsatellites by PCR. Broadly two strategies are used for the isolation of microsatellite markers.

2.3.1 Cosmid Derived Microsatellite Marker

In this strategy, the genomic DNA, after digestion with restriction enzymes is cloned into suitable vector mostly cosmid, thus forming a cosmid genomic library. The cosmid are then screened with a labeled $(CA)_n$ or $(GT)_n$ polynucleotied probe. The clones that are hybridized to the probes are detected by autoradiography. The positive clone is isolated and the insert (microsatellite), which they harbour, is sequenced and characterized. Appropriate primers are designed from the flanking regions.

2.3.2 Microdissected Chromosome Derived Microsatellite Markers

A chromosome spread is obtained from a blood culture in this methodology and chromosome of interest is identified under a microscope. This chromosome is dissected using a micromanipulator. The microdissected chromosomal fragments are then used to construct genomic DNA library which are then screened with radiolabeled $(CA)_n$ or $(GT)_n$ probes. Positive clones are isolated and subjected to PCR amplification. The PCR products are sequenced are checked for uniqueness to develop PCR primers (Zao et al. 1999). A modification of this method involves amplification of the microdissected chromosomal fragments by PCR using degenerate oligonucleotide primers. To the amplified products biotinylated $(CA)_n$ probes added. After denaturation and annealed DNA is added to streptavidin paramagnetic particles and incubated to capture DNA fragments hybridized to biotinylated $(CA)_n$ probes. The bound DNA is eluted and amplified using appropriate primers. The amplified products are purified and sequenced to be used as markers (Dukhanina and Sverdlov, 1998; Sarkar et al. 2001).

2.4 MICROSATELLITE POLYMORPHISM

Microsatellites are the highly polymorphic which makes them attractive as genetic markers (Goldstein and Shlotterer, 1999). Since microsatellite polymorphisms are visualized through PCR, each microsatellite locus can be considered a sequence tagged site (STS) (Olson et al., 1989) as the two primers serve as the sequence tags. PCR based microsatellites have facilitated the construction of genome maps in most live stock species because of their abundance in the genome, specificity of primers, high degree of polymorphism which yield several alleles and easy detection. The degree of polymorphism and heterozygosity detected by microsatellites is much greater than that of protein markers.

For the naturally evolving DNA sequences, the amount of the polymorphism is expected to be directly proportional to the mutation rate (Kimura, 1983). Hence it has been assumed that the mutation rate to form new length variants at the microsatellite loci is appreciable, and this idea has been substantiated by direct observation of the spontaneous events of the germline mutation from the pedigree analysis (Weber and Wong, 1993; Goldstein and Pollock, 1997).

The eukaryotic DNA sequences mutate at the rate of approximately 10^{-9} per nucleotide per generation (Crow, 1993). The mutation rate of the microsatellites is several orders of the higher magnitude, often quoted in the range of the 10^{-3} to 10^{-4} per locus per generation (Weber and Wong, 1993). The most recent data from the large-scale human genome mapping or paternity testing based on many different loci suggest an even higher rate of about 2×10^{-3} per meiosis (Ellegren, 2000a; Kayser et al., 2000). The degree of polymorphism, at least for the mammalian $(CA)_n$ repeats is positively correlated with the average number of the repeat units.

The length of microsatellite repeats may have an effect on the mutation rate such that longer are more polymorphic than shorter once (Weber, 1990). This is probably because the opportunity for a stable misaligned configuration is greater for longer repeat arrays. The second parameter that influences microsatellite stability is the purity of the repeat. Interrupted microsatellite repeats seems to have lower mutation rates than perfect repeats. Many studies have shown that microsatellite loci involve more gains than losses of repeat units (Weber and Wong, 1993). Mutational process of microsatellite seems to be very complex process. It is very likely that these processes are heterogeneous with difference between loci and alleles.

2.5 EVOLUTION OF MICROSATELLITE

Microsatellites are highly variable. To “evolve” simply means to change. Microsatellite alleles change (mutate) over time. In a population, there may exist many alleles of a single microsatellite locus. Microsatellite alleles differ in the number of repeats. For example, one allele may have 7 repeats of a CT motif and another 8 repeats. In a population, there may exist many alleles (upto 70 or 80) at a single locus, with each allele having a different length. An individual who is homozygous for a locus will have the same number of repeats on both chromosomes, whereas a heterozygous individual will have different numbers of repeats on the two chromosomes. The regions surrounding the microsatellite locus, called the flanking regions,

may still have the same sequence. This is important because the flanking regions can therefore be used as PCR primers when amplifying microsatellite loci, and can be conserved across genera or sometimes-even families. Below, the two lines represent the sequences on two homologous chromosomes in a diploid organism.

Homozygous: (Both strands have 7 CT repeats).

……...TAGCCTTGCATCCTT**CTCTCTCTCTCTCT**ATCGGTACTA …

……..TAGCCTTGCATCCTT**CTCTCTCTCTCTCT**ATCGGTACTA …

5’ Flanking region Microsatellite locus 3’ Flanking region

Heterozygous: (One strand has 7 repeats and the other has 8 repeats).

…...AGCCTTGCATCCTT**CTCTCTCTCTCTCT**ATCGGTACTA ……...

…...AGCCTTGCATCCTT**CTCTCTCTCTCTCTCT**ATCGGTACTA……

5’ Flanking region Microsatellite locus 3’ Flanking region

It is estimated that microsatellites mutate 100 to 10,000 times as fast as base pair substitutions (transitions and transversions). This makes microsatellites useful for studying evolution over short time spans (100 to 1000 years); where as base pair substitutions are more useful for studying evolution over long time spans (millions of years).

2.6 REPLICATION SLIPPAGE

SSRs are known to be prone to a high rate of mutation via a mechanism called slipped strand mispairing (or replication slippage). It involves transient dissociation of the replicating DNA strand followed by misalignment reassociation (Richards and Sutherland 1994); strand slippage replication appears to be predominant mode at microsatellite evolution (fig. 5). The *in vivo* experiments clearly demonstrate that the microsatellite sequences have the intrinsic ability to undergo the DNA slippage (Schlotter and Tautz, 1992). When DNA replicates, the two strands sometimes disassociate. In non-repetitive DNA, the strands reassociate the same way they were before the slippage event, with matching base pairs on the opposing strands. But in repetitive DNA, since there are so many possible matching base pair alignments, sometimes the strands realign differently, forming an unmatched loop on one of the strands.

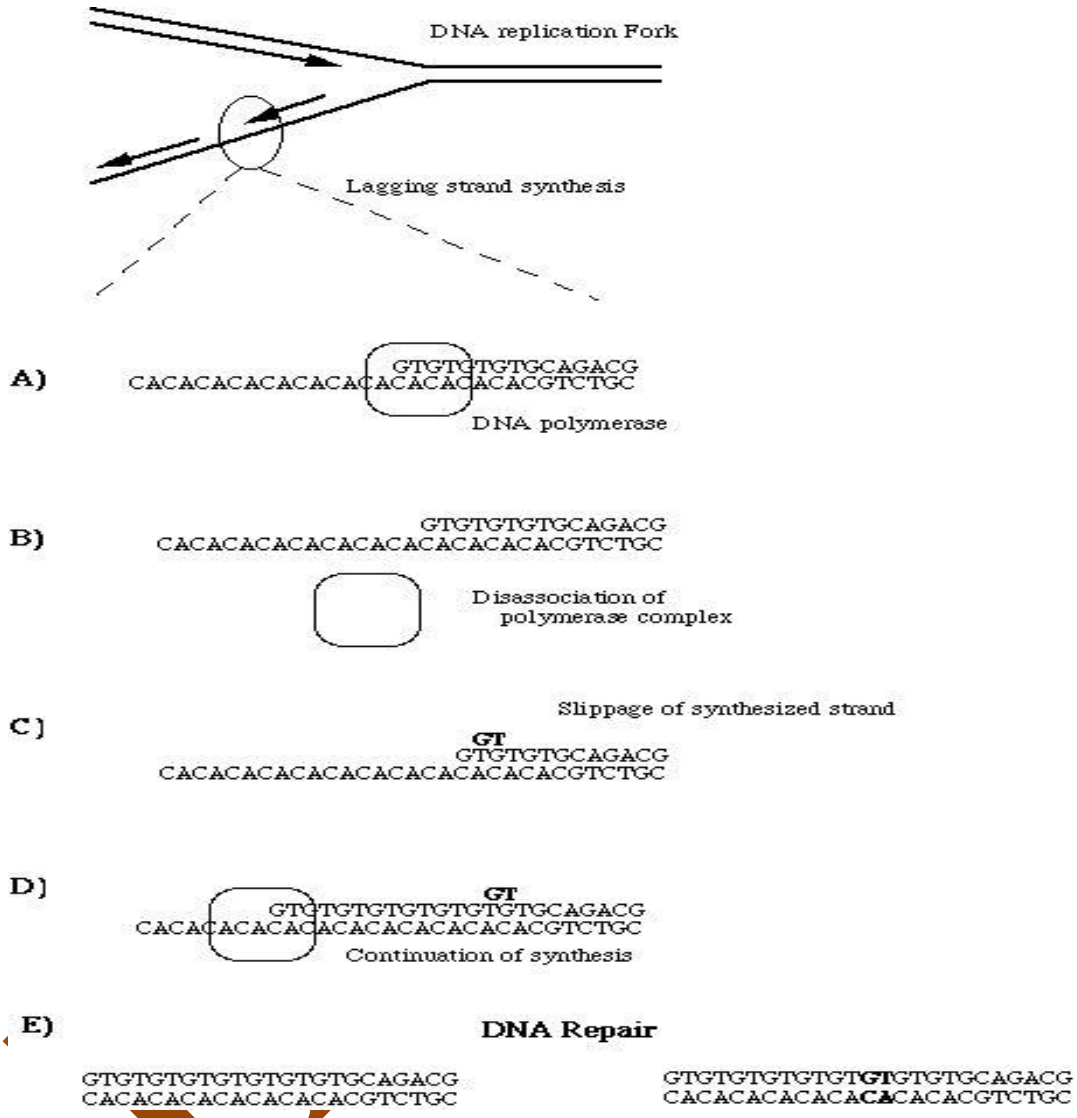


Fig. 1: Diagrammatic Representation of DNA Polymerase Replication Slippage Model

When the two strands completely dissociate and begin replication anew the strand, which had the loop, will contain a longer microsatellite than the opposing strand. The microsatellite on the template strand will always have the same length before and after the slippage event. If the loop is on the template strand, then the microsatellite on the replicating strand will be shorter; and if the loop is on the replicating strand, then the microsatellite is on its side will be longer. These primary mutations, which depend exclusively on the DNA replication machinery, occur at rates several orders of magnitude higher than the observed mutation rate and are countered by the DNA mismatch repair machinery. Thus the observed mutations are those replication

slippage events that have escaped repair. The longer microsatellites present more opportunity for slippage, which would expect mutation rates to increase as a function of microsatellite length; this prediction is experimentally supported. Interrupted microsatellites have lower mutation rates as interruption allow less realignment after a possible slippage event. A further complication to measuring microsatellite variability is that insertions/deletions in the flanking regions can also affect the PCR fragment length.

2.7 FUNCTIONAL SIGNIFICANCE OF MICROSATELLITE

Although SSRs are usually considered just as evolutionary neutral DNA markers, the functional significance of a substantial part of SSRs has been proven by critical tests in various biological phenomena. The functional significance of the microsatellites remains to be clearly understood. Several probable roles have been proposed, all more or less relating to the unique capability of the alternating purine-pyrimidine residues to form the Z-DNA (Hamada et al., 1984b). The $(GC)_n$ or $(CA)_n$ can change from the B-DNA to Z-DNA, *in vitro*, in response to the various environmental factors like the elevated ionic strength (Klysik et al., 1981), negative torsional stress (Hanniford and Pulleyblank, 1993), presence of intercalators and the modification of guanine or cytosine residues (Rich et al., 1984). The identification of nuclear proteins that preferential bind to the Z-DNA and *in situ* application of the antibodies raised against the Z-DNA provide indirect evidence of possible existence of the Z-DNA *in vivo* (Nordheim et al., 1982).

When the Z-DNA structure had been demonstrated for alternating purine-pyrimidine residues, the involvement of this structure in recombination process became the tempting suggestion (Jelinek et al., 1980). The stickiness of the left-handed DNA could facilitate the correct pairing of the homologous chromosomes during the meiotic recombination (Blaho and Wells, 1989). The DNA sequences associated with synaptonemal complexes are rich in the microsatellite sequences (Pearlman et al., 1992). The presence of the microsatellite loci may also affect the reciprocal meiotic exchange (Schultes and Szostak, 1991). The $d(CA)_n.d(GT)_n$ microsatellites have been demonstrated to inhibit rec-A promoted strand exchange *in vitro* (Gendrel et al., 2000).

The simple tandem repeats are absent in the bacteria and are more frequent in the Euchromatin than the heterochromatin (Stallings, 1992). Their conformational properties may provide a suitable condition for the repeated packaging and condensing of DNA during the cell cycles (Stalling et al., 1991). These highly repetitive and conserved sequences might also function as the repository of the species. It is also a possibility that they might not have any function at all and these are just the "junk" DNA that is carried along by the process of replication and segregation of the chromosomes. However, the validity of these postulated functions of the microsatellites needs the further investigation.

2.8 MOLECULAR DETAILS OF MICROSATELLITES

2.8.1 Mutation Rates

A variety of *in vivo* and *in vitro* studies indicate the microsatellite loci are highly unstable, having some of the highest mutation rates observed at molecular loci. Microsatellite loci are highly unstable having some of the highest mutation rates observed at molecular loci means adding or subtracting a small number of perfect repeats. It is due to polymerase slippage (Chevinson and Gutmah 1987) mutation may high as 10^{-3} (Jaffrey et al.1998, Kelley et al.1991). It is possible, but not proven, that under conditions where a mutation in a microsatellite would be favorable. The eukaryotic DNA sequences mutate at the rate of approximately 10^{-9} per nucleotide per generation (Crow, 1993). The mutation rate of the microsatellites is several orders of the higher magnitude, often quoted in the range of the 10^{-3} to 10^{-4} per locus per generation (Weber and Wong, 1993).

2.8.2 Distribution of mutation sizes

The majority of observed mutation is of a single step (one repeat unit) or two steps, a significant minority of mutations may be of longer size. The occurrence of mutations longer than one or two steps is confirmed by studies of trinucleotide expansions in which alleles beyond a certain size threshold have asymmetric distribution of mutations including some of larger size. Atypically large alleles are hyper mutable, leading to the production of expanded, symptomatic alleles; this suggests a potential mechanism of size constraint.

2.8.3 Asymmetry of mutation distribution

A tendency to mutate to alleles of larger size (positive asymmetry) was first observed in the asymmetric mutation distribution for large alleles at trinucleotide expansion loci (Ashley & Warren, 1995). The behavior of expanded trinucleotide alleles, suggest that the degree of asymmetry may depend on allele size. In assessing asymmetry it will therefore be important not only to consider differences among loci but also differences among alleles within a locus.

2.8.4 Range constraints

The most compelling evidence that the number of repeats at microsatellite loci is under some form of constraint is simply the absence of alleles of very large size. Given the high mutation rate, and very large number of loci that have been characterized, it is clear that if the process were an unconstrained random walk we would expect to regularly observe loci with very large alleles. In fact, with the exception of trinucleotide expansion loci, alleles much greater that 60 repeat are very rarely observed (Primmer et al. 1996).

2.8.5 Dependence of the mutation process on allele size and sequence

In trinucleotide repeat expansion loci, the rate and distribution of mutations change dramatically as allele size pass from the pre-mutation (atypically large but non-symptomatic) to full mutation state (Ashley & Warren, 1995). A number of population studies have also tested the dependence of the mutation rate on allelic size by correlating observed levels of variation with average allele size. *Goldstein and Clark* (1995) analyzed the dependence of the allelic variance on the repeat count itself, considering both the average size and the maximum size at a

locus. The increase in mutation rate with repeat size is not linear, or that some other assumption of the stepwise mutation model is violated. The imperfection in the repeat array tends to stabilized microsatellites (Goldstein and Clark, 1995).

2.9 MUTATIONAL MODELS AND THEIR PROBABLE MECHANISMS

Genotypic data are multivariate: the frequency of each allele at each locus is usually different in each population. Analysis of genotypic data from natural loci is an important method for describing the patterns of genetic variation within species and inferring the evolutionary processes that give rise to those patterns. Genetic distances are metrics that summarize these differences in overall measures of differentiation for a pair of populations. Generally, a matrix of pair wise genetic distances between a set of population is estimated. This metrics is often visualized with phenograms, isolation by distance plots, principal component analysis, or multi-dimensional scaling plots. To estimate population differentiation measures and genetic distances from microsatellite data, theoretical mutational models for the evolutionary process of microsatellites are needed. By ignorant of the actual mutational details, two different assumptions are used to model mutations. These are based on standard models from molecular population genetics, and are:

2.9.1 Infinite Alleles Model (IAM)

Infinite alleles model is a term used to refer to a model where each new mutant is different (i.e., there are an infinite number of states that an allele can mutate too, hence each mutation is assumed to be unique. The motivation is considering a long DNA sequence, were any two changes are extremely likely to given two different sequences). If each new mutant is different, then we don't have to correct for mutations that we did not see and hence did not score, and (as the fig. 6 below shows) we can simply look at the differences between marker loci and count the actual number of mutations that have occurred.

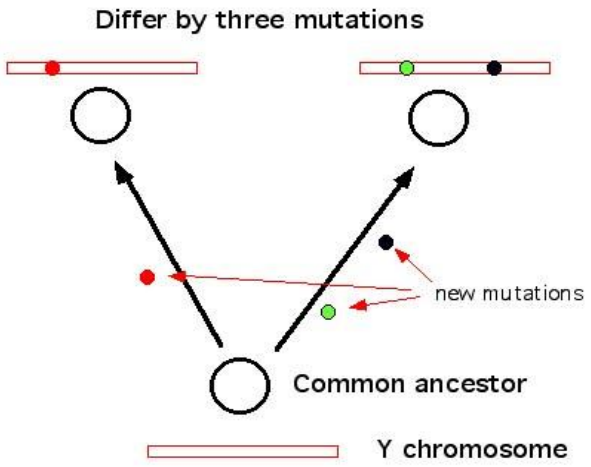


Fig. 2: Diagrammatic Representation of Microsatellite Evaluation Using Infinite Allele Model

[The red mutation on the left individual's Y chromosome and the black and green on the right individual's]

In the infinite alleles model (Kimura and Crow, 1964) mutation can involve any number of tandem repeats and always results in a new allele state not previously existed in population. But this model does not confer with the slipped strand mispairing mechanism responsible for microsatellite length variation. Most classical distance measures, however, are based either on IAM; or upon multidimensional geometric considerations without reference to a particular evolutionary model.

2.9.2 Stepwise Mutation Model (SMM)

The stepwise mutational model was introduced in 1970s. The model assumes that the entire sequence of the allelic states can be expressed as integers and mutation results in a change in a one repeat unit either by insertion or deletion (Ohta and Kimura, 1978). SMM tries to better account for the actual mutational process that occurs at microsatellite markers. The SMM has been applied to microsatellite allele frequencies (Valdes et al., 1993). Three to five base pair repeat evolve via the SMM while one–two base pair repeats are by TPM. The SMM is generally the perfect model when calculating relatedness between individuals and population subtracting although there is problem of homoplasy.

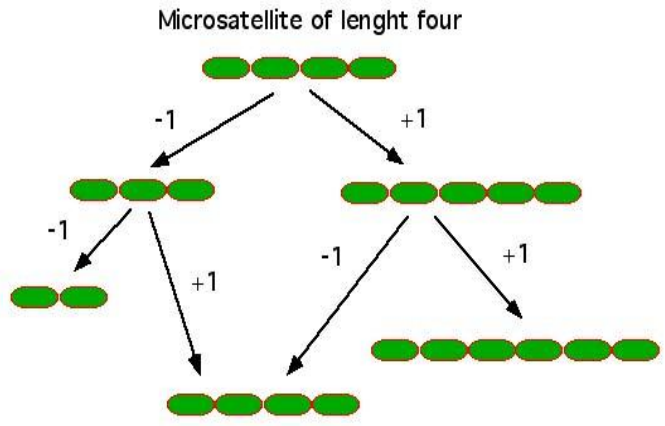


Fig. 3: Diagrammatic Representation of Microsatellite Evaluation Using Stepwise mutation model

[Two mutations, (-1) mutation followed by a (+1) mutation returns the marker to its original length.]

SMM follows the changes in length, allowing a step up (or down) by one at each mutation (fig. 7). The SMM developed for allozymes provides better description for these kinds of evolutionary processes. Different kinds of repeat number variance estimators based on SMM have been developed for estimating phylogenetic relationship, genetic distance and population differentiation [$(\delta\mu)^2$; Goldstein, 1995 a, b; Dsw, Shriver, 1995; R_{st}, Slatkin, 1995]; from microsatellite data.

The stepwise mutation model is the oldest model of microsatellite evolution, originally proposed for electrophoretic alleles. In this model the number of repeat units is equally likely to increase or decrease by one at a rate independent of the microsatellite's length, subject to the constraint that the number of repeat units cannot become smaller than one. Let X be the length of the microsatellite, then

$$X \to X + 1 \text{ at rate } \gamma \text{ ; and}$$

$$X \to X - 1 \text{ at rate } \gamma$$

This is a symmetric random walk with a lower boundary condition. Numerous complications to this basic model have been introduced.

The first complication we will discuss is allowing the mutation rate to depend on the microsatellite's length. For example, Kruglyak *et al.* proposed a proportional slippage model where the mutation rate increases the microsatellite's length.

$$X \to X + 1 \text{ at rate } b(X-1) \text{ ; and}$$

$$X \to X - 1 \text{ at rate } b(X-1)$$

Another complication is to allow the mutation rates to be asymmetric. Walsh proposed a linear birth death chain, *i.e.*, a proportional slippage model where the mutation rate increases linearly with the microsatellite's length in the presence of biased contractions.

$X \rightarrow X + 1$ at rate bX

$X \rightarrow X - 1$ at rate dX

For X $\{2,3\ldots\infty\}$ and $1 \rightarrow 2$ at a much smaller birth rate υ. He showed that a stationary distribution exists for this model when $d/b > 1$. Fu and Chakraborty proposed a model, which allows larger mutations according to a geometric distribution in the presence of a constant mutational bias.

The final model complication will consider is point mutations. Point mutations can interrupt a microsatellite, for example transforming $(AT)_{20}$ into $(AT)_{12}GT(AT)_7$. Since most researchers measure the length of microsatellites without sequencing they would not detect this transformation. Bell and Jurka proposed that such point mutations constrain the growth of microsatellites. Kruglyak *et al.* (1998) proposed a model with two processes:

1. Single-step proportional slippage :

$X \rightarrow X + 1$ at rate $b(X - 1)$, and

$X \rightarrow X - 1$ at rate $b(X - 1)$

2. Point mutations:

$X \rightarrow j$, where $j < X$ at rate a.

Thus a point mutation is uniformly likely to affect any of the repeat units. Durrett and Kruglyak showed that this model has a stationary distribution.

2.9.3 Two-Phase Model (TPM)

In addition to this model, Di Rienzo et al. (1994) described the two-phase model (TPM), where a limited proportion of mutations involve several repeats. The two-phase model, incorporates the mutational process of the SMM, but allows mutations of longer magnitude to occur.

2.9.4 K-Allele Model (KAM)

Although rarely cited in microsatellite literature, a K-allele model (KAM) could also be considered for microsatellites. Under this model, there are K possible allelic states, and any allele has a constant probability of mutating towards any of the other K-allelic states (Crow and Kimura, 1970). Due to size constraint acting on microsatellite loci, the KAM seems to be more realistic than the IAM.

Theoretical mutation models like SMM and TPM may provide adequate measures if populations are relatively closely related, but these simple models become inadequate when divergence between populations and especially between species increases (Takezaki & Nei 1996).

There are several ways to study microsatellite evolution. First, theoretical studies attempt to model the process of microsatellite evolution by applying assumptions to a range of parameters considered to be important to the mutational process. After computer simulations the resulting data can be compared to the observed distribution of allele frequencies and/or the

heterozygosity of a locus (e.g. Deka *et al.* 1991, Shriver *et al.* 1993). Valdes *et al.* (1993) and Di Rienzo *et al.* (1994) used an alternative method where they compared the empirical and modeled allele frequency distributions. These studies have shown that SMM and TPM can explain relatively well the evolutionary processes of microsatellites.

It is very likely that these are allele and locus-dependent processes (Ellegren 2000b). Theoretical mutation models, such as SMM and TPM may accurately represent the evolutionary processes of microsatellites when closely related populations are considered. However, over long evolutionary distances the mutation process seems to be more complex. Thus, theoretical mutation models that can more accurately represent the evolutionary processes of microsatellites needed to obtain better estimates of population differentiation measures.

2.10 ADVANTAGES OF MICROSATELLITES

Earlier studies have been demonstrated that considerable numbers of microsatellite loci are conserved across livestock species. Cross species utilization of microsatellite loci would not only save time and efforts put forth in the detection of suitable panel of polymorphic microsatellite marker for genetic evaluation of specific animal species but also enable the construction of comparative maps between species. Recently microsatellites have been increasingly used as the marker of choice. There are some advantages to utilizing microsatellites over the other markers, which make them desirable. Compared to the DNA profiling with VNTR loci, commonly applied in the microsatellite typing offers several advantages (Devlin et al., 1990).

Part of the appeal of microsatellite over RFLPs and RAPDs is that the genetic bases of microsatellite variability are readily apparent: unique primers amplify a genomic region including a well- defined repeat structure that is responsible for the observed variation. This allows the development of inferential methods based on explicit models of microsatellite evolution (Slatkin 1995a,b; Goldstein et al. 1995a,b; Goldstein et al. 1996, Feldman et al. 1996, Pollock et al. 1996). These advantages suggest that microsatellite will enjoy a lengthy region in population studies. Some of the more advantages of microsatellites are as follows:

- Microsatellites are very abundant and spread over the entire genome of all living organisms, both in protein coding and non-coding regions. SSRs in exons are less abundant than in non-coding regions (Hancock, 1995). Because of there high mutability, microsatellites are thought play a significant role in genome evolution by creating and maintaining quantitative genetic variation (Tautz et al. 1986; Kashi et al. 1997).
- Microsatellites are highly polymorphic due to this nature higher average per locus heterozygosity and Polymorphic Information Content (PIC) in any population can be easily calculated.

- Alleles size can be determined with an accuracy of one base pair, allowing accurate comparison across different gels and different microsatellites may be multiplexed in PCR or on gel.
- Microsatellite primers developed for one species frequently amplify in related species.
- For captive or endangered species microsatellites can serve as tools to evaluate inbreeding levels (F_{IS}). Using tools such as F-statistics and genetic distances, genetic structure of sub-population and population can be known.
- Microsatellites can be used to assess demographic history (e.g., to look for evidence of population bottlenecks), to assess effective population size (N_e) and to assess the magnitude and directionality of gene flow between populations.
- Microsatellite may affect DNA replication (Field & Wills, 1996), gene translation and enzyme controlling cell cycles.
- Microsatellites located in promoter regions may affect gene activity and in intronic region can also affect gene transcription.
- Numerous microsatellite and minisatellite DNAs have been proposed as hot spot for recombination (Jeffreys et al. 1998; Templeton et al. 2000).
- Microsatellite in the Eukaryotic DNA mis-match repair (MMR) gene as modulators of evolutionary mutation rate. DNA MMR proteins correct replication errors and actively inhibit recombination between diverged sequences (Chen & Jinks-Robertson, 1998; Kolodner & Marsischky, 1999).
- Some microsatellites, found in upstream activation sequences, serve as binding sites for a variety of regulatory proteins (Lue et al. 1989, Csink & Henikoff 1998) and may also affect protein binding.
- The great increase in successful studies in a number of species has resulted from major improvements of reference genetic linkage maps consisting of STRs (Fries et al., 1989).
- Microsatellite typing is also used as tool for identity or paternity testing by detection of hyper variable sequences.
- They serve as a role in biomedical diagnosis as marker of certain disease conditions i.e., certain microsatellite alleles are associated with certain mutations in coding regions of the DNA that can cause a variety of medical disorders.
- They have also become the primary marker for DNA testing in forensic (court) contexts-both for human and wildlife cases.

2.11 APPLICATIONS OF MICROSATELLITE

Besides, the obvious advantage of PCR based analysis, the applicability of the microsatellite markers in genome analysis primarily depends on the three inherent properties: abundance, hyper variability and Mendelian inheritance. These properties make the microsatellites very

informative markers in the genome analysis and are used for various applications in Bubalus bubalis *and other species.*

2.11.1 Genetic analysis of closely related breeds/ populations

The allelic frequency data obtained after the PCR-based genome scoring can be utilized for studying the evolutionary relationship of the closely related breeds/ populations of a species or closely related species (Bowcock et al., 1994; Laval et al., 2000). The high degree of polymorphism make them the marker of choice for such studies over the conventional marker like the restriction fragment length polymorphism, which generally have only two alleles, and hence a maximum theoretical heterozygosity of 50% (Botstein et al. 1980). The microsatellites gives discriminating and significantly concordant result as compared to RAPD (Bart-Delabasses et al., 2001). The more and more reports are now appearing in the farm animals including buffaloes describing genetic characterization of the breeds using microsatellite markers (Arora et al., 2004; Vijh et al., 2004).

Ali et al. (1993), used a synthetic oligodeoxyribonucleotide probe (OAT36) comprising nine repeats of 5' GACA 3' and several enzymes were used to analyse cow, (*Bos taurus*) and buffalo (*Bubalus bubalis*) genomes and a number of monomorphic loci were detected in both the species. Different animals from the same species showed an almost 'similar' monomorphic hybridization pattern but animals from two separate species showed a different 'genome specific' pattern. The overall hybridization with any enzyme and probe combination was found to be unique to one species. This forms the basis of genome specific hybridization that is substantiated by our zoo-blot hybridization studies. The evolutionary aspect of these loci in the context of sequence polymorphisms is discussed.

Moore et al. (1995), report on the ability of a set of eighty bovine derived DNA microsatellite primers to amplify sequences in the two types (swamp and river) of water buffalo (*Bubalus bubalis*). Number of alleles and percent heterozygosities in a large number of animals were determined on a subset of microsatellite loci selected on the robustness of the primers. These loci will form the basis of a set of polymorphic DNA markers for use in water buffalo.

Jonh and Ali, (1997) uses a series of synthetic oligonucleotide probes were developed as markers for genetic analysis and molecular systematics of *Bubalus bubalis* and other eutherian mammals. A very high level of heterozygosity (approximately 95%) was observed in the bubaline genome, with an overall band-sharing probability of 2.08×10^{-11} using $(TGG)_6$ and *Hinf I* probe-enzyme combination. Breed affiliation studies on different buffalo breeds, viz. Toda, Surti, Mehsana, and Murrah, revealed that the semi-wild Toda breed belonged to a distinct group. Employing a DNA fingerprinting approach using these markers also identified the desired genotypes in each successive generation in an actual breeding program. Estimation of genetic distances by calculating the mean allelic frequencies at $(CA)_n$, $(TGG)_n$, and $(GGAT)_n$ repeat loci between buffalo and other related animals such as horse (order *Perisodactyla*), rabbits (order *Lagomorpha*), pigs, cattle, goat, and sheep (order *Artiodactyla*) revealed that with respect to the genetic distance coefficient (GDC), goat (caprine) was closer to buffalo (GDC =

0.0005) than sheep (ovine) (GDC = 0.0007). The genetic distance between horse and buffalo was calculated to be 0.4085, indicating that compared to other animals, horse was distantly related to buffalo. The understanding of overall allelic variations and breed affiliation of the bubaline genome will contribute to the propagation and conservation of the desired germplasm and better management of this species.

Mattallil and Ali, (1999), studied the distribution and evolutionary pattern of the conserved microsatellite repeat sequences $(CA)_n$, $(TGG)_6$, and $(GGAT)_4$ were studied to determine the divergence time and phylogenetic position of the water buffalo, *Bubalus bubalis*. The mean allelic frequencies of these repeat loci showed a high level of heterozygosity among the euartiodactyls (buffalo, cattle, sheep, and goat). Genetic distances calculated from the allelic frequencies of these microsatellites were used to position *Bubalus bubalis* in the phylogenetic tree. The tree topology revealed a closer proximity of the Bubalus bubalis to the Ovis aries (sheep) genome than to other domestic species. The estimated time of divergence of the water buffalo genome relative to cattle, goat, sheep, pig, rabbit, and horse was found to be 21, 0.5, 0.7, 94, 20.3, and 408 million years, respectively. Although water buffaloes share morphological and biochemical similarities with cattle, their study using the microsatellite sequences placed the bubaline species in an entirely new phylogenetic position. Their results also suggested that with respect to these repeat loci, the water buffalo genome shared a common ancestry with sheep and goat after the divergence of subfamily *Bovinae* (*Bos taurus*) from the family *Bovidae*.

Arora et al. (2004), the genetic diversity in two buffalo populations of northern India, the Bhadawari and the Tarai was assessed using a set of 22 heterologous (bovine) microsatellite markers. The average number of alleles across all loci in both populations was found to be 4.7, indicating that this set of 22 bovine microsatellite markers could be used to study genetic variation in buffalo species also. The overall polymorphic information content (PIC) value for these markers was 0.54. The average observed and expected heterozygosities for both populations were 0.59 and 0.64, respectively. Common alleles with varying allele frequencies in both populations also represented the genetic variability existing between Bhadawari and Tarai buffaloes. However the θ estimates for population differentiation indicated low levels of differentiation between the two populations. This was further supported by the low genetic distance (0.155) between Bhadawari and Tarai, which was calculated using Nei's standard genetic distance method. The present study on Bhadawari and Tarai populations represents a much-needed preliminary effort that could be extended to other local buffalo populations of India as well.

Vijh et al (2005) the blood samples from 104 individuals were collected from unrelated animals belonging to three buffalo populations. The populations were Tarai, Bhadawari and Local populations of Kerala. The DNA was isolated and microsatellite data on 24 loci was generated. The data was subjected to analysis for the estimation of genetic distances based on arithmetic and geometrical considerations viz; Nei's D_a, Allele sharing distance Das, Cavalli Sforza Edward's Chord distance D_c, Prevosti distance C_p and Ronger's distance D_r. The

phylogenetic tree/topology was prepared using both NJ and UPGMA method. The distances utilized for this purpose were population, as well as inter-individual distance. The genetic distances were least among Tarai and Bhadawari populations and were closely related and this may be attributed due to the fact that the two populations are in geographical contiguity.

Vijh et al. (2005) reported on comparative evaluation of three buffalo populations viz; Bhadawari Tarai and Kerala buffaloes using microsatellite markers. Data generated on 24 loci from these populations was subjected to analysis for estimation of genetic distances. The genetic distances calculated were Nei's minimum, Nei's standard, Latter's Fst, Reynolds, distances (IAM). Genetic distances based on SMM were also calculated. These were average squared distance (ASD), delta mu sequare $[(\delta\mu)^2]$, Shrivers' Dsw and similarity index (D_R). The dendrograms wre reparaed using both UPGMA and NJ algorithms. The Tarai and Bhadawari populations were close to one another and this was expected because of the contiguity of their breeding tract resulting in increased gene flow. The results of all the genetic distances were similar except ASD, which is not the right method for closely related populations of a species. The inter-individual distances were also estimated and clustering of individuals was done utilizing UPGMA and NJ methods.

2.11.2 Linkage analysis and gene mapping

Microsatellite marker used for mapping of QTL by linkage. This information of gene of economic important can be used in breeding programmes of either within breed's manipulation like marker-assisted selection of young sires or between breeds introgression programmes. Microsatellite also sever as reference point for mapping gene of interest and further in transgenic breeding. Microsatellite are also helpful for making strategy in moving the genes across the breeding barriers of animals. Microsatellite can be also used to study the effects and level of inbreeding (Beaumont and Bruford 1999, Pemberton et al.1999, Sweigart et al. 1999). The average heterozygosity of an individual measured from microsatellite data should realistically reflect the level of inbreeding.

Nahas et al. (1999), examined ninety-nine loci in river buffalo chromosomes, 67 of which are coding genes and 32 of which are anonymous DNA segments (microsatellites). Sixty-seven assignments were based on co-segregation of cellular markers in somatic cell hybrids (synteny), whereas 39 were based on in situ hybridization of fixed metaphase chromosomes with labeled DNA probes. Seven loci were assigned by both methods. Of the 67 assignments in somatic cell hybrids, 38 were based on polymerase chain reaction (PCR), 11 on isozyme electrophoresis, 10 on restriction endonuclease digestion of DNA, 4 on immunofluorescence, and 4 on chromosomal identification. A genetic marker or syntenic group has been assigned to each arm of the five sub-metacentric buffalo chromosomes as well as to the 19 acrocentric autosomes, and the X and Y chromosomes. These same markers map to the 29 cattle autosomes and the X and Y chromosomes, and without exception, cattle markers map to the buffalo chromosome or chromosomal region predicted from chromosome banding similarity.

VanHoof et al. (1999) used bovine autosomal microsatellite markers for population genetic studies on African buffalo were investigated. A total of 168 microsatellite markers were tested for PCR amplification on a test panel of seven African buffalo. Amplification was observed for 139 markers (83%), and 101 markers were studied further with 91 (90%) being polymorphic. The mean number of alleles per marker was 5·0 (*SE* = 0·2) and the mean heterozygosity per marker was 0·61 (*SE* = 0·03). Considering the overall high level of polymorphism, it was concluded that most bovine microsatellite markers are applicable in African buffalo (*Syncerus caffer*).

Navani, (2000), utilised the one hundred and eight microsatellite primer pairs, originally identified for cattle, were evaluated for their applicability in buffalo. Eighty-one primer pairs (75%) amplified discrete products, and of these, 61 pairs (56%) gave polymorphic band patterns on a panel of 25 buffaloes. The mean number of alleles per polymorphic marker was 4.50 +/- 0.20, and the mean heterozygosity per polymorphic marker was 0.66 +/- 0.02. Successful genotyping of buffaloes using cattle specific primers suggested that the latter could be a valuable resource for genome analysis in bubaline species.

Sreekumar et al. (2002), a 4400-bp genomic sequence and a 332-bp truncated cDNA sequence of the interleukin-2 (IL-2) gene of Indian water buffalo (Bubalus bubalis) were amplified by polymerase chain reaction and cloned. The coding sequence of the buffalo IL-2 gene was assembled from the 5' end of the genomic clone and the truncated cDNA clone. This sequence had 98.5% nucleotide identity and 98% amino acid identity with cattle IL-2. Three amino acid substitutions were observed at positions 63, 124 and 135. Comparison of the predicted protein structure of buffaloes IL-2 with that of human and cattle IL-2 did not reveal significant differences. The putative amino acids responsible for IL-2 receptor binding were conserved in buffalo, cattle and human IL-2. The amino acid sequence of buffalo IL-2 also showed very high identity with that of other ruminants, indicating functional cross-reactivity.

Iannuzzi et al. (2003), sixty-four genomic BAC-clones mapping five type I (ADCYAP1, HRH1, IL3, RBP3B and SRY) and 59 type II loci, previously FISH-mapped to goat (63 loci) and cattle (SRY) chromosomes, were fluorescence in situ mapped to river buffalo R-banded chromosomes, noticeably extending the physical map of this species. All mapped loci from 26 bovine syntenic groups were located on homeologous chromosomes and chromosome regions of river buffalo and goat (cattle) chromosomes, confirming the high degree of chromosome homeologies among bovids. Furthermore, an improved cytogenetic map of the river buffalo with 293 loci from all 31 bovine syntenic groups is reported.

2.11.3 Individual Identification of Parentage Testing

The detection of the hyper variable sequences, the PCR based microsatellite typing provide the powerful tool for the identity or paternity testing (Hagelberg et al., 1991). With the selected microsatellite loci, the multiplex PCR system and electrophoresis in one gel lane, forms a highly discriminating, extremely powerful for parentage testing (Heyen et al., 1997; Peelman et al., 1998; Luikat et al., 1999). However, in the case of exclusions based on microsatellite

polymorphism, special attention must be given to whether the offspring and the parent in question are homozygous for the alleles. In these cases, it cannot be excluded that the non-paternity / non-maternity would be incorrectly diagnosed, due to the allele non- amplification. The precision of the allele designation across the gel is sufficient for the five or six loci, thus allowing the comparison between the newly analyzed and the stored samples. The microsatellites have been successfully and extensively employed for the parentage testing and individual identification and for breed allocation etc., in various domestic animals.

Tonhati et al. (2000) investigated genetic trends of some productive and reproductive traits in a herd of Murrah buffalo raised in Sao Paulo, Brazil. Variance components for milk production (MP), length of lactation (LL), calving interval (CI) and age of first calving (AFC) were estimated by the restricted maximum likelihood method, using an animal model. Estimated heritability values were 0.38; 0.01; 0.10 and 0.20 for MP, LL, CI and AFC, respectively. Estimated repeatability values were 0.50, 0.13 and 0.20 for MP, LL and CI, respectively. Means of predicted breeding values for cows; dams and sires according to calving year and the genetic correlations were presented.

Microsatellite maker are also used as a diagnostic tool to detect diseases (Weiss Enbach et al., 1992). Now day's microsatellite are regularly used in population and ecological studies. Microsatellite are excellent marker for studying gene flow, effective population size (Ne), dispersal and migration related issues and parentage and relatedness (Taylor et al., 1994; Caulson et al., 1998; Ciofi and Bruford, 1999; Goldstein et al., 1999; Luikart and England, 1999).

2.12 SOME TECHNICAL DRAWBACKS OF MICROSATELLITES

2.12.1 Null Alleles

Failure of amplification of some alleles due to mutations in the binding regions results in reduction or loss of PCR products. These are termed as null alleles and may lead to serious underestimation of heterozygosity. Heterozygotes may be misclassified as homozygotes when null-alleles occur due to mutation in the primer annealing sites. This can lead to serious underestimation of heterozygosity, compared with that expected on the basis of Hardy Weinberg equilibrium (Callen et al., 1993, Paetkau and Stroback, 1995). In a heterozygote of two different alleles, if one allele fails to amplify due to primer annealing difficulties then the phenotype will appear as a single banded homozygote. This problem may be overcome by designing new primers but it a very tedious task. The use of homologous primers reduces the occurrence of null alleles but Callen et al. (1993) identified null alleles using homologous primers.

2.12.2 Slippage

Problem associated with the PCR process itself; Taq polymerase generates slippage during PCR and the tendency of Taq polymerase to add an additional dATP to PCR products can

sometimes make allele scoring problematic (Ginot et al. 1996, Gill et al., 1997). The activity of the Taq DNA polymerase used in the PCR. During PCR amplification, the thermo-polymerase tends to 'slip leading to production of differently sized products. These are less intense and also referred to as shadow bands. Further the Taq polymerase has a tendency to add an additional ATP at the 3' end of the amplified PCR products. This can also lead to difficulties in scoring bands.

2.12.3 Homoplasy

Homoplasy can be defined as the co-occurrence of alleles that are identical by descent. If two alleles are inherited without any mutation from the same ancestral allele they are identical by descent. But two alleles may have the same structure and even the same sequence but may not have been inherited from the same ancestral allele. Such alleles are identical in state. Homoplasy, therefore, helps in estimating the actual divergence between populations. Homoplasy due to different forward and backward mutations may underestimate genetic divergence. The two fragments of the same length are not derived from the same ancestral sequence introducing the possibility of size homoplasy. Under the IAM there should not be any homoplasy, but SMM and TPM can generate size homoplasy.

2.12.4 Linkage Disequilibrium

Linkage disequilibrium is the non-random association of alleles between two loci on a chromosome. This may be caused by epistatic natural selection and random genetic drift (Ohta, 1982, 1982a). This could create opportunities as well as difficulties in gene mapping. In such case, the presence of all alleles at one locus would predict the location of the allele at the other locus making one locus redundant for mapping purposes. The non-random association between alleles of linked marker reflects the size of chromosomal segments remaining intact in a population. Linkage disequilibrium has proven to be powerful tool for mapping of disease genes for monogenic disorders (Mohlke et al., 2001).

Some of the other problems may also associated with microsatellite marker such as stutter bands on gels may complicate accurate scoring of polymorphisms, underlying mutation model (infinite alleles model or stepwise mutation model) largely unknown and high development costs in case primers are not yet available.

One perceived difficulty with microsatellites is the long lead-time in identifying and characterizing microsatellites in new taxanomic groups. This problem is partially alleviated, however, by the continuing popularity of microsatellites in genetic mapping. Another practical long-term difficulty with microsatellite markers is the requirement of determining fragment lengths, which would seem to complicate automation. Ultimately the future may belong to markers amenable to yes/no tests which can be set up on dense chips, as for example single nucleotide polymorphism (SNP). In contrast with their importance in intraspecific studies, microsatellites have yet to make any real contribution to phylogeny reconstruction. Although it is not yet entirely clear why microsatellites have not been more successful in reconstructing phylogenies, part of the difficulty certainly stems from restrictions to divergence imposed by

range constraints, irregularities and asymmetries in the mutation process, and the degradation of microsatellites over time.

Conclusively, the microsatellites are the genetic markers that can be useful in addressing the questions at a variety of scales. Most specially, this genetic tool can help in solving the problems ranging from the individual specific, such as the questions of relatedness and parentage, the genetic structure of populations, the comparison among breeds/populations/species to the linkage analysis and gene mapping. Further it has several technical and analytical advantages that make it superior to the genetic markers whose domains are far smaller. Thus, microsatellites are markers of the choice for the genome analysis studies.

3. MATERIALS AND METHODS

3.1 MATERIAL REQUIRED

3.1.1 Chemicals

Ammonium Chloride (1M), Potassium bicarbonate (1M), EDTA (0.5M), Sodium (1M), Tris-Cl (1M) pH 8.0, Proteinase K, SDS (20%), Phenol (Tris equilibrated) pH 8.0, Chloroform, Isoamyl alchohol, Sodium acetate (3M) pH 5.2, Chilled absolute Ethanol, T.E. buffer (pH 8.0), 70% Ethanol, *Taq* Polymerase Enzyme, Deoxyribose nucleotides, PCR Buffer, Primers, Liz-standard, Sequencing Buffer, POP4 and Hi-Di Formamide. All reagents used in the study were of molecular biology grade.

3.1.2 Equipments

Refrigerated Centrifuge (Sigma 3K30), Rotar Plate Centrifuge (Sigma 4-15), Thermocycler (ABI Perkin Elmer System and Bio-Rad), Tube Centrifuge (Sigma 4-15), pH Meter (Thermo-Orion), Electronic Balance (Sartorius), C24KC Refrigerated Incubator Shaker (Edison, NJ.USA), Magnetic Strrier (Biosan), Vortex Shaker CM101 (Remi Equipments) Horizontal Electrophoresis Unit (Bio-Rad), Automated DNA Sequencer (ABI System Avant- 3100), UV Transilluminator and Microwave Oven.

3.2 METHODS PERFORMED

3.2.1 Blood Sample Collection

Two populations of buffalo Marathwada and Surti were selected and 44 blood samples from each breed were collected at random from field conditions from the breeding tract for these populations (fig. 8 and 9). 10 ml of whole blood was collected aseptically from juglar vein of each animal using heparinised vacutainer tubes and transported to laboratory at 0-5^{o}C.

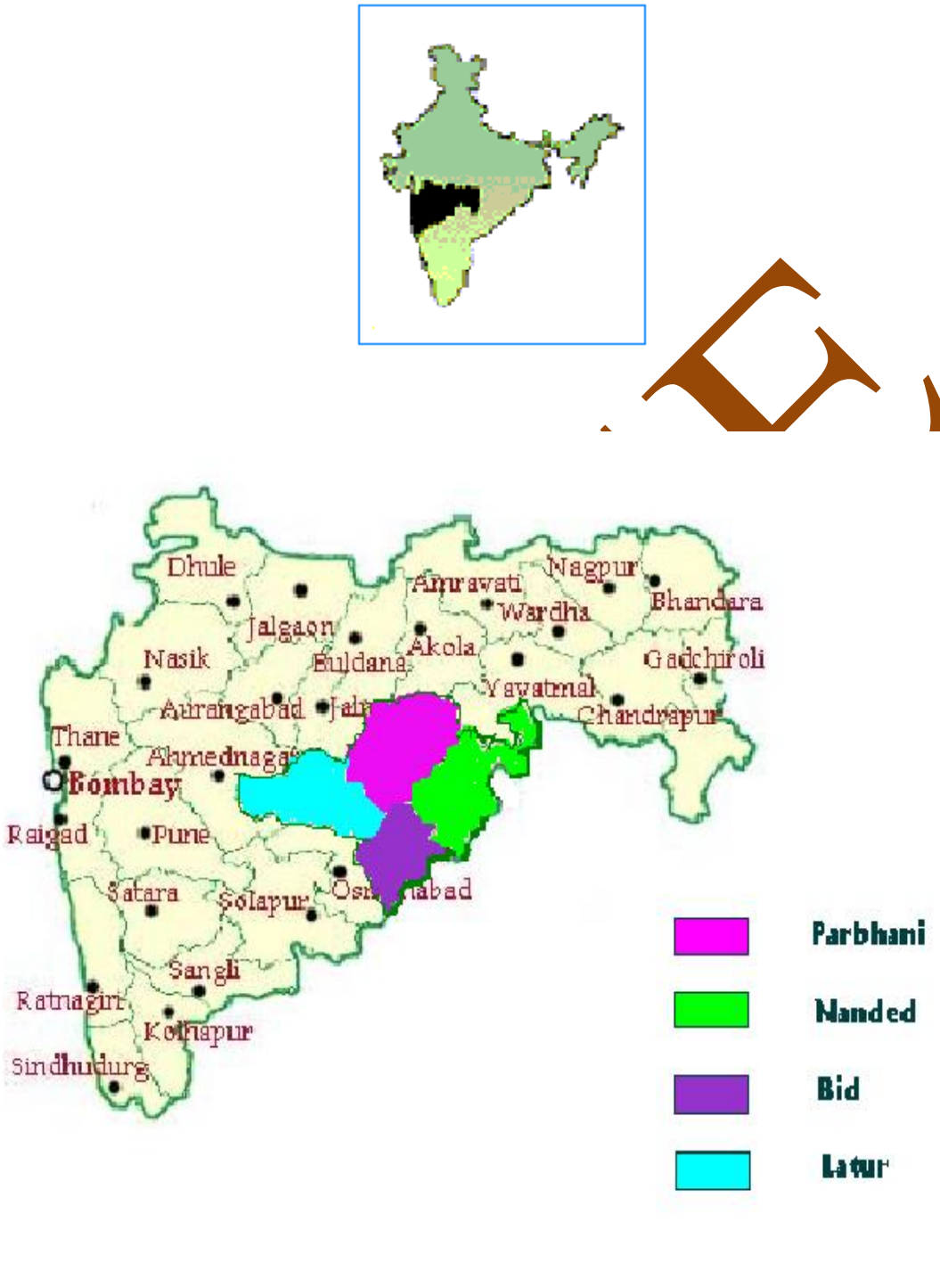


Fig. 4 Geographical Distribution And Breeding Tract Of Marathwada Buffaloes

(IJIASE) 2016, Vol. No. 2, Jan-Dec e-ISSN: 2454-9258, p-ISSN: 2454-809X

Fig. 5 Geographical Distribution And Breeding Tract Of Surti Buffaloes

3.2.2 DNA Isolation

For isolation of DNA from collected sample, the blood was transferred to autoclaved Oakridge centrifuged tubes. The RBCs were lysed with lysis buffer (Ammonium chloride 155 mM, Potassium bicarbonate 10 mM and EDTA 0.1 mM). In all the blood samples, double amount of lysis buffer was added. The samples were then mixed gently and kept in ice for 10 minutes. The tubes were centrifuged at 12,000 rpm for 10 minutes at 4^{o}C in the refrigerated centrifuge. The supernatant was carefully decanted and the pallet was redissolved in lysis buffer and washed three times as described above or until white pallet was obtained.

The white pellet was resuspended in 10 ml digestion buffer (Sodium chloride 75mM, Tris-Cl 1 M, pH8.0 and EDTA 0.5 mM). Vortexed gently and to it added 20% Sodium sulphate (200 µl/sample) and Proteinase K (1 mg/sample) and incubated at 57^{o}C overnight in a water bath. After incubation, digested solution was obtained to which equal amount of Tris equilibrated phenol (pH 8.0) was added, mixed gently by moving the tubes gently '8' fashions for 10 minutes and centrifuge at 12,000 rpm for 10 minutes at 25^{o}C. After centrifugation the aqueous phase and organic phase were separated. The DNA remained dissolved in aqueous phase while the protein was retained at the interphase. The aqueous phase was transferred carefully by Pasture pipette to another Oakridge tube without disturbing the interphase. To the aqueous phase, equal amount of phenol: chloroform: isoamyl alcohol (25:24:1) was added and mixed by moving the tubes gently in '8' fashions for 10 minutes at 25^{o}C to separated aqueous phase and organic phase. Again the aqueous phase was carefully pipetted to another tube without disturbing the interphase. The aqueous phase was mixed with chloroform: isoamyl: alcohol (24:1). The solution was subjected to centrifugation at 12,000 rpm for 10 minutes at 25^{o}C and aqueous phase was carefully transferred to glass culture tubes. The DNA was precipitated by adding 1/10th volume of Sodium acetate (3M, pH 5.2) and 2.5 volume of aqueous phase of chilled absolute alcohol and cover the tubes with paraffin and mix slightly, the DNA get precipitated (visible as white stringy strands).

3.2.3 Washing and Storage of Isolated DNA

The DNA was spooled out into eppendrof tubes washed twice with 70% ethanol to remove the salts. After washing the alcohol was allowed to evaporate and DNA was redissolved in 500 µl Tris EDTA buffer (Tris 10mM, EDTA 10mM, pH 8.0) and kept out eppendrof tubes at 65^{o}C for 1 hour. The stock DNA was stored at –20^{o}C.

3.2.4 Estimation of DNA quantity and purity

DNA quantification can be done using spectrophotometric measurement of UV absorption at wavelengths 230, 260 and 280 nm. Measures of DNA purity can be determined by the A_{260}:A_{280} and A_{260}:A_{230} ratios. These ratios provide indications of protein, and polyphenol and carbohydrate contamination, respectively (Manning, 1991). The DNA should show a clear

absorbance peak at 260 nm. The A_{260} value provides a measure of concentration (roughly 1.0 reading at A_{260} is equivalent to 50 mg per ml). A pure DNA solution has an A_{260}:A_{280} ratio of 1.8 ± 0.1.

The concentration of unknown double stranded DNA samples was estimated using the formula:

$$\text{DNA concentration (µg/µl)} = \frac{\text{OD 260 X dil. Factor X 50 µg/ml}}{1000}$$

Quantification of the DNA can be achieved by running the DNA samples on 0.6 % agarose gel stained with ethidium bromide (0.5 mg per ml) (Ethidium bromide is a fluorescent chemical that intercalates between base pairs in a double stranded DNA molecule). Aliquots of the DNA extracts are loaded [e.g. 6 µl of DNA mix (1 µl of each DNA sample and 5 µl of 1X loading dye)] alongside a range [7 µl each of 5, 10, 25 and 50 ng per µl] of uncut lambda DNA standards. High molecular weight DNA appeared as a well-resolved band alongside the lambda DNA bands whilst the smearing below the band indicates mechanical or chemical degradation. A smeared band towards the bottom of the gel is an indication of the presence of RNA in the extract. A rough estimate of DNA content (± 50 ng/µl) may be obtained by comparing band intensities of the DNA extract and the standard by eyes.

3.3 AGAROSE GEL ELECTROPHORESIS FOR DILUTION OF GENOMIC DNA

The technique of electrophoresis is based on the fact that DNA is negatively charged at neutral pH due to its phosphate backbone. For this region, when an electrical potential is placed on the DNA, it will moves toward the positive pole.

3.3.1 Dissolving of Agarose

- Prepared the 1X TAE buffer by diluting 50X TAE buffer.
- Weighed out 0.9 gm of agarose (0.6%) and add 150 ml 1X TAE.
- In microwave oven heat the slurry for 3 minutes in one minute incriminates, swirling the solution gently between heating cycles to release trapped air and resuspend any agarose particles caught on the side of the Erlenmeyer flask.
- Ethidium bromide was added to the warm solution and cooling the solution about 50^{o}C. [Ethidium bromide is intercalating dye and carcinogenic, handle this gel only while wearing gloves].

3.3.2 Casting the gel

After cooling the solution it was poured into a casting tray containing a sample comb and allowed to solidify at room temperature. The gel was 3-5 mm thick. There was no air bubbles

under or between the teeth of comb. In case, if bubbles form, they can be removed by pocking them with the pointed end of pipette tips before the gel has set. After the gel has solidified, the comb was removed carefully by wriggling back and forth gently and then lifting up carefully, not rip the bottom of the well.

3.3.3 Preparing the sample

While the gel was cooling, prepared the DNA sample by adding 5 µl of tracking dye to 1 µl of each DNA sample. The tracking dye contains Bromophenol Blue and Xylene cynol FF and 50% glycerol. Adding tracking dye to the sample increased its density so it falls into the well of the gel and provided a visible marker to monitor the progress of electrophoresis. Also prepared molecular size standard by mixing 5 µl of the 1 kb ladder with 1 µl of tracking dye. Bromophenol Blue migrated through agarose gel approximately 2.2 fold faster than Xylene cynol FF dye, independent to agarose gel concentration.

3.3.4 Loading and running the gel

The gel on the tray, inserted horizontally into the electrophoresis chamber and flooded the top of the gel with fresh running buffer (1X TAE) to cover the gel to depth of about 1 mm. Sucked the solution (DNA sample with dye) into the pipette, placed the tip in the top of the well and gently expelled the solution into the well. The lid and power leads were placed on the apparatus, 40-80 V current is applied, and current flowing was confirmed by observing bubbles coming off from the electrodes. Run the gel until the Bromophenol blue and Xylene cynol FF was migrated the appropriate distance about three-fourth of the way across the gel.

3.3.5 Visualization Of DNA

Bright orange color DNA bands were visualized by placing the tray onto High Performance UV Transilluminator.

3.4 POLYMERASE CHAIN REACTION AMPLIFICATION OF MICROSATELLITE LOCI

Polymerase chain reaction (PCR), is the major scientific development of the last quarter century, sometime referred to as "**molecular photocopying**", that can characterize, analyze, and synthesize any specific piece of DNA or RNA. Polymerase chain reaction is an in vitro method for analyzing defined sequence of DNA. Kary Mullis, while working Cetus Corporation, California invented PCR in 1985 for which he was awarded the Nobel Prize in chemistry in 1993.

There are three major steps in PCR, were carried out in the same vial but at different temperatures and are repeated for 30 cycles. This was done on an automated thermocycler, which automatically heated and cooled the tubes with the reaction mixture in a very short time.

Denaturation at 94^{o}C: During the denaturation, the first part of the process, the double strand melted, open to single stranded DNA, all enzymatic reactions stop (e.g., the extension from a previous cycle).

Annealing at 55-60°C: The primers were jiggling around, caused by the Brownian motion. Ionic bonds were constantly formed and broken between the single stranded primer and the single stranded template. The polymerase enzyme attached on double stranded DNA (template and primer) and started copying the template. The vial is cooled to 55-60°C. At this temperature, the primers bind or "anneal" to the end of the DNA strands.

Extension at 72°C: This was the ideal working temperature for the polymerase. Primers that were on positions with no exact match get loose (because of the higher temperature) and didn't give an extension of the fragment.

3.4.1 Criteria For Selection of Primers

The primers were selected on the basis of recommendation given by diversity analysis group of Food and Agriculture Organization (Barker, 1995). The main criteria for selection of microsatellite loci for biodiversity analysis were:

1. The followed Mendalian heritance.
2. They were polymorphic in nature with minimum three alleles.
3. The microsatellite loci information is available in public domain.
4. They are suitable for cross amplification among related species.
5. They should be at linkage equilibrium.
6. They should be representative of whole genome.

Table-1a: List of Cattle Primers Screened in Buffalo Breeds:

Primer	Sequences	Repeats	Dye	Accession Number
ILSTS87	*F- AgCAgACATgATgACTCAgC* *R-CTgCCTCTTTTCTTgAgAgC*	$(CA)_{14}$	Ned	L37279
ILSTS59	F- AgTATggTAAggCCAAAggg R- CgACTTgTgTTgTTCAAAgC	$(gT)_{21}$	Vic	L37266
ILSTS62	F- CTgTCCTTTAAgAACAAACC R- TgCAACTTAggCTATTgACg	$(CA)_{11}$	Pet	L37222
CSSM29	F- TCTCCATTATgCACATgCCATgCT R- CgTgAgAACCgAAAgCACACATTC	$(AC)_{18}$	Ned	U03807
ILSTS11	F- GCTTgCTACATggAAAgTgC R- CTAAAATgCAgAgCCCTACC	$(CA)_{11}$	Vic	L23485
BM1818	F- AgCTgggAATATAACCAAAgg R- AgTgCTTTCAAggTCCATgC	$(Tg)_{13}$	Pet	G18391
ILSTS72	F- ATgAATgTgAAAgCCAAggg R- CTTCCgTAAATAATTgTggg	$(CA)_{14}$	Ned	L37272
ILSTS49	F- CAATTTTCTTgTCTCTCCCC R- gCTgAATCTTgTCAAACAgg	$(CA)_{9}$	Pet	L37261
ILSTS05	F- GgAAgCAATgAAATCTATAgCC	$(Tg)_{9}$	Vic	L23481

	R- TgTTCTgTgAgTTTgTAAgC			
ILSTS58	F- GCCTTACTACCATTTCCAgC R- CATCCTgACTTTggCTgTgg	$(gT)_{15}$	Vic	L37225
CSSM43	F- AAAACTCTgggAACTTgAAAAACTA R- gTTACAAATTTAAgAgACAgAgTT	$(CA)_{19}$	Ned	U03824
CSSM45	F- TAgAggCACAAgCAAACCTAACAC R- TTggAAAgATgCAgTAgAACTCAT	$(CA)_{14}$	Ped	NW_381320

Table-1b: List of Cattle Primers Screened in Buffalo Breeds:

Primer	***Sequences***	**Repeats**	**Dye**	**Accession Number**
CSSM47	F- TCTCTgTCTCTATCACTATATggC R- CTgggCACCTgAAACTATCATCAT	$(gT)_{12}$	Ned	U03821
ILSTS30	F- CTgCAgTTCTgCATATgTgg R- CTTAgACAACAggggTTTgg	$(gT)_{10}$	Vic	L37212
CSSM08	F- CTTggTgTTACTAgCCCTggg R- gATATATTTgCCAgAgATTCTgCA	$(Tg)_{15}$	Pet	NW_375905
CSSM33	F- CACTgTgAATgCATgTgTgTgAgC R- CCCATgATAAgAgTgCAgATgACT	$(Tg)_{16}$	Pet	U03805
CSRM60	F- AAgATgTgATCCAAgAgAgAggCA R- AggACCAgATCgTgAAAggCATAg	$(CA)_{17}$	Vic	AF232758
ETH152	F- TACTCgTAgggCAggCTgCCTg R- gAgACCTCAgggTTggTgATCAg	$(CA)_{17}$	Ned	Z14040
CSSM19	F- TTgTCAgCAACTTCTTgTATCTTT R- TgTTTTAAgCCACCCAATTATTTg	$(Tg)_{18}$	Vic	AF232761
CSSM06	F- AgCTTCTgACCTTTAAAgAAAATg R- AgCTTATAgATTTgCACAAgTgCC	$(Tg)_{13}$	Vic	U03787
ILSTS29	F- TgTTTTgATggAACACAgCC R- TggATTTAgACCAgggTTgg	$(AC)_{19}$	Ned	L37252
CSSM57	F- TgTggTgTTTAACCCTTgTAATCT R- gTCgCTggATAAACAATTTAAAgT	$(gT)_{16}$	Pet	U03840
ILSTS38	F- GggCATTATTTTgTTTCCC R- CCACTTCTgggTAATTATCC	$(gT)_{14}$	Pet	L37256

3.4.2 Taq DNA Polymerase

Taq DNA polymerase is obtained from the thermophilic archaea family bacteria Thermus aquaticus. It possesses a 5'→ 3' polymerase activity and a double strand specific 5'→ 3' exonuclease activity.

3.4.3 PCR Buffer

PCR Buffer (Sigma) contains Tris–Cl (100 mM, pH 8.3 at 25°C), KCl (500 mM), $MgCl_2$ (15 mM), gelatin (0.01%). It supports the activity of Taq polymerase.

3.4.4 Magnesium Chloride concentration

Magnesium concentration is a crucial factor affecting the performance of Taq DNA polymerase. Reaction components, including template DNA, chelating agents present in the sample (e.g., EDTA), dNTPs can affect the amount of free magnesium. In the absence of adequate free magnesium, Taq DNA polymerase is inactive. Conversely, excess free magnesium reduces enzyme fidelity and may increase the level of nonspecific amplification. The optimal $MgCl_2$ concentration is use for each reaction.

3.5 PRE-PCR PREPARATION

The PCR was performed under standard condition as described by Kaul et al., 2001. The genomic DNA was diluted so as to contain about 50 ng DNA/µl.

3.5.1 Cocktail Preparation

Table 2: The cocktail for PCR one reaction (14 µl) consisted of:

	Volume	Concentration
10X PCR Buffer	1.5 µl	-
DNTPs	0.1 µl	200 µM
Primer (Forward)	1.0 µl	4 pmol
Primer (Reverse)	1.0 µl	4 pmol
Taq Polymerase	0.08 µl	0.4 Unit
Distilled Water	10.32 µl	-
Total	14 µl	

Beside this 2 mM $MgCl_2$ (nuclease free) was added separately for ETH152, CSSM08, ILSTS38 primers and 1.5 mM $MgCl_2$ was added for ILSTS29 primer.

3.5.2 PCR programme

1.0 µl of the template DNA was directly added into the cocktail in each lane of the PCR plate. The PCR was carried out in Applied Bio-system and Bio-Rad thermocycler.

Table 3: For the specific conditions, the PCR machine is programmed at:

Steps	*Programme*	Temperature	Time	Number of cycle
Step- 1	Initial Denaturation	95°C	5 minutes	1 cycle
Step- 2	Denaturation	94°C	45 seconds	30cycles
Step- 3	Annealing	55°C (or standardized)	45 seconds	
Step- 4	Polymerization	72°C	45 seconds	
Step- 5	Final Extension	72°C	5 minutes	1 cycle

Step- 6	Final Temperature	4°C	Until removed from PCR machine

3.6 AGAROSE GEL ELECTROPHORESIS OF PCR AMPLIFIED DNA

After completion of the PCR programme, the products were checked on 2% agarose gel for the amplification. Before loading into the well, gel-loading dye (xylene cynol FF, bromophenol blue in glycerol) was added to the sample and the samples were run under constant voltage conditions (80 V) till the two dyes were separated. Amplified product appeared as sharp orange color bands under UV Transilluminator due to the intercalation of ethidium bromide.

3.7 POST PCR MULTIPLEXING

3.7.1 Pooling Ratios

PCR products of different sizes and dyes were pooled for one capillary injection for maximizing the throughput. It was important to pool PCR products together at the correct ratios, in order to get similar florescent intensities across all loci in the pooling. The fluorescent dyes were detected with different efficiencies. The pooling ratio, or amount of each dye-labeled added with respect to the other products in the pool and was adjusted to ensure an appropriate detection of all the loci.

Dilution series was carried on each combination of pooling to determine the optimal fluorescent intensity, which gives clean sharp peaks and correct banding patterns, which was allowed for accurate allele calling. The loci in a multiplex if overlapping, were labeled with different dyes.

After determining the optimal pooling ratio and dilution ratio for a set of primers, the same dilutions were used for subsequent analyses, as PCR yields were fairly consistent. In a pool the following ratio was used:

Vic labeled PCR product - 1.5 µl
Ned labeled PCR product - 2.0 µl
Pet labeled PCR product - 2.0 µl

3.7.2 Size Standard

Gene Scan- 500 LIZ™ Size Standard or ROX standard developed by Applied Biosystems was used for fragment sizing, LIZ size standard yield size fragments between 50- 500bp poviding16 single – standard labeled fragments of 35, 50, 75, 100, 139, 150, 160, 200, 250, 300, 340, 350, 400, 450, 490, and 500 bases. Each of the DNA fragments labeled with a proprietary fluorophore, which results in a single peak when run under denaturing condition. Internal lane size standard was run with every sample for accurate sizing. The genotyping reaction components were:

Pooled PCR Product - 1 µl

Hi-Di Foramide - 8.75 µl
Liz Size standard - 0.25 µl

3.8 DENATURATION

The above components were mixed well and denatured at 95ºC for 5 minutes. The 96 well plate was loaded in Automated DNA sequencer for genotyping.

3. 9 GENOTYPING

Automated DNA Sequencer- ABI 3100 Avant, which was provided with 4 capillaries, performed Genotyping and different array sizes can be used. We used 36 cm array size for performing genotyping. The large surface area of a capillary allowed heat generated during electrophoresis to be dissipated efficiently, allowing high voltage electrophoresis. The result was rapid, high resolution separation of DNA fragments. Polymer POP 4 (Performance Optimized Polymer) was used for sizing and separating of DNA fragments. Plate records were prepared and size standard was added to Automated DNA sequencer prior to setup of the run and then 96 well plate was linked and started the run.

3.10 DATA COLLECTION AND EXTRACTION

The data extracted from Automated DNA sequencer were sized using Gene Scan software and the sizes were extracted using Gene Mapper software version 3.1. The data was exported as text file and imported into excel sheet before submitting it to further Statistical Analysis (fig. 11 and 12).

3.11 STATISTICAL ANALYSIS OF DATA

The data generated using the microsatellite loci were subjected to statistical analysis. The microsatellite data was subjected to Ewens-Watersson test of Neutrality to check for the loci being selectively neutral.

The statistical analysis was carried out using POPGENE software (Yeh *et al*., 1999). The heterozygosity measures were calculated using the following formulae given by Nei (1978).

3.11.1 Allele number

Alleles are a set of alternative forms of the same gene occupying the same relative position or locus on homologous chromosomes. Allele number is the total number of alleles for a given marker/locus in a population, which is counted with a non-zero frequency. The allele number for each locus can be determined manually from the silver stained gels/autoradiograms.

3.11.2 Allele Frequency/Gene Frequency

The frequency of an allele ‘A’ is the number of ‘A’ alleles in the population divided by the total number of allele/genes. It gives the indication of the most or least prevalent alleles in the

population. The allele frequency is affected over time by force such as genetic drift, mutation and migration.

3.11.3 Effective Allele Number (Kimura & Crow, 1964)

The effective number of alleles (n_e) is the reciprocal of the sum of the square of allele frequencies.

$$n_e = 1 / \Sigma P_i^2$$

where P_i is the frequency of the i^{th} allele.

3.11.4 Heterozygosity

Heterozygosity is the state of possessing different alleles at given locus in regard to given character. It is a measure of heterozygotes or genic variation in a population. The population heterozygosity at a locus is given by the formula:

$$H = 1 - \Sigma P_i^2$$

where, Σ stands for summation over all alleles (Nei, 1978) and P_i is the frequency of the i^{th} allele at a locus in a population. The average heterozygosity per locus (H) is defined as the mean of H overall structural loci in the genome.

However, the unbiased estimate of the expected heterozygosity at a locus is (if N <50): N is sample size**.**

$$H_E = \frac{2N}{2N-1}\left[1 - \sum_{i=1}^{n} \bar{P}_i^2\right]$$

3.11.5 Polymorphism Information Content (PIC)

Polymorphism Information Content a measure of informativeness of marker and the PIC value of each locus was calculated according to Botstein et al. (1980) using the given formula:

$$PIC = 1 - \left(\sum_{i=1}^{n} P_i^2\right) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2P_i^2 P_j^2$$

P_i is the frequency of the i^{th} allele at a locus in a population. P_i is the frequency of the j^{th} allele at a locus in a population.

3.11.6 Hardy Weinberg Equilibrium

The Hardy Weinberg equilibrium was tested using χ^2 and G^2 statistics.

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

where, O is observed heterozygosity and E is expected heterozygosity.

The log likelihood ratio was calculated as

$$G^2 = 2 \times O \times \ln (O/E)$$

where, O is observed heterozygosity and E is expected heterozygosity.

3.11.7 F-statistic

F-statistics can be thought of as a measure of the correlation of alleles within individuals and are related to inbreeding coefficients. An inbreeding coefficient is really a measure of the nonrandom association of alleles within an individual. As such, F- statistics describe the amount inbreeding-like effects within subpopulations, F_{IS} among subpopulations, F_{ST} and within the entire population. F_{IT} these can also be classified as IAM and SMM/TPM based estimators of F-statistics. Some of the more useful measures of population subdivision are the F-statistics developed by Wright (1965).

Inbreeding within a subpopulation is caused by the nonrandom mating of the members of that subpopulation, in that mating occurs more often than by chance alone, between closely related individuals. As closely related individuals will contain a large proportion of the same alleles due to common descent, their offspring will have a higher level of homozygosity, and conversely, a lower level of heterozygosity then expected. A within subpopulations F-statistic can be estimated from a ratio of the observed to expected heterozygosity where,

$$F_{IS} = \frac{\bar{H}_S - H_I}{\bar{H}_S}$$

where H_S is the average expected heterozygosity estimated from each subpopulation by,

$$H_S = 1 - \sum_{i=1}^{k} p_i^2,$$

and H_I is the average observed heterozygosity,

$$H_I = \sum_{i=1}^{k} \frac{H_i}{k}$$

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

for **k** subpopulations.

Population substructure will also lead to inbreeding-like effects, i.e. a reduction in observed heterozygosity when compared to expected. This effect is known as Wahlunds' effect. This relationship shows that as allele frequencies in two subpopulations deviate, the average expected heterozygosity in those populations will always be less than that expected from the pooled allele frequencies. H_T an among subpopulations F-statistic can be estimated from this ratio.

$$F_{ST} = \frac{H_T - \bar{H}_S}{H_T},$$

where,

$$H_T = 1 - \sum_{i=1}^{k} \bar{p}_i^2,$$

and P_i is the frequency of the **i**th allele averaged over all subpopulations. It should be noted that as allele frequencies deviate, the difference in H_S and H_T will increase and F_{ST} will therefore also serve as a measure of genetic distance among sub-populations.

The measure of the correlation of alleles for the entire population is thus a combination of both the within and among subpopulation effects, and can be estimated from,

$$F_{IT} = \frac{H_T - H_I}{H_T}.$$

Slatkin (1995) proposed measures of average sum of squares of the difference in allele size are related to F_{ST}. From coalescent theory it can be shown that F_{ST} is related to the average coalescence times, such that,

$$F_{ST} = \frac{\bar{t} - t_0}{\bar{t}},$$

where $\bar{t}$ is the average time to coalescence of any two randomly chosen alleles from the entire population, and t_0 is the average time to coalescence of any two randomly chosen alleles drawn from the same subpopulation (Slatkin 1991). In two populations with divergent alleles, the average time to a common ancestor for any two alleles in a subpopulation is going to be less than the average time to coalescence of any 2 alleles drawn from the entire population. Based on

this relationship, Slatkin (1995) developed an estimation of F_{ST} that would incorporate the mutational history contained within the alleles under the SMM, $R_{ST.}$

$$R_{ST} = \frac{\bar{S} - S_W}{\bar{S}},$$

where S_W and S are the average sum of squares of the difference in allele size within a subpopulation and the entire population as describe above. Slatkin noted that the E (S_W) = $2\mu t_0\sigma^2_m$ and that the E (S) = $2\mu t\sigma^2 m$. Thus,

$$\frac{E(\bar{S}) - E(S_W)}{E(\bar{S})} = \frac{\bar{t} - t_0}{\bar{t}} = F_{ST}.$$

3.11.8 Number of Migrants

The number of migrants were estimated using F_{ST} values as follows:

$$\mathbf{F_{ST} = 0.25(1\text{-}F_{ST})/F_{ST}}$$

3.11.9 Genetic Distance

The most commonly used distance measure was developed by Nei (1987). Nei's standard genetic distance D_S is calculated from

$$D_S = -\ln I,$$

where I is a measure of genetic identity. Identity is estimated from

$$I = \frac{J_{xy}}{\sqrt{J_x J_y}},$$

where Jxy, Jx and Jy are the means for all loci of $\Sigma x_i y_i$, Σx_i^2 and Σy_i^2 for each locus.

In present case since there were only two populations we calculated interindividual distance among the member of the two populations.

3.11.10 Phylogenetic Tree

We utilized the genetic distance matrix of interindividual distance to prepare phylogenetic tree utilizing Neighbour Joining (NJ) and UPGMA algorithm.

3.12 TESTING BUFFALO POPULATIONS FOR MUTATION DRIFT EQUILIBRIUM

3.12.1 Qualitative Method

The method is graphical representation of the genetic bottleneck. The data of allele frequency was classified into 10 allelic frequency classes and then plotting a frequency histogram. The 10 allelic frequency classes were 0.001-0.100, 0.101-0.200, ……..0.901-1.00. The low frequency allelic classes were 0.001- 0.100 and high frequency allelic classes were

0.901-1.000 the rest were termed as intermediate classes. The above classification system was arbitrary but suited graphical qualitative assessment of allele frequency distributions. The classification of the data in large number of allelic classes might have defied the very purpose of meaningful assessment of the distribution of allele frequencies. This graphical method concludes that a population has been recently bottlenecked if fewer alleles are found in low frequency class than in one or more intermediate frequency classes.

3.12.2 Quantitative Methods

The population bottlenecks induce a transient excess of heterozygosity. Any population, that has experienced a recent bottleneck, will show higher than the expected (equilibrium) heterozygosity for a large majority of loci. A simple test for this shall be (observed - expected) heterozygosity across all loci in a population sample.

Establishing distribution of heterozygosity under the null hypothesis of mutation drift equilibrium. The probability distributions are actually conditioned by the number of alleles. Consequently the parameter θ is known, since a range of values of θ is compatible with the observed number of alleles. All the three models of microsatellite evolution i.e Infinite Allele Model. (IAM), Stepwise Mutation Model (SMM) and Two Phase Model (TPM) have different properties and thus different solutions were applied to the three models.

For the IAM, the number of alleles is a non-decreasing function of the number of mutations in the genealogy of the sampled genes i.e. a new mutation increase the number of different alleles by zero (when it hits an edge already bearing a mutation) or one (when it hits a new edge). Therefore, to obtain a simulated sample with exactly K_0 alleles, we first simulated a genealogy of the n genes in a population at mutation - drift-equilibrium (with edge lengths following usual exponential law). Then one mutation at one time was added till such time there were exactly K_0 different alleles in the sample. It is known that the mutations are distributed according to a Poisson law; the probability of one mutation affecting a given edge is proportional to the relative length of this edge. The distribution of heterozygosity in samples simulated as explained above is θ free.

The same rationale could not however be applied to the SMM, because the number of alleles can increase, decrease or remain unchanged when adding a new mutation. Therefore the classical simulation process of the coalescent process (assuming mutation drift equilibrium) was used with two modifications.

i) The iterations which lead to the observed number of alleles were considered (the number K_o actually observed in the sample) and

ii) For each iteration the value of θ was taken at random following a probability distribution defined according to the Bayesian approach. Assuming a uniform prior distribution of θ, the condition $k= k_0$ is accounted for by considering a posterior distribution of θ proportional to the probability of getting k_0 alleles given θ. This was achieved by simulating the coalescent

process of a sample of n genes using a set of values of θ and counting for each θ, the proportion of iteration in which $k = k_0$.

The various steps performed were as follows:

(a) Extremes of the distribution of the likelihood of θ given k_0 and n, starting with $\theta_{min} = 10^{-7}$, 200 iterations were performed. If none of them produced the correct number of alleles, a new θ_{min}. is taken by multiplying the previous by 10. This process is repeated till at least one (iteration out of 200) produced the observed number of alleles. The final minimum is taken as the last θ_{min} divided by 10. An analogue was conducted to get maximum (θ_{max}).

(b) Defining a step between two successive values of θ in such a way that θ_{min} - θ_{max} covered at different values of θ (preferably 12) so as it provided a compromise of both precision and speed of execution.

(c) For each of the 12 θ's, the proportion of iterations (out of 1000 in present study), which gave exact K_o alleles, was computed.

After these three steps the coalescent process was simulated taking at random one of the 12 proceeding θ's with a probability proportional to the proportion computed in the last step (c) disregarding all the iteration that produced a number of alleles different than k_0. For each locus, the average heterozygosity and standard deviation was estimated through the above simulation process with a level of precision controlled by the number of simulation replications.

The sample size in terms of haploid genomes, observed heterozygosity and number of alleles observed were calculated. The heterozygosity was calculated using the formula of Nei (1987) for each locus.

3.12.3 Sign-Test

The 23 polymorphic microsatellite loci with 40 individuals of Marathwada and 48 individuals of Surti buffalo per population (Marathwada and Surti) and 44 samples for Marathwada buffaloes and with no bottleneck (mutation drift equilibrium), there is approximately an equal probability of getting a positive or negative difference between the observed and the expected heterozygosities. In contrast if there had been a recent bottleneck, a positive difference (in the number of loci with heterozygosity excess) is likely to be observed more often than a negative difference. Here we applied a simple test if the number of loci for which there is heterozygosity excess is significantly larger than 12 assuming an *apriori* binomial distribution of parameters 23 loci (L) and ½ (12).

In a finite sample, heterozygosity can have only a finite number of values and its probability distribution shall be discrete and asymmetric. Consequently for any locus there is a specific probability (slightly different from 0.5) of heterozygosity excess. This probability can however, be computed if we know the theoretical distribution of the heterozygosity in a sample of n individuals assuming mutation drift equilibrium. This distribution was established using 1000

simulations under all the possible models [Infinite Allele Model (IAM), Stepwise Mutation Model (SMM) and Two Phase Model (TPM)].

The probability Pr_L (l) of having l loci among L, for which there is a heterozygosity excess, utilizing usual rationale in which the drawing of additional locus is considered as a Markovian process in which states are the number of loci showing an heterozygosity excess (0,1,2…..L). Starting with zero loci and a probability distribution Pr_o (l) equal to (1,0,0.…0) the recurrence relationship was estimated for m = 1 to L.

$$Pr_m [l\ ] = [(1-p_{lm})\ pr_{m-1}(l\ )] + [p_{lm}\ pr_{m-1}(l\ -1)\ ]$$

where, plm is the probability of heterozygosity excess at locus m.

$$\text{Let } Q = \sum_{l=o}^{L} prL(l\ )$$

where, Q is the probability of getting at least l_o loci with heterozygosity excess. If Q is less than 0.05, the null hypothesis (mutation drift equilibrium) is rejected in favour of overall heterozygosity excess and a recent genetic bottleneck hypothesis, the difference between the observed (H_o) and the expected heterozygosity.

3.12.4 Standardized Differences Test

The sign-test does not take into account the magnitude of the heterozygosity excess/deficiency. Under the null heterozygosity (H_e) is the output of a random variable with an expectation equal to zero for all loci. If we divide these differences by a σ (standard deviation) of the corresponding distributions of heterozygosities, we can get standardized deviates. The sum of these standardized deviates approaches a Gaussian distribution with null expectation and variance equal to L, when L increases by virtue of the Central Limit theorem. Thus the test consists in comparing the statistics T_2 to N (0,1) distribution.

$$T_2 = L^{(-)\,0.5} \sum_{l=1}^{L} [\Delta l\ \sqrt{\sigma_l}\ ]$$

With $\Delta l = (H_o-H_e)$ for the l^{th} locus and σl is the standard deviation of the distribution of H at the l^{th} locus. If the alternative hypothesis was heterozygosity excess (One tailed test, we rejected the null hypothesis of population at mutation drift equilibrium) at 5% level if T_2 values were larger than or equal to 1.645 (value from table of normal distribution).

3.12.5 Wilcoxon Rank Test

A Wilcoxon sign Rank (Luikart, 1997) for the median difference in this type of paired data to be zero consists of sorting the ranks to the absolute values (rank 1 to the smallest, rank 2 to the next smallest and so on) and then finding the sum of the ranks of the positive differences. If the null hypothesis is true, the sum of the ranks of positive differences should be about the same as the sum of the ranks of negative differences. The probabilities values P value of the test can be estimated. Here the empirical values for each of the locus were having a comparable value

found from the distribution of heterozygosity under the null hypothesis. The Wilcoxon rank test is more powerful than the previous two tests i.e; Sign Rank test and standardised differences test.

If n (>8), Z_+ than the distribution is approximately normal and the mean and variance were calculated as:

$$\mu = \frac{n(n+1)}{4}$$

$$\sigma^2 = \frac{n(n+1)(2n+1)}{24}$$

Sum of Z values = $Z_{(+)} + Z_{(-)} = \frac{n(n+1)}{2}$

Test Statistic W is calculated as $W = \frac{Z_+ - \mu}{\sigma}$

The significant can be tested as supposing w is test statistics

Prob (W≥w) ≤0.05 : Reject the null hypothesis; for one tailed test.

Prob (W|≥|w|) = 2* Prob (W≥w) ≤0.05; for two tailed test.

4. RESULTS AND DISCUSSION

A total of 23 loci were utilized in study. All the loci selected were found to be polymorphic. The 23 loci were not specific for the *Bubalus bubalis* but we are taken from the gene bank accession of microsatellite loci for cattle available in the NCBI database. These primers were amplified in the related species of buffaloes and thus were heterologous in nature, since this represented cross species amplification. The sizes of alleles are likely to be different from the size range available for cattle. The loci were amplified and scored for understanding the level of polymorphism in buffalo and to estimate gene and genetic diversity among Marathwada and Surti buffaloes.

4.1 NUMBER OF ALLELES AND ALLELE FREQUENCY

The amplified PCR products were sized using Avant 3100 Automated DNA Sequencer with Liz 500 was taken as internal size standard. The sizing of the alleles was extrapolated from the regression curve drawn using Liz squares and genetic algorithm using Gene Mapper software (V-3.0). Since most of the loci taken in the study were dinucleotide in nature, the

mutation rate is relatively of higher magnitude compared to tri and tetra nucleotide repeats. The accession number, repeat and product range of each locus is given in table no. 4. The details of the parameters for the loci are given locus wise.

Table No.4a: Standarised PCR conditions, repeats, allelic range and primer sequence is label for each microsatellite loci used in the present study.

Locus	**Sequences**	**Dye Label**	**Repeats**	**A.T.**	**Alleles**	**Allele size**	**Acc No**	**$MgCl_2$ Conc.**	*Primer Cycle*
ILSTS87	*F- AgCAgACATgATgACTCAgC R-CTgCCTCTTTTCTTgAgAgC*	Ned	$(CA)_{14}$	55^0C	5	102-114	L37279	1.5 mM	Step 1 (1x) - 95- 5’ Step 2 (30x)- 95- 45’’ 55- 45’’ 72- 45’’ Step 3 (1x) - 72- 5’ Step 4 - 10- ∞
ILSTS59	F-AgTATggTAAggCCAAAggg R-CgACTTgTgTTgTTCAAAgC	Vic	$(gT)_{21}$	55^0C	11	151-187	L37266	1.5 mM	Step 1 (1x) - 95- 5’ Step 2 (30x)- 95- 45’’ 55- 45’’ 72- 45’’ Step 3 (1x) - 72- 5’ Step 4 - 10- ∞
ILSTS52	F-CTgTCCTTTAAgAACAAACC R-TgCAACTTAggCTATTgACg	Pet	$(CA)_{11}$	55^0C	10	138-178	L37222	1.5 mM	Step 1 (1x) - 95- 5’ Step 2 (30x)- 95- 45’’ 55- 45’’ 72- 45’’ Step 3 (1x) - 72- 5’ Step 4 - 10- ∞
CSSM29	F-TCTCCATTATgCACATgCCATgCT R-CgTgAgAACCgAAAgCACACATTC	Ned	$(AC)_{18}$	60^0C	10	156-186	U03807	1.5 mM	Step 1 (1x) - 95- 5’ Step 2 (30x)- 95- 45’’ 60- 45’’ 72- 45’’ Step 3 (1x) - 72- 5’ Step 4 - 10- ∞
ILSTS11	F-GCTTgCTACATggAAAgTgC R-CTAAAATgCAgAgCCCTACC	Vic	$(CA)_{11}$	58^0C	6	246-270	L23485	1.5 mM	Step 1 (1x) - 95- 5’ Step 2 (30x)- 95- 45’’ 58- 45’’ 72- 45’’ Step 3 (1x) - 72- 5’ Step 4 - 10- ∞

Table No.4b: Standarised PCR conditions, repeats, allelic range and primer sequence is label for each microsatellite loci used in the present study.

Locus	**Sequences**	**Dye Label**	**Repeats**	**A.T.**	**All eles**	**Allele size**	**Acc No**	**$MgCl_2$ Conc.**	*Primer Cycle*
BM1818	F-AgCTgggAATATAACCAAAgg R-	Pet	$(Tg)_{13}$	58^0C	16	226-288	G18391	1.5 mM	Step 1 (1x) - 95- 5’ Step 2 (30x)- 95- 45’’ 58- 45’’

	AgTgCTTTCAAggTCCATgC								72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ILSTS72	F- ATgAATgTgAAAgCCAAggg R- CTTCCgTAAATAATTgTggg	Ned	$(CA)_{14}$	55^0C	4	137-149	L37272	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ILSTS49	F- CAATTTTCTTgTCTCTCCCC R- gCTgAATCTTgTCAAACAgg	Pet	$(CA)_9$	55^0C	6	136-198	L37261	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ILSTS05	F- GgAAgCAATgAAATCTATAgCC R- TgTTCTgTgAgTTTgTAAgC	Vic	$(Tg)_9$	55^0C	6	165-197	L23481	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ILSTS58	F- GCCTTACTACCATTTCCAgC R- CATCCTgACTTTggCTgTgg	Vic	$(gT)_{15}$	55^0C	15	114-146	L37225	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞

Table No.4c: Standarised PCR conditions, repeats, allelic range and primer sequence is label for each microsatellite loci used in the present study.

Locus	**Sequences**	**Dye Label**	**Repeats**	**A.**	**Alleles**	**Allele size**	**Acc No**	**$MgCl_2$ Conc.**	*Primer Cycle*
CSSM43	F- AAAACTCTgggAACTTgAAAAACTA R- gTTACAAATTTAAgAgACAgAgTT	Ned	$(CA)_{19}$	55^0 C	8	219 - 257	U03824	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
CSSM45	F- TAgAggCACAAgCAAACCTAACAC R- TTggAAAgATgCAgTAgAACTCAT	Ped	$(CA)_{14}$	58^0 C	10	98-130	NW_381320	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 58- 45'' 72- 45'' Step 3 (1x) - 72- 5'

									Step 4 - 10- ∞
CSSM47	F-TCTCTgTCTCTATCACTATATggC R-CTgggCACCTgAAACTATCATCAT	Ned	$(gT)_{12}$	55^0 C	17	126 - 164	U03821	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ILSTS30	F- CTgCAgTTCTgCATATgTgg R- CTTAgACAACAggggTTTgg	Vic	$(gT)_{10}$	55^0 C	8	144 - 166	L37212	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
CSSM08	F- CTTggTgTTACTAgCCCTggg R-gATATATTTgCCAgAgATTCTgCA	Pet	$(Tg)_{15}$	55^0 C	6	181 - 193	NW_375905	2 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞

Table No.4d: Standarised PCR conditions, repeats, allelic range and primer sequence is label for each microsatellite loci used in the present study.

Locus	**Sequences**	**Dye Label**	**Repeats**	**A.T.**	**Alleles**	**Allele size**	**Acc No**	**$MgCl_2$ Conc.**	*Primer Cycle*
CSSM33	F-CACTgTgAATgCATgTgTgTgAgC R-CCCATgATAAgAgTgCAgATgACT	Pet	$(Tg)_{16}$	58^0C	10	150-176	U03805	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 58- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
CSRM60	F-AAgATgTgATCCAAgAgAgAggCA R-AggACCAgATCgTgAAAggCATAg	Vic	$(CA)_{17}$	60^0C	12	88-136	AF232758	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 60- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ETH152	F-TACTCgTAgggCAggCTgCCTg R-gAgACCTCAgggTTggTgATCAg	Ned	$(CA)_{17}$	60^0C	9	190-216	Z14040	2 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 60- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
CSSM19	F-TTgTCAgCAACTTCTTgTATC	Vic	$(Tg)_{18}$	55^0C	6	125-145	AF232761	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45''

	TTT R- TgTTTTAAgCCACCCAATTATTTg								55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
CSSM06	F- AgCTTCTgACCTTTAAAgAAAATg R- AgCTTATAgATTTgCACAAgTgCC	Vic	$(Tg)_{13}$	55^0C	13	187-221	U03787	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞

Table No.4e: Standarised PCR conditions, repeats, allelic range and primer sequence is label for each microsatellite loci used in the present study.

Locus	**Sequences**	**Dye Label**	**Repeats**	**A.T.**	**Alleles**	**Allele size**	**Acc No**	**$MgCl_2$ Conc.**	*Primer Cycle*
ILSTS29	F- TgTTTTgATggAACACAgCC R- TggATTTAgACCAgggTTgg	Ned	$(AC)_{19}$	60^0C	7	150 - 164	L37252	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 60- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
CSSM57	F- TgTggTgTTTAACCCTTgTAATCT R- gTCgCTggATAAACAATTTAAAgT	Pet	$(gT)_{16}$	60^0C	11	114 - 138	U03840	1.5 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 60- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞
ILSTS38	F- GggCATTATTTTgTTTCCC R- CCACTTCTgggTAATTATCC	Pet	$(gT)_{14}$	55^0C	5	154 - 164	L37256	2 mM	Step 1 (1x) - 95- 5' Step 2 (30x)- 95- 45'' 55- 45'' 72- 45'' Step 3 (1x) - 72- 5' Step 4 - 10- ∞

CSSM06

A total of 13 alleles were obtained for this locus. The size range varied from 187-221. The allelic frequencies in the two populations have been depicted in figure 13a. The Marathwada population had 3 alleles specific to it. These were of product size 187, 199 and 209. The Surti population also had 3 alleles private to it. They were of size 213, 217 and 221.

CSSM19

A total of 6 alleles were obtained in the study. The size range varied from 125-145. The allelic frequencies in the two populations have been exhibited in figure 13a. The Marathwada and Surti population had no private alleles.

CSSM57

A total of 11 alleles were obtained for this locus. The size range varied from 114-138. The allelic frequencies in the two populations have been depicted in figure 13a. The Marathwada population had 2 alleles specific to it. These were of product size 114 and 122. The Surti population however, had 3 alleles private to the population. They were of size 128, 132 and 138.

ILSTS 38

A total of 5 alleles were obtained in the study. The size range varied from 154-164. The allelic frequencies in the two populations have been presented in figure 13a. The Marathwada population had 3 alleles specific to it. These were of product size 158, 160 and 164. The Surti population had no private alleles.

ILSTS29

A total of 7 alleles were obtained in the study. The size range varied from 150-164. The allelic frequencies in the two populations have been shown in figure 13a. The Marathwada population had no private allele. The Surti population also had 3 alleles private to the population. They were of size 150, 154 and 162. Rest of the 4 alleles occurred in both the populations.

ILSTS30

A total of 8 alleles were obtained for this locus. The size range varied from 144-166. The allelic frequencies in the two populations have been depicted in figure 13a. The Marathwada population had no private allele. The Surti population also had 2 alleles private to it. They were of size 144 and 152. Rest of the 4 alleles occurred in both the populations.

ETH152

A total of 9 alleles were obtained in the study. The size range varied from 190-216. The allelic frequencies in the two populations have been exhibited in figure 13a. The Marathwada population had 2 alleles specific to it. These were of product size 208 and 214. The Surti population however, had no alleles private.

CSSM47

A total of 17 alleles were obtained for this locus. The size range varied from 126-164. The allelic frequencies in the two populations have been depicted in figure 13a. The Marathwada

population had 4 alleles specific to the population. These were of product size 132, 148, 156 and 160. The Surti population also had 4 alleles private to it. They were of size 128, 144, 154 and 158. Rest of the 9 alleles occurred in both the populations.

CSSM33

A total of 10 alleles were obtained in the study. The size range varied from 150-176. The allelic frequencies in the two populations have been presented in figure 13b. The Marathwada population had 1 private allele with the product size166. The Surti population also had 3 alleles private to the population. They were of size 150, 160 and 176.

CSSM08

A total of 6 alleles were obtained for this locus. The size range varied from 181-193. The allelic frequencies in the two populations have been shown in figure 13b. The Marathwada population had 1 private allele with the product size 185. The Surti population also had 1 allele private to the population with the product size 189.

CSRM60

A total of 12 alleles were obtained for this locus. The size range varied from 88-136. The allelic frequencies in the two populations have been depicted in figure 13b. The Marathwada population had 3 alleles specific to the population. These were of product size 94, 104 and 136. The Surti population also had 1 allele private to it. The product size was 126.

CSSM43

A total of 8 alleles were obtained in the study. The size range varied from 219- 257. The allelic frequencies in the two populations have been exhibited in figure 13b. The Marathwada population had 1 allele specific to the population. The product size was 219. The Surti population also had 1 allele private to it. The product size was 247.

CSSM45

A total of 10 alleles were obtained in the study. The size range varied from 98-130. The allelic frequencies in the two populations have been shown in figure 13b. The Marathwada population had 3 alleles specific to it. These were of product size 126, 128 and 130. The Surti population also had 1 allele private with the product size 100. Rest of the 6 alleles occurred in both the populations.

ILSTS05

A total of 6 alleles were obtained for this locus. The size range varied from 165-197. The allelic frequencies in the two populations have been presented in figure 13b. The Marathwada population had 2 alleles specific to it. These were of product size 165 and 181. The Surti population also had 1 allele private with the product size 197.

ILSTS49

A total of 6 alleles were obtained in the study. The size range varied from 136-198. The allelic frequencies in the two populations have been depicted in figure 13b. The Marathwada population had 1 allele specific to it with the product size 140. The Surti population also had 3 alleles private to the population. They were of size 160, 188 and 198.

ILSTS58

A total of 15 alleles were obtained for this locus. The size range varied from 114-146. The allelic frequencies in the two populations have been deducted in figure 13b. The Marathwada population had 3 alleles specific to it. These were of product size 144, 118 and 120. The Surti population also had 3 alleles private to it. They were of size 124, 126 and 132. Rest of the 9 alleles occurred in both the populations.

ILSTS72

A total of 4 alleles were obtained for this locus. The size range varied from 137-149. The allelic frequencies in the two populations have been shown in figure 13c. The Marathwada population had 1 allele specific to the population. The product size was 146. The Surti population also had 1 private allele with the product size141.

BM1818

A total of 16 alleles were obtained in the study. The size range varied from 226-288. The allelic frequencies in the two populations have been presented in figure 13c. The Marathwada population had 6 alleles specific to the population. These were of product size 226, 228, 232, 246, 250 and 276. The Surti population however, had 1 allele private with the product size 288. Rest of the nine alleles occurred in both the populations.

CSSM29

A total of 10 alleles were obtained for this locus. The size range varied from 156-186. The allelic frequencies in the two populations have been depicted in figure 13c. The Marathwada population had 4 alleles specific to the population. These were of product size 156, 160, 170 and 180. The Surti population also had 1 private allele with the product size 164.

ILSTS11

A total of 6 alleles were obtained in the study. The size range varied from 246-270. The allelic frequencies in the two populations have been exhibited in figure 13c. The Marathwada population had 2 alleles specific to it. These were of product size 246 and 252. Surti population however, had no private allele.

ILSTS52

A total of 10 alleles were obtained for this locus. The size range varied from 138-178. The allelic frequencies in the two populations have been shown in figure 13c. The Marathwada population had 1 allele specific to the population. The product size was 156. The Surti population also had 2 private alleles with sizes 138 and 142.

ILSTS59

A total of 11 alleles were obtained for this locus. The size range varied from 151- 187. The allelic frequencies in the two populations have been depicted in figure 13c. The Marathwada population had 4 alleles specific to the population. These were of product size151, 155, 183 and 185. The Surti population also had 2 alleles private to it. They were of size 179 and 187. Rest of the five alleles occurred in both the populations.

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

ILSTS87

A total of 4 alleles were obtained for this locus. The size range varied from 102-114. The allelic frequencies in the two populations have been presented in figure 13c. The Marathwada population had 2 private alleles with product size 102 and 114. The Surti population also however, had no private allele to the population.

The relative allele frequencies for the 23 loci in the two populations are exhibited

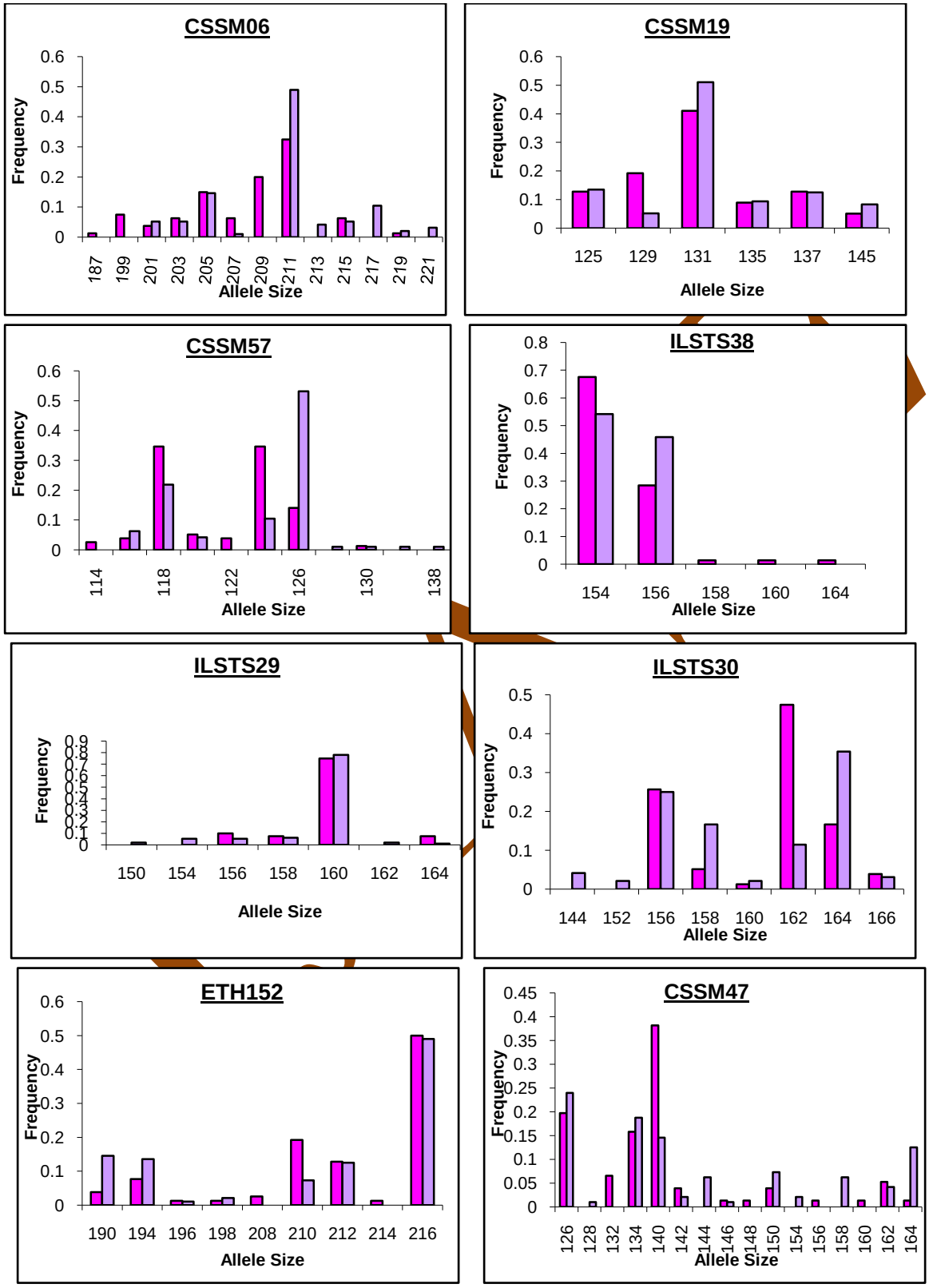
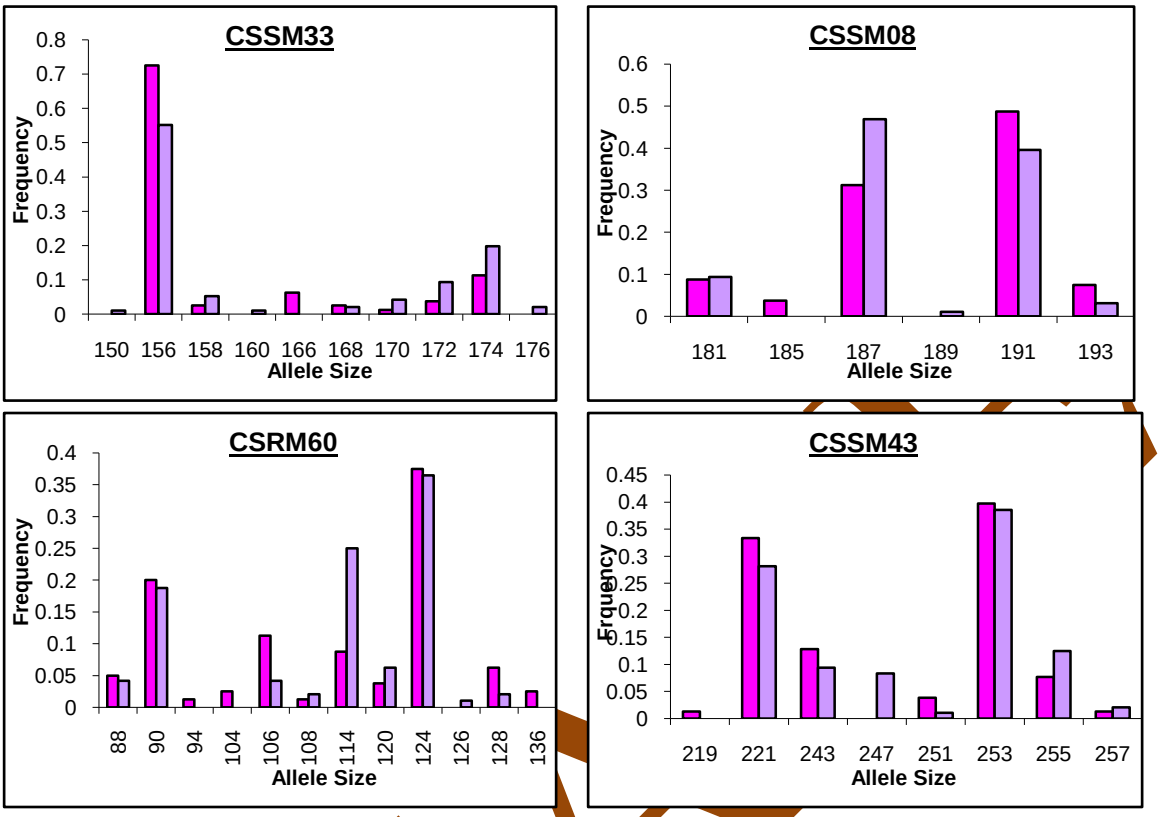

CSSM06
Frequency
Allele Size
187 199 201 203 205 207 209 211 213 215 217 219 221
CSSM19
Frequency
Allele Size
125 129 131 135 137 145
CSSM57
Frequency
Allele Size
114 118 122 126 130 138
ILSTS38
Frequency
Allele Size
154 156 158 160 164
ILSTS29
Frequency
Allele Size
150 154 156 158 160 162 164
ILSTS30
Frequency
Allele Size
144 152 156 158 160 162 164 166
ETH152
Frequency
Allele Size
190 194 196 198 208 210 212 214 216
CSSM47
Frequency
Allele Size
126 128 132 134 140 142 144 146 148 150 154 156 158 160 162 164

Fig. 5a Locus wise allele frequency distribution in Marathwada and Surti buffaloes.



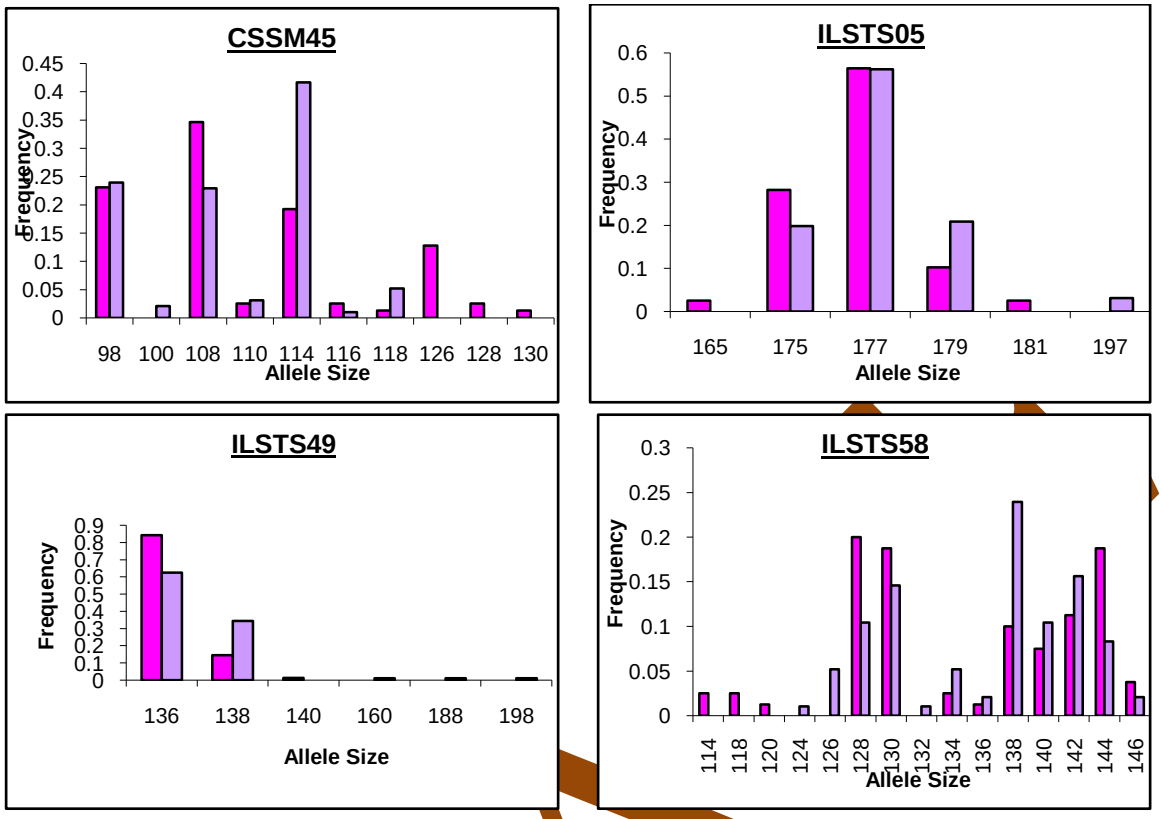
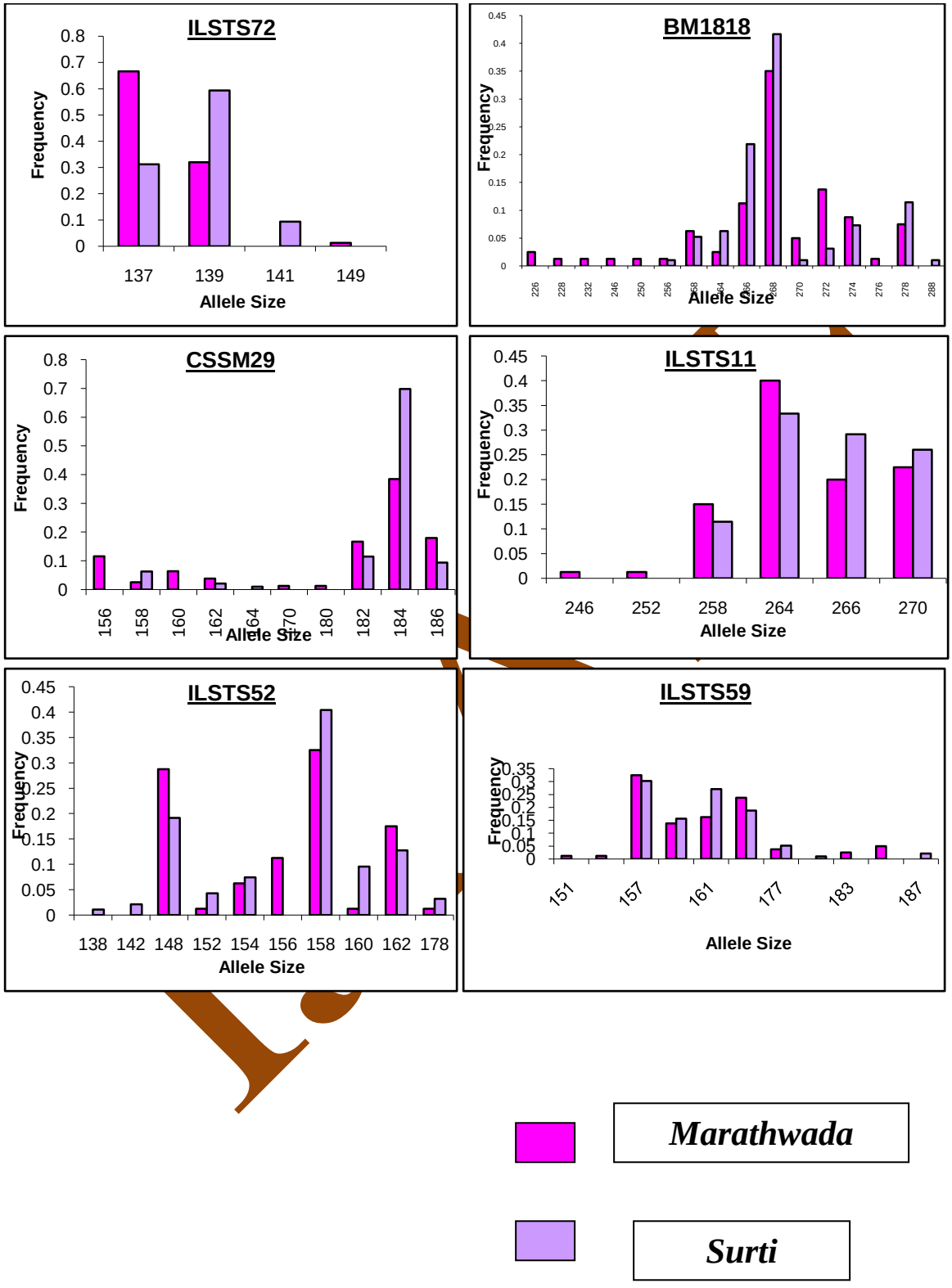


Fig. 5b Locus wise allele frequency distribution in Marathwada and Surti buffaloes


ILSTS72
Frequency
Allele Size
137
139
141
149
BM1818
CSSM29
ILSTS11
246
252
258
264
266
270
ILSTS52
138 142 148 152 154 156 158 160 162 178
ILSTS59
151
157
161
177
183
187
Marathwada
Surti

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

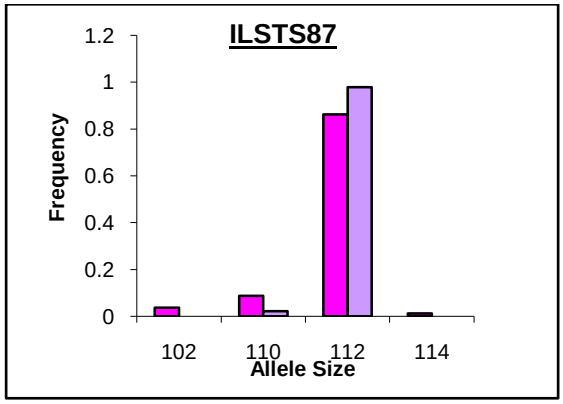


Fig. 5cLocus wise allele frequency distribution in Marathwada and Surti buffaloes

4.2 GENIC VARIATION

The genic variation of the two populations Marathwada and Surti are depicted in table no. 5. Number of alleles observed in each of the two populations has also been given in table no. 6 along with the effective number of allele of two populations. It is evident from the table that the effective number of alleles are almost half of the total number of alleles observed (fig. 14). This is due to the fact that a large number of alleles are present at very low frequency. The mean number of alleles (na) over 23 loci were found to be 7.52 while the mean effective number of alleles (ne) were 3.54 in Marathwada buffaloes and 6.96 and 3.38 for Surti buffaloes respectively. The effective numbers of alleles depicted the number that shall be available and shall not be lost from the population by chance. The Shannon's Information Index, which represents the relative abundance of information of a specific locus to the total information available over all loci. The Shannon's Information Index is thus a function of number of alleles observed for a specific locus.

Table No. 5: Genic Variation Statistics for all loci in pooled data set

Locus	Sample Size	na*	ne*	I*
CSSM06	176	13	4.6013	1.9794
CSSM19	174	6	3.6155	1.5374
CSSM57	174	11	3.9412	1.6086
ILSTS38	170	5	1.9751	0.7647
ILSTS29	176	7	1.6636	0.9102
ILSTS30	174	8	4.3789	1.6293
ETH152	174	9	3.3521	1.5426
CSSM47	172	16	6.3485	2.1512
CSSM33	176	10	2.3168	1.2925
CSSM08	176	6	2.7717	1.1972
CSRM60	176	12	4.6068	1.8345
CSSM43	174	8	3.6886	1.5338
CSSM45	174	10	4.1554	1.6461
ILSTS05	174	6	2.5050	1.1306
ILSTS49	172	6	1.7085	0.7044
ILSTS58	176	15	7.8262	2.2310
ILSTS72	174	4	2.2377	0.8919
BM1818	176	16	4.8340	1.9801
CSSM29	174	10	2.8264	1.4542
ILSTS11	176	6	3.6824	1.3834
ILSTS52	174	10	4.4367	1.7640
ILSTS59	176	11	4.6357	1.7242
ILSTS87	174	4	1.1639	0.3247
Mean	174	9.0870	3.6205	1.4442
St. Dev		3.6170	1.5625	0.4858

* na = Observed number of alleles
* ne = Effective number of alleles [Kimura and Crow (1964)]
* I = Shannon's Information index [Lewontin (1972)]

Table No. 6: Genic Variation of all loci in two buffalo populations

Locus	Marathwada Bufffaloes				Surti buffaloes			
	Sample Size	na*	ne*	I*	Sample Size	na*	ne*	I*
CSSM06	80	10	5.3422	1.9185	96	10	3.5310	1.6967
CSSM19	78	6	4.0185	1.5780	96	6	3.1956	1.4568
CSSM57	78	8	3.7602	1.5634	96	9	2.8818	1.4000
ILSTS38	74	5	1.8601	0.7968	96	2	1.9862	0.6897
ILSTS29	80	4	1.7131	0.8346	96	7	1.6112	0.8828
ILSTS30	78	6	3.0978	1.3349	96	8	4.3025	1.6631
ETH152	78	9	3.2055	1.5111	96	7	3.3247	1.4802
CSSM47	76	12	4.5338	1.8534	96	12	6.8776	2.1232
CSSM33	80	7	1.8349	1.0146	96	9	2.7910	1.4132
CSSM08	80	5	2.8571	1.2443	96	5	2.5888	1.0998
CSRM60	80	11	4.7548	1.8888	96	9	4.1853	1.6753
CSSM43	78	7	3.4103	1.4306	96	7	3.8528	1.5414
CSSM45	78	9	4.3707	1.6795	96	7	3.4751	1.4351
ILSTS05	78	5	2.4414	1.1014	96	4	2.5003	1.0793
ILSTS49	76	3	1.3694	0.4815	96	5	1.9642	0.8035
ILSTS58	80	12	7.0330	2.1293	96	12	7.2339	2.1556
ILSTS72	78	3	1.8270	0.6909	96	3	2.1787	0.8949
BM1818	80	15	5.6838	2.1296	96	10	4.0350	1.7145
CSSM29	78	9	4.3896	1.7307	96	6	1.9476	1.0226
ILSTS11	80	6	3.6571	1.4181	96	4	3.6085	1.3242
ILSTS52	80	8	4.2384	1.6121	94	9	4.2645	1.7380
ILSTS59	80	9	4.7128	1.7495	96	7	4.3969	1.6014
ILSTS87	80	4	1.3278	0.5186	94	2	1.0435	0.1030
Mean	79	7.5217	3.5408	1.4004	96	6.9565	3.3816	1.3476
St. Dev		3.1315	1.5027	0.4898		2.8521	1.4955	0.4746

* na = Observed number of alleles
* ne = Effective number of alleles [Kimura and Crow (1964)]

* I = Shannon's Information index [Lewontin (1972)]

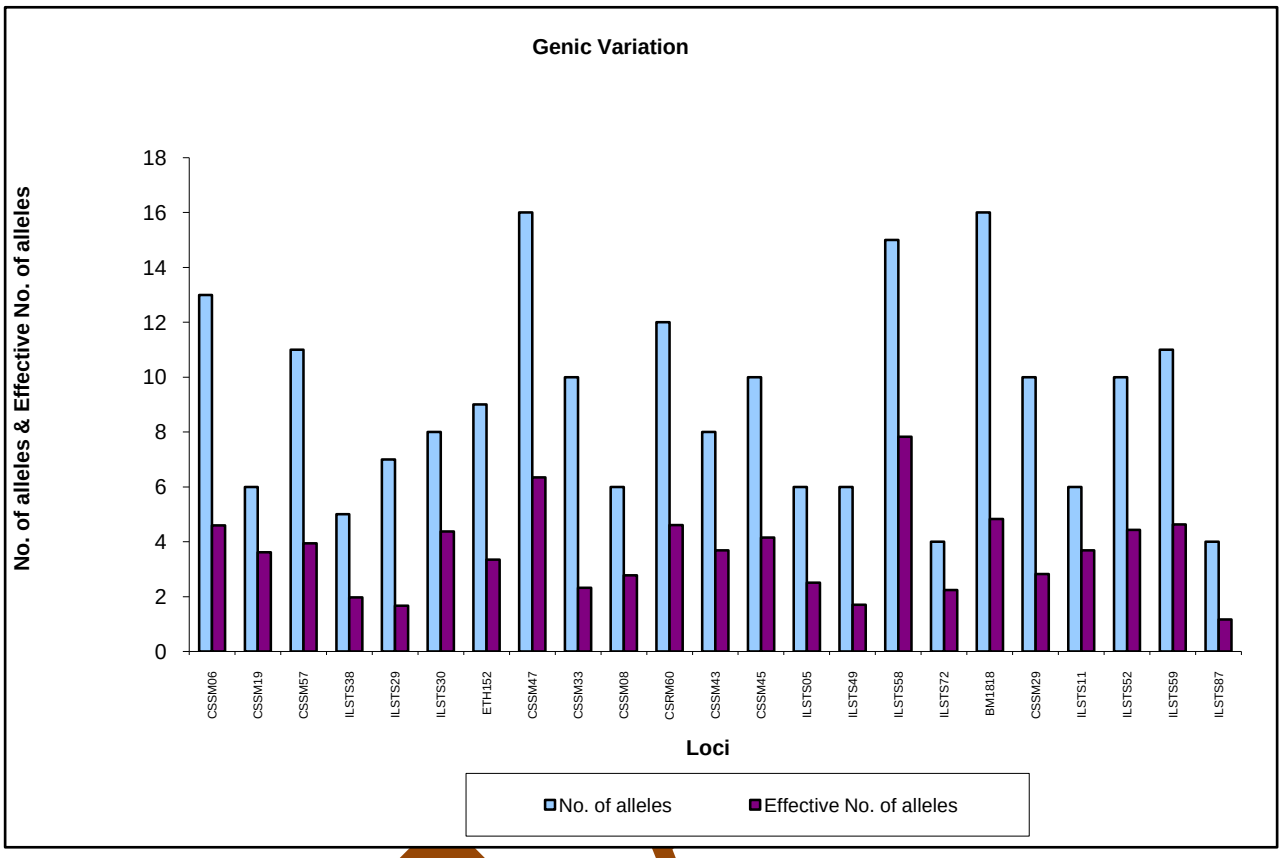


Fig. 6 Graphical representation of number of alleles and effective number of alleles of pooled data set of Marathwada buffaloes and Surti buffaloes

4.3 HETEROZYGOSITY AND POLYMORPHISM INFORMATION CONTENT

The observed homozygosity and heterozygosity, expected homozygosity and heterozygosity, average heterozygosity and the PIC are deducted in table no. 7 for pooled data set of two buffalo populations. The observed and expected homozygosity and heterozygosity are given in table no. 8 for Marathwada buffaloes. The average number of samples was 80 for Marathwada populations (applied genome). The observed heterozygosity ranged from 0.108 (ILSTS38) and 0.875 (ILSTS58). The mean observed heterozygosity was found to be 0.591 with a standard error of 0.025. The expected heterozygosity, which is a function of alleles and its frequency, was highest for locus ILSTS58 (0.87) and minimum for locus ILSTS87 (0.25). The PIC, which is estimated from deduction of expected homozygosity and cross homozygosity from 1.0. The PIC gives the information that the locus is likely to provide in terms of gene diversity. The values are slightly less then the expected heterozygosity values (fig. 15). The locus with least PIC was ILSTS87 (0.25) and highest PIC was ILSTS58 (0.86). The Nei's average heterozygosity also given in table no. 8 for Marathwada buffaloes. The PIC of these loci was 0.65, which is quit high, and thus reaffirm the utility of these 23 loci for genetic diversity analysis.

The observed and expected homozygosity and heterozygosity are given in table no. 9 for Surti buffaloes. The average number of samples was 96 for Surti population (applied genome). The observed heterozygosity ranged from 0.00 (ILSTS87) and 0.9167 (CSSM47). The mean observed heterozygosity was found to be 0.575 with a standard error of 0.024. The expected heterozygosity, which is a function of alleles and its frequency, was highest for locus ILSTS58 (0.87) and minimum for locus ILSTS87 (0.04). The PIC, which is estimated from deduction of, expected homozygosity and cross homozygosity from 1.0. The PIC gives the information that the locus is likely to provide in terms of gene diversity. The values are slightly less then the expected heterozygosity values (fig. 16). The locus with least PIC was ILSTS87 (0.04) and highest PIC was ILSTS58 (0.86). The Nei's average heterozygosity also given in table no. 9 for Surti buffaloes. The PIC of these loci is 0.64, which is quit high, and thus reaffirm the utility of these 23 loci for genetic diversity analysis.

Table No. 7: Heterozygosity estimates for all loci in pooled data set buffaloes as a measure of variability of microsatellite loci

Locus	Sample Size	Obs_Hom	Obs_Het	Exp_Hom*	Exp_Het*	PIC*	Ave_Het

CSSM06	176	0.3068	0.6932	0.2129	0.7871	0.7827	0.7648
CSSM19	174	0.4138	0.5862	0.2724	0.7276	0.7234	0.7191
CSSM57	174	0.4713	0.5287	0.2494	0.7506	0.7463	0.6935
ILSTS38	170	0.8824	0.1176	0.5034	0.4966	0.4937	0.4795
ILSTS29	176	0.6705	0.3295	0.5988	0.4012	0.3989	0.3978
ILSTS30	174	0.2644	0.7356	0.2239	0.7761	0.7716	0.7224
ETH152	174	0.3333	0.6667	0.2943	0.7057	0.7017	0.6936
CSSM47	172	0.2326	0.7674	0.1526	0.8474	0.8425	0.817
CSSM33	176	0.4545	0.5455	0.4284	0.5716	0.5684	0.5484
CSSM08	176	0.4091	0.5909	0.3571	0.6429	0.6392	0.6319
CSRM60	176	0.2273	0.7727	0.2126	0.7874	0.7829	0.7754
CSSM43	174	0.2529	0.7471	0.2669	0.7331	0.7289	0.7236
CSSM45	174	0.2529	0.7471	0.2363	0.7637	0.7593	0.7417
ILSTS05	174	0.5287	0.4713	0.3957	0.6043	0.6008	0.5952
ILSTS49	172	0.6047	0.3953	0.5829	0.4171	0.4147	0.3803
ILSTS58	176	0.1023	0.8977	0.1228	0.8772	0.8722	0.8598
ILSTS72	174	0.5517	0.4483	0.4437	0.5563	0.5531	0.4968
BM1818	176	0.4886	0.5114	0.2023	0.7977	0.7931	0.7881
CSSM29	174	0.4368	0.5632	0.3501	0.6499	0.6462	0.6294
ILSTS11	176	0.2955	0.7045	0.2674	0.7326	0.7284	0.7247
ILSTS52	174	0.2184	0.7816	0.2209	0.7791	0.7746	0.7648
ILSTS59	176	0.2841	0.7159	0.2112	0.7888	0.7843	0.7802
ILSTS87	174	0.9195	0.0805	0.8583	0.1417	0.1408	0.1443
Mean	174	0.4175	0.5825	0.3332	0.6668	0.6629	0.6466
St. Dev		0.2074	0.2074	0.1727	0.1727	0.1718	0.1708

* Observed homozygosty and heterozygosity
* Expected homozygosty and heterozygosity
* PIC Polymorphic Information Content

Table No. 8: Different estimates of heterozygosity in Marathwada buffaloes

Locus	Sample Size	Obs_Hom	Obs_Het	Exp_Hom*	Exp_Het*	PIC*	Avg_Het
CSSM06	80	0.3000	0.7000	0.1769	0.8231	0.8128	0.7648
CSSM19	78	0.3846	0.6154	0.2391	0.7609	0.7512	0.7191
CSSM57	78	0.4615	0.5385	0.2564	0.7436	0.7341	0.6935
ILSTS38	74	0.8919	0.1081	0.5313	0.4687	0.4624	0.4795

ILSTS29	80	0.6000	0.4000	0.5785	0.4215	0.4163	0.3978
ILSTS30	78	0.4103	0.5897	0.3140	0.6860	0.6772	0.7224
ETH152	78	0.3590	0.6410	0.3030	0.6970	0.6880	0.6936
CSSM47	76	0.4211	0.5789	0.2102	0.7898	0.7794	0.817
CSSM33	80	0.5250	0.4750	0.5392	0.4608	0.4550	0.5484
CSSM08	80	0.3000	0.7000	0.3418	0.6582	0.6500	0.6319
CSRM60	80	0.2750	0.7250	0.2003	0.7997	0.7897	0.7754
CSSM43	78	0.2564	0.7436	0.2840	0.7160	0.7068	0.7236
CSSM45	78	0.2051	0.7949	0.2188	0.7812	0.7712	0.7417
ILSTS05	78	0.6410	0.3590	0.4019	0.5981	0.5904	0.5952
ILSTS49	76	0.6842	0.3158	0.7267	0.2733	0.2697	0.3803
ILSTS58	80	0.1250	0.875	0.1313	0.8687	0.8578	0.8598
ILSTS72	78	0.5385	0.4615	0.5415	0.4585	0.4527	0.4968
BM1818	80	0.4000	0.6000	0.1655	0.8345	0.8241	0.7881
CSSM29	78	0.1282	0.8718	0.2178	0.7822	0.7722	0.6294
ILSTS11	80	0.2250	0.7750	0.2642	0.7358	0.7266	0.7247
ILSTS52	80	0.1250	0.8750	0.2263	0.7737	0.7641	0.7648
ILSTS59	80	0.3250	0.6750	0.2022	0.7978	0.7878	0.7802
ILSTS87	80	0.8250	0.1750	0.7500	0.2500	0.2469	0.1443
Mean	79	0.4090	0.5910	0.3400	0.6600	0.6516	0.6466
St. Dev		0.2128	0.2128	0.1811	0.1811	0.1789	0.1708

* Observed homozygosty and heterozygosity
* Expected homozygosty and heterozygosity
* PIC Polymorphic Information Content

The number of polymorphic loci is : 23
The percentage of polymorphic loci is : 100.00 %

Table No. 9: Different estimates of heterozygosity in Surti buffaloes

Locus	Sample Size	Obs_Hom	Obs_Het	Exp_Hom	Exp_Het*	PIC	Ave_Het
CSSM06	96	0.3125	0.6875	0.2757	0.7243	0.7168	0.7648
CSSM19	96	0.4375	0.5625	0.3057	0.6943	0.6871	0.7191
CSSM57	96	0.4792	0.5208	0.3401	0.6599	0.653	0.6935
ILSTS38	96	0.8750	0.1250	0.4982	0.5018	0.4965	0.4795
ILSTS29	96	0.7292	0.2708	0.6167	0.3833	0.3793	0.3978

ILSTS30	96	0.1458	0.8542	0.2243	0.7757	0.7676	0.7224
ETH152	96	0.3125	0.6875	0.2934	0.7066	0.6992	0.6936
CSSM47	96	0.0833	0.9167	0.1364	0.8636	0.8546	0.817
CSSM33	96	0.3958	0.6042	0.3515	0.6485	0.6417	0.5484
CSSM08	96	0.5000	0.5000	0.3798	0.6202	0.6137	0.6319
CSRM60	96	0.1875	0.8125	0.2309	0.7691	0.7611	0.7754
CSSM43	96	0.2500	0.7500	0.2518	0.7482	0.7405	0.7236
CSSM45	96	0.2917	0.7083	0.2803	0.7197	0.7122	0.7417
ILSTS05	96	0.4375	0.5625	0.3936	0.6064	0.6000	0.5952
ILSTS49	96	0.5417	0.4583	0.5039	0.4961	0.4909	0.3803
ILSTS58	96	0.0833	0.9167	0.1292	0.8708	0.8618	0.8598
ILSTS72	96	0.5625	0.4375	0.4533	0.5467	0.541	0.4968
BM1818	96	0.5625	0.4375	0.2399	0.7601	0.7522	0.7881
CSSM29	96	0.6875	0.3125	0.5083	0.4917	0.4865	0.6294
ILSTS11	96	0.3542	0.6458	0.2695	0.7305	0.7229	0.7247
ILSTS52	94	0.2979	0.7021	0.2263	0.7737	0.7655	0.7648
ILSTS59	96	0.2500	0.7500	0.2193	0.7807	0.7726	0.7802
ILSTS87	94	1.0000	0.0000	0.9579	0.0421	0.0416	0.1443
Mean	96	0.4251	0.5749	0.3516	0.6484	0.6417	0.6466
St. Dev		0.2383	0.2383	0.1816	0.1816	0.1797	0.1708

* Observed homozygosty and heterozygosity
* Expected homozygosty and heterozygosity
* PIC Polymorphic Information Content

The number of polymorphic loci is : 23
The percentage of polymorphic loci is : 100.00 %

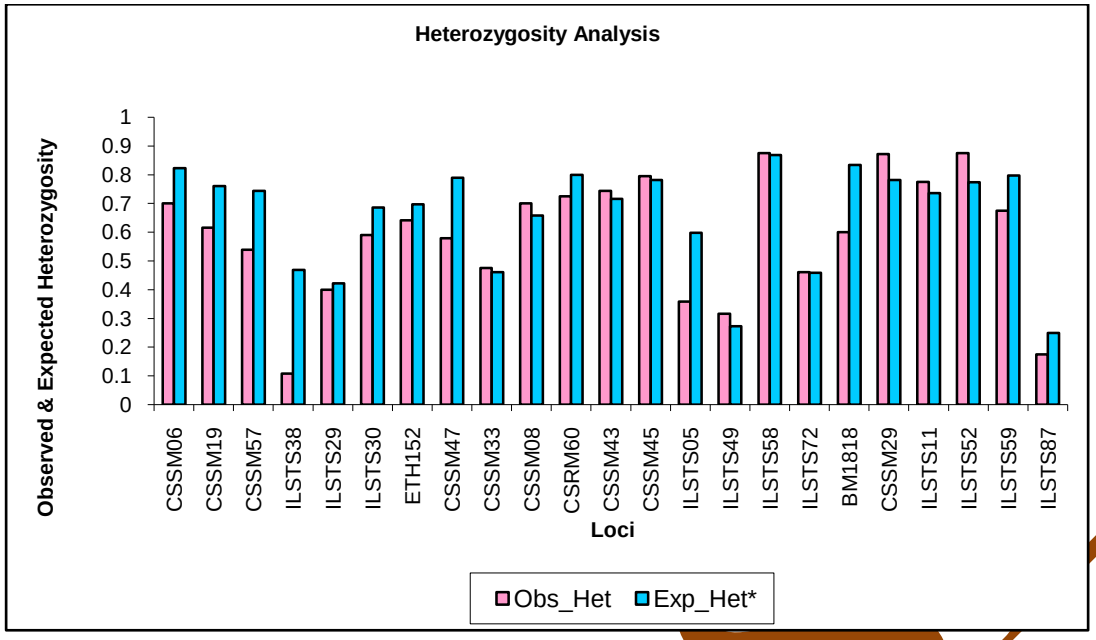


Fig. 7 Graphical representation of observed and expected heterozygosity of Marathwada buffaloes

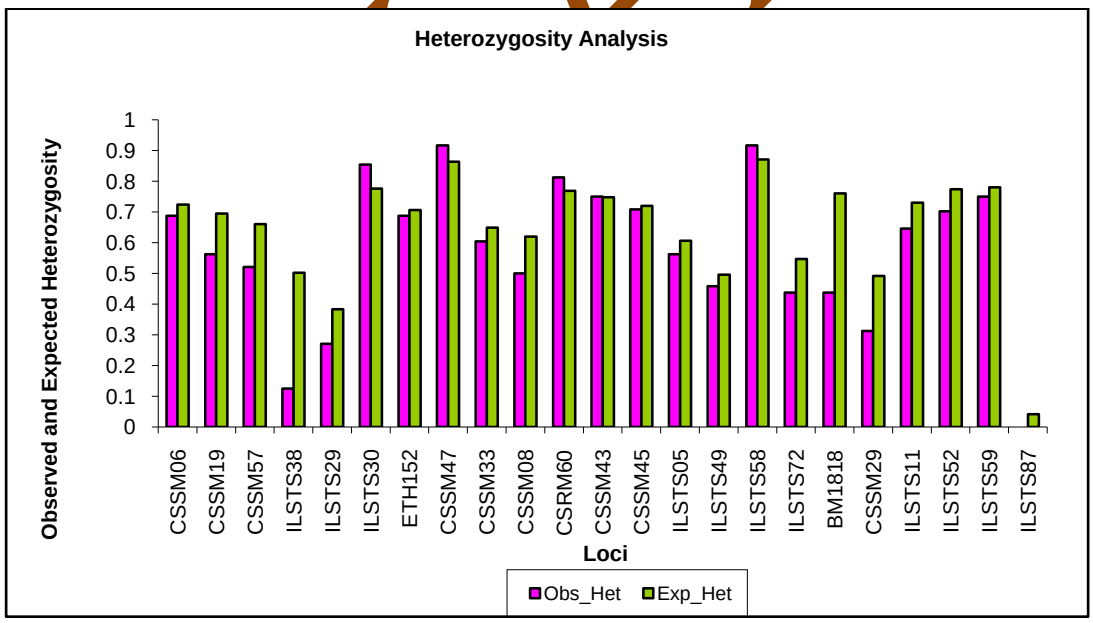


Fig. 8 Graphical representation of observed and expected heterozygosity of Surti buffaloes

4.4 HARDY WEINBERG EQUILIBRIUM

The Hardy Weinberg equilibrium states that in a large random mating population, the gene and genotypic frequency does not change from generation to generation in the absence of mutation, migration and selection. Two methods (statistics) were applied to test for the population of Marathwada and Surti. They were χ^2 test and G^2 or likelihood ratio test. The values obtained for each locus and populations have been depicted in table no.10 and table no.11 in Mehsana and Pandharpuri buffaloes respectively. Thirteen loci (CSSM06, CSSM57, ILSTS30, ILSTS38, ETH152, CSRM60, ILSTS05, ILSTS58, BM1818, CSSM45, ILSTS11, ILSTS59 and ILSTS87) deviated from Hardy Weinberg equilibrium using χ^2 test while three loci (ILSTS30, ILSTS38, and ILSTS05) using likelihood ratio test deviated from Hardy Weinberg equilibrium.

Similarly nine loci (CSSM19, ILSTS38, ILSTS29, ILSTS49, ILSTS72, BM1818, ILSTS11, CSSM29 and ILSTS87) in Surti deviated in χ^2 test and 7 loci (CSSM19, ILSTS38, ILSTS72, BM1818, ILSTS11, CSSM29 and ILSTS87) deviated in G^2 test. The deviation from Hardy Weinberg equilibrium point may be either towards small population size, mutation, migration or selection. All these factors have to be analyzed in details for pinpointing the deviations.

Table No. 10: Testing Marathwada buffalo population for Hardy Weinberg equilibrium using Chi square and likelihood test:

Locus	*Chi square Test*			Likelihood Ratio Test		
	Chi square	Degree of Freedom	Probability	G^2	Degree of Freedom	Probability
CSSM06	62.586211	45	0.042349*	44.948835	45	0.474104
CSSM19	17.458486	15	0.292203	19.725970	15	0.182699
CSSM57	42.671856	28	0.037431*	35.081090	28	0.167519
ILSTS38	102.236327	10	0.000000*	40.886638	10	0.000012*
ILSTS29	5.5551790	6	0.474813	4.789924	6	0.571026
ILSTS30	42.574762	15	0.000183*	31.203945	15	0.008249*
ETH152	90.839141	36	0.000001*	24.420133	36	0.928628
CSSM47	109.948948	66	56.644436	56.644436	66	0.787426
CSSM33	15.309558	21	0.807105	11.111071	21	0.960580
CSSM08	7.641808	10	0.663779	8.698296	10	0.560956
CSRM60	132.990416	55	0.000000*	52.870397	55	0.556423
CSSM43	8.998153	21	0.989229	9.353433	21	0.986149
CSSM45	93.251192	36	0.000001*	26.551694	36	0.874659

ILSTS05	91.959368	10	0.000000*	21.716989	10	0.016613*
ILSTS49	1.211310	3	0.750293	2.075650	3	0.556855
ILSTS58	135.868003	66	0.000001*	53.541304	66	0.864824
ILSTS72	0.487179	3	0.921699	0.791520	3	0.851494
BM1818	308.683108	105	0.000000*	87.288259	105	0.894700
CSSM29	29.900190	36	0.753073	30.069387	36	0.745912
ILSTS11	88.139455	15	0.000000*	20.733890	15	0.145549
ILSTS52	15.619760	28	0.971122	16.231033	28	0.962273
ILSTS59	96.430811	36	0.000000*	27.062479	36	0.858892
ILSTS87	13.460419	6	0.036280*	7.547300	6	0.273179

*Deviation from Hardy Weinberg equilibrium (<0.05)

Table No. 11: Testing Surti buffalo population for Hardy Weinberg equilibrium using Chi square and likelihood test:

Locus	Chi square Test			Likelihood Ratio Test		
	Chi square	Degree of Freedom	Probability	G^2	Degree of Freedom	Probability
CSSM06	29.849585	45	0.959941	26.191717	45	0.988754
CSSM19	52.105392	15	0.000005*	30.340507	15	0.010751*
CSSM57	36.205836	36	0.459043	36.659704	36	0.438084
ILSTS38	27.647096	1	0.000000*	30.847873	1	0.000000*
ILSTS29	57.440300	21	0.000031*	24.478807	21	0.270442
ILSTS30	24.844573	28	0.636281	27.863969	28	0.471674
ETH152	28.466233	21	0.127417	17.945395	21	0.652457
CSSM47	74.709971	66	0.216354	50.637885	66	0.919007
CSSM33	20.111434	36	0.984997	18.913745	36	0.991445
CSSM08	7.812669	10	0.647129	8.362080	10	8.362080
CSRM60	24.656250	36	0.923587	25.965923	36	0.891399
CSSM43	20.827941	21	0.469494	18.459692	21	0.619757
CSSM45	15.543885	21	0.794706	14.244678	21	0.858862
ILSTS05	12.467971	6	0.052307	11.669929	6	0.069751
ILSTS49	97.016906	10	0.000000*	13.326229	10	0.206000
ILSTS58	58.352654	66	0.737145	47.930750	66	0.954024
ILSTS72	7.864578	3	0.048895*	9.948370	3	0.019010*
BM1818	107.391938	45	0.000001*	70.956711	45	0.008090*
CSSM29	34.731494	15	0.002685*	27.812764	15	0.022771*
ILSTS11	15.635623	6	0.015849*	14.382091	6	0.025647*
ILSTS52	49.935170	36	0.061216	39.656774	36	0.310230
ILSTS59	19.701647	21	0.540223	20.017028	21	0.520183
ILSTS87	93.010989	1	0.000000*	11.065278	1	0.000880*

*Deviation from Hardy Weinberg equilibrium (<0.05)

4.5 EWENS-WATTERSON TEST FOR NEUTRALITY

The microsatellite loci are considered neutral and thus selection cannot operate on them. However, if these microsatellite loci are present close to a trait of interest on which the selection is operating. The microsatellite loci shall also be subjected to selection pressure. In order to test the neutrality of microsatellite loci used in diversity analysis, we performed Ewens-Watterson test for neutrality given by Manly, 1985. The data was simulated 1000 times to generate the lower and upper limits. The number of observation, number of alleles, observed frequency and its upper and lower limits at 95% of confidence are given in table no. 12. All the loci selected for the present analysis were neutral in character. These shows that all the microsatellites selected for diversity analysis were suitable for the purpose.

Table No. 12 Ewens Watterson test for neutrality

Locus	n	k	Obs. F	L95*	U95*
CSSM06	176	13	0.2173	0.1315	0.489

CSSM19	174	6	0.2766	0.2453	0.8374
CSSM57	174	11	0.2537	0.1532	0.5466
ILSTS38	170	5	0.5063	0.2745	0.8869
ILSTS29	176	7	0.6011	0.2176	0.7888
ILSTS30	174	8	0.2284	0.1920	0.7107
ETH152	174	9	0.2983	0.1802	0.6585
CSSM47	172	16	0.1575	0.1120	0.3673
CSSM33	176	10	0.4316	0.1690	0.5990
CSSM08	176	6	0.3608	0.2430	0.8195
CSRM60	176	12	0.2171	0.1429	0.5229
CSSM43	174	8	0.2711	0.1953	0.7197
CSSM45	174	10	0.2407	0.1599	0.5855
ILSTS05	174	6	0.3992	0.2455	0.8473
ILSTS49	172	6	0.5853	0.2413	0.8356
ILSTS58	176	15	0.1278	0.1169	0.4085
ILSTS72	174	4	0.4469	0.3296	0.9437
BM1818	176	16	0.2069	0.1100	0.3735
CSSM29	174	10	0.3538	0.1701	0.6077
ILSTS11	176	6	0.2716	0.2413	0.8300
ILSTS52	174	10	0.2254	0.1602	0.5908
ILSTS59	176	11	0.2157	0.1532	0.5518
ILSTS87	174	4	0.8592	0.3293	0.9437

* These statistics were calculated using 1000 simulated samples.

4.6 F- STATISTICS

It provides the analysis pertaining to nonrandom union of gametes i.e., inbreeding as well as population differentiation. The F_{IS} values provide the nonrandom union of gametes in the population i.e., the mating among the individuals in the populations is more/less related than the average relationship among the individuals. The more the value of F_{IS} the more the inbreeding coefficient and thus the individuals are more closely related to one another. The negative values of F_{IS} point towards out breeding i.e., the mating of individuals who are less related than the average relationship of the population.

In case of Marathwada the F_{IS} values are significantly different from zero and ten loci point towards the outbreeding while thirteen loci point (table no. 13) towards mating among the individuals of the population more closely related than the average relationship.

In case of Surti five loci have negative values (table no. 13), which point towards outbreeding in population i.e., the mating of the individuals is among the individuals less related than the average relationship of the population. Eighteen other loci however point towards existence of population structure.

The F_{ST} values, which point toward the population differentiation, have positive value of 0.0253, which is significantly different from zero (table no. 14). More so the F_{ST} value point towards the low level of population differentiation among the Marathwada and Surti populations. In terms of analysis of molecular variance (AMOVA) between the variations is 2.53% within the population variation is 97.47%. In term of member of migrants estimated from F_{ST} value, the vales are quite high 9.6244 meaning there by a moderate exchange of migrants between the two populations. The exchange of moderate member of migrants among Marathwada and Surti populations may be the reason for low level of population differentiation between the two populations. It may also be possible that the semen of the bulls (Murrah) may have been utilized in the two populations and thus the allele frequencies in the two populations have become similar and the differentiation between the two population have been reduced substantially.

Table No. 13: Wright's Fixation Index (F_{IS}) of Marathwada and Surti buffaloes

Locus	F_{IS} of Marathwada	F_{IS} of Surti
CSSM06	0.1388	0.0409
CSSM19	0.1807	0.1813
CSSM57	0.2665	0.2024
ILSTS38	0.7662	0.7483
ILSTS29	0.0390	0.2860
ILSTS30	0.1291	-0.1128
ETH152	0.0683	0.0168
CSSM47	0.2572	-0.0726
CSSM33	-0.0440	0.0585
CSSM08	-0.0769	0.1853
CSRM60	0.0819	-0.0676
CSSM43	-0.0521	-0.0129
CSSM45	-0.0307	0.0055
ILSTS05	0.3920	0.0626
ILSTS49	-0.1707	0.0663
ILSTS58	-0.0200	-0.0637
ILSTS72	-0.0196	0.1913
BM1818	0.2719	0.4183
CSSM29	-0.1290	0.3577
ILSTS11	-0.0667	0.1066

ILSTS52	-0.1452	0.0828
ILSTS59	0.1432	0.0292
ILSTS87	0.2911	1.0000

Table No. 14: F-Statistics and Gene Flow in buffalo populations

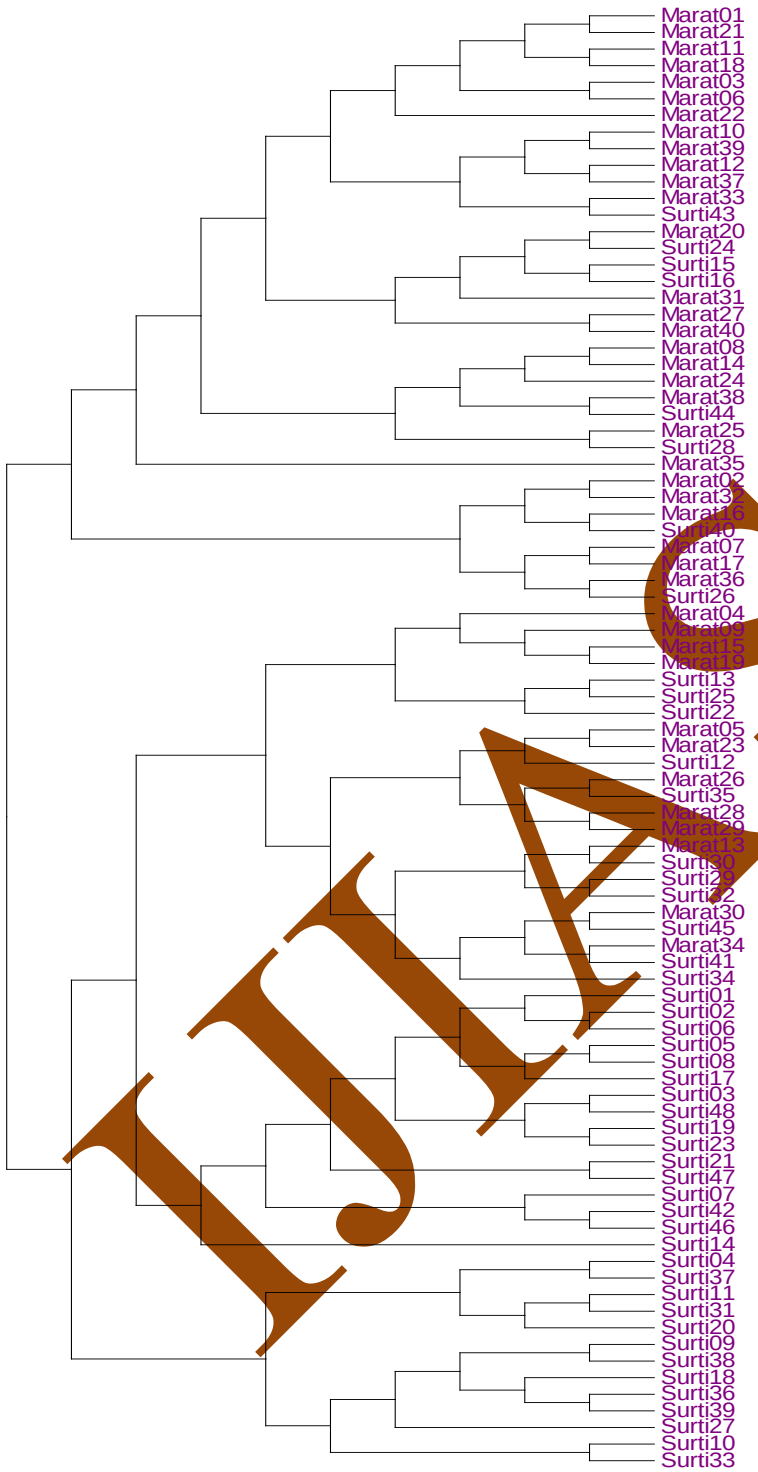
Locus	Sample Size	Fis	Fit	Fst	Nm*
CSSM06	176	0.0929	0.1187	0.0285	8.5299
CSSM19	174	0.1810	0.1897	0.0106	23.3472
CSSM57	174	0.2363	0.2948	0.0766	3.0127
ILSTS38	170	0.7569	0.763	0.0249	9.79
ILSTS29	176	0.1568	0.1627	0.0070	35.5792
ILSTS30	174	0.0006	0.0593	0.0587	4.0080
ETH152	174	0.0423	0.0527	0.0108	22.9564
CSSM47	172	0.0847	0.1081	0.0255	9.5513
CSSM33	176	0.016	0.0364	0.0208	11.7877
CSSM08	176	0.0504	0.0639	0.0142	17.4139
CSRM60	176	0.0085	0.0198	0.0114	21.7174
CSSM43	174	-0.0320	-0.0270	0.0049	50.5484
CSSM45	174	-0.0133	0.0144	0.0274	8.8725
ILSTS05	174	0.2259	0.2326	0.0086	28.9508
ILSTS49	172	-0.0177	0.0375	0.0542	4.3594
ILSTS58	176	-0.0419	-0.0272	0.0142	17.392
ILSTS72	174	0.0952	0.1814	0.0952	2.3767
BM1818	176	0.3418	0.3487	0.0105	23.5943
CSSM29	174	0.0591	0.1046	0.0483	4.9265
ILSTS11	176	0.0197	0.0250	0.0054	46.2535
ILSTS52	174	-0.0311	-0.018	0.0127	19.4524
ILSTS59	176	0.0868	0.0924	0.0061	40.4128
ILSTS87	174	0.3935	0.4132	0.0326	7.4148
Mean	174	0.0982	0.1210	0.0253	9.6244

4.7 GENETIC DISTANCES

The interindividual genetic distance estimated and the constitution of phylogenetic tree can also used as an assignment method (fig. 17a-d). In case of Marathwada and Surti, the nine individuals of Marathwada clubbed with Surti while six individuals of Surti clubbed with Marathwada. In all there were fifteen wrong assignments based on the phylogenetic tree

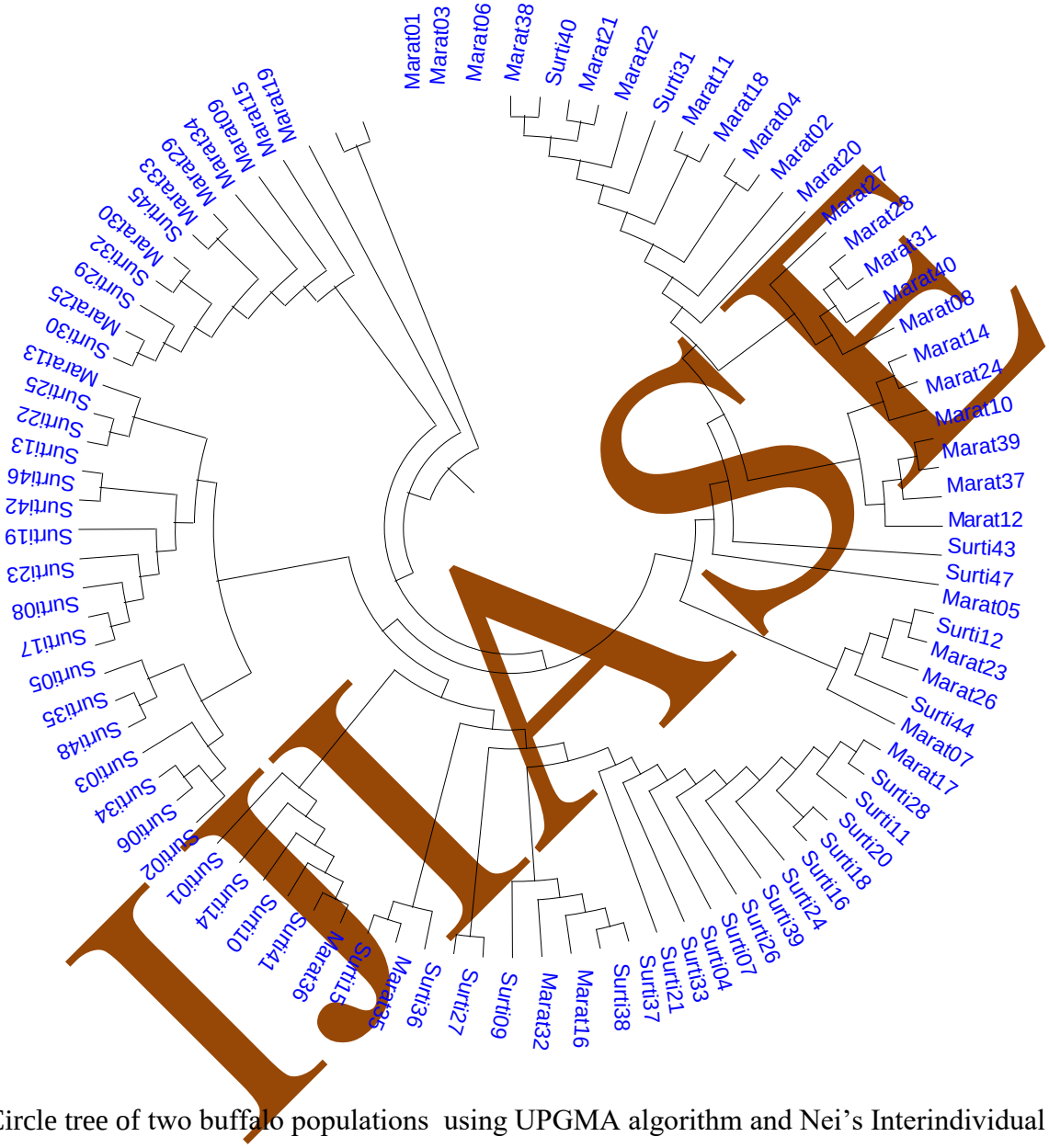
constructed. This contributes approximately 17% wrong assignment meaning there by the two populations are quite distinctive in terms of their genotypes.

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**



Topology of Surti and Marathwada populations using Neighbour joining Algorithm.

Radiation tree using Neighbour joining algorithm and Nei's standard genetic distance.



Circle tree of two buffalo populations using UPGMA algorithm and Nei's Interindividual genetic distance.

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

Curved tree using UPGMA algorithm and Inter individual distances based on allelic frequency data.

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

4.8 MUTATION DRIFT EQUILIBRIUM

The population bottlenecks or the founder events are very important for population genetic studies, speciation theory and conservation biology. It is well known that population bottlenecks invariably lead to increased rate of inbreeding and loss of genetic variation and may eventually lead to lower adaptive value of the individuals. The population bottlenecks may cause extensive genetic changes. For the microsatellite loci, which are neutral, the allele number and frequency distribution in a natural population results from a dynamic equilibrium between mutation and genetic drift. When a population experiences a reduction of its effective size, it generally develops heterozygosity excess at selectively neutral loci. The heterozygosity excess persists only for a certain number of generations until a new equilibrium is established. Thus the bottlenecks can be detected for populations who have experienced it in recent times upto few dozen generations. The bottleneck causes the effective population size to be dramatically reduced. This is coupled with a correlative and progressive reduction of the allele number and heterozygosity. The allelic diversity is reduced faster than is the heterozygosity during a bottleneck (Nei *et al*., 1975 and Denniston, 1978). The allelic deficiency is a complex function of four parameters; the time since the beginning of the bottleneck, effective population size ratio before/after the beginning of the bottleneck, the mutation rate of the loci and the sample size of genes. Maruyama and Fuerst (1985) have reported that the magnitude of an allele deficiency first increases with t, reaches a maximum and then decreases asymptotically toward zero, corresponding to new mutation drift equilibrium. Cornuet and Luikart (1996) utilized this knowledge to develop a test based on detecting an allelic deficiency. The test is based on the principle that if a significantly high proportion of loci exhibit an allelic deficiency, one can conclude that the population is not at mutation drift equilibrium due to a recent reduction of the effective population size. The difference between the observed heterozygosity and heterozygosity expected from the observed number of alleles is utilized in the test to find the number of loci exhibiting heterozygosity excess. The populations that exhibit significant heterozygosity excess would be considered as having experienced a recent genetic bottleneck. The other test called "shift mode" test given by Luikart *et al*., (1998) is based on the principle that the population at mutation drift equilibrium for selectively neutral loci are expected to have a large proportion of alleles at low frequency. It is also known that the alleles at low frequency (<0.1) are always expected to be more abundant than the alleles at intermediate frequencies, regardless of the mutation rate and model (Nei and Li, 1976). The populations, which have undergone recent bottleneck, are expected to lose the alleles at low frequency (Allendorf 1986; Denniston, 1978; Nei *et al*., 1976). Thus the distribution of allele frequency provides a test for a recent population bottleneck.

Two buffalo populations Marathwada (Maharashtra) and Surti (Gujarat) were selected for the study. We utilized four tests viz. Sign test, Standardized differences test, Wilcoxon test

and Shift mode test to find out which of the two buffalo population had undergone population bottleneck in recent times. We also utilized the three models of microsatellite evolution i.e. IAM, SMM, and TPM. The sample size in terms of haploid genome, He (Average Heterozygosity), the heterozygosity observed, K_o (number of alleles observed) for the data have been given in table no. 16 and 17 for the two buffalo populations.

In Marathwada the sample size was 40 i.e., 80 haploid genomes. The values of the haploid genomes for all 23 loci separately have been given in table no. 16a and 16b. The expected numbers of loci with heterozygotic excess were 13.60 out of 23 loci under IAM, 13.65 under the TPM and 13.67 under the SMM (table no. 15). The numbers of loci with heterozygosity excess observed in the study were 15, 9 and 2 loci respectively for the three mutational models. The probability values were 0.35635 (IAM), 0.04038 for (TPM) and 0.00000 for (SMM) respectively. The sign test reveals no bottleneck in the population of Marathwada buffaloes under IAM i.e., the null hypothesis of mutation drift equilibrium is accepted under IAM but rejected under SMM and TPM in favour of heterozygotic deficiency. The results obtained for the standardized difference test in which the T2 values were calculated and have been given in table no. 15. The T2 values were less than 1.645 and thus null hypothesis of mutation drift equilibrium is accepted for IAM while it is rejected for TPM and SMM in favour of heterozygotic deficiency. The Wilcoxon test for Marathwada population also reveals similar results the probability values being 0.01633 (<0.05) i.e. rejecting the null hypothesis of mutation drift equilibrium under IAM while accepted under TPM and SMM. The mode shift reveals no distortion of allelic frequency distribution and the graphic representation is normal L-shaped (fig 18).

Table No. 15: Tests for Null Hypothesis under three microsatellites mutation models

Buffaloes	**IAM**		**TPM**		*SMM*	
	Sign Test (No. of loci with heterozygosity excess)					
	Expected	Observed	Expected	Observed	Expected	Observed
Marathwada	13.60	15	13.65*	9	13.67*	2
Surti	13.15	16	13.27	12	13.24*	5
Standardized differences test (T2 values)						
Marathwada	1.362		-1.744*		-7.699*	
Surti	2.218*		-0.943		-6.9944*	
Wilcoxon rank test (probability of heterozygosity Excess)						
Marathwada	0.01633*		0.89441		1.0000	
Surti	0.00558*		0.62301		0.99893	

(IJIASE) 2016, Vol. No. 2, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

* Rejection of Null Hypothesis of mutation drift equilibrium bottleneck

† The negative values show heterozygotic deficiency.

Table No. 16a. Testing heterozygosity excess at 23 polymorphic microsatellite loci in Marathwada buffaloes using Standardised deviation test

Locus	CSSM06	CSSM19	CSSM57	ILSTS38	ILSTS29	ILSTS30	ETH152	CSSM47	CSSM33	CSSM08	CSRM60	CSSM43
Sample size (Haploid)	80	78	78	74	80	78	78	76	80	80	80	78
Ho	0.823	0.761	0.744	0.469	0.422	0.686	0.697	0.790	0.461	0.658	0.800	0.716
No. of Alleles (ko)	10	6	8	5	4	6	9	12	7	5	11	7
IAM												
He	0.746	0.589	0.685	0.537	0.447	0.595	0.725	0.802	0.646	0.524	0.777	0.643
Std. Deviation(SD)	0.092	0.145	0.116	0.153	0.171	0.138	0.095	0.069	0.125	0.158	0.086	0.124
[(Ho-He)/SD]	0.839	1.187	0.511	-0.447	-0.150	0.660	-0.293	-0.178	-1.485	0.849	0.269	0.588
Probability(H>He)	0.1830	0.0670	0.3700	0.2880	0.4010	0.2960	0.2920	0.3230	0.0950	0.2170	0.5040	0.3290
TPM												
He	0.808	0.669	0.749	0.606	0.519	0.666	0.784	0.846	0.714	0.596	0.825	0.717
Std. Deviation(SD)	0.051	0.096	0.076	0.123	0.140	0.103	0.061	0.041	0.084	0.124	0.050	0.086
[(Ho-He)/SD]	0.302	0.956	-0.073	-1.115	-0.694	0.196	-1.433	-1.373	-3.016	0.502	-0.498	-0.0008
Probability(H>He)	0.4500	0.1490	0.3760	0.1470	0.2350	0.5170	0.0840	0.0920	0.0170	0.3630	0.2440	0.4000
SMM												
He	0.849	0.726	0.806	0.680	0.590	0.735	0.830	0.877	0.771	0.675	0.864	0.770
Std. Deviation(SD)	0.030	0.072	0.042	0.078	0.1006	0.016	0.037	0.023	0.055	0.089	0.031	0.054
(Ho-He)/SD]	-0.852	0.484	-1.497	-2.697	-1.599	-0.800	-3.577	-3.808	-5.663	-0.186	-2.069	-1.010
Probability(H>He)	0.1630	0.3770	0.0840	0.0270	0.0810	0.1780	0.0080	0.0050	0.0010	0.3220	0.0350	0.1390

Table No. 16b. Testing heterozygosity excess at 23 polymorphic microsatellite loci in Marathwada buffaloes using Standardised deviation test

Locus	CSSM45	ILSTS05	LISTS49	ILSTS58	ILSTS72	BM1818	CSSM29	ILSTS11	ILSTS52	ILSTS59	ILSTS87
Sample size (Haploid)	78	78	76	80	78	80	78	80	80	80	80
Ho	0.781	0.598	0.273	0.869	0.459	0.834	0.782	0.736	0.774	0.798	0.250
No. of Alleles (ko)	9	5	3	12	3	15	15159	6	8	9	4
IAM											
He	0.719	0.524	0.336	0.797	0.344	0.846	0.721	0.590	0.683	0.719	0.440
Std. Deviation(SD)	0.100	0.154	0.180	0.072	0.184	0.052	0.104	0.144	0.117	0.103	0.171
[(Ho-He)/SD]	0.623	0.480	-0.346	0.993	0.621	-0.224	0587	1.011	0.780	0.762	-1.107
Probability(H>He)	0.3140	0.3890	0.4000	0.1030	0.3570	0.3040	0.3210	0.1330	0.2310	0.2430	0.1840
TPM											
He	0.786	0.611	0.411	0.845	0.402	0.880	0.785	0.661	0.754	0.785	0.518
Std. Deviation(SD)	0.058	0.13	0.164	0.041	0164	0.032	0.061	0.202	0.068	0.062	0.149
[(Ho-He)/SD]	-0.079	-0.115	-0.837	0.589	0.344	-1.397	-0.039	0.730	0.287	0.212	-1.806
Probability(H>He)	0.3920	0.3690	0.2400	0.3230	0.4720	0.0800	0.4010	0.2590	0.4940	0.5080	0.0760
SMM											
He	0.827	0.676	0.482	0.877	467	0.906	0.830	0.735	0.803	0.829	0.590
Std. Deviation(SD)	0.404	0.083	.0.126	0.023	0.138	0.020	0.036	0.061	0.043	0.037	0.109
(Ho-He)/SD]	-1.144	-0.940	-1.657	-0.367	-0.064	-3.538	-1.323	0.011	-0.684	-0.835	-3.109

Probability(H>He)	0.1160	0.1450	0.0850	0.3070	0.3780	0.0050	0.1080	0.4320	0.2100	0.1740	0.0120

Table No. 17a. Testing heterozygosity excess at 23 polymorphic microsatellite loci in Surti buffaloes using Standardised deviation test

Locus	CSSM45	ILSTS05	LISTS49	ILSTS58	ILSTS72	BM1818	CSSM29	ILSTS11	ILSTS52	ILSTS59	ILSTS87
Sample size (Haploid)	96	96	96	96	96	96	96	96	94	96	[illegible]
Ho	0.720	0.606	0.496	0.871	0.547	0.760	0.492	0.730	[illegible]	[illegible]	[illegible]
No. of Alleles (ko)	7	4	5	12	3	10	6	4	9	7	[illegible]
IAM											
He	0.627	0.432	0.515	0.777	0.327	0.738	0.578	0.429	[illegible]	[illegible]	[illegible]
Std. Deviation(SD)	0.136	0.179	0.163	0.083	0.182	0.097	0.143	0.176	[illegible]	[illegible]	[illegible]
[(Ho-He)/SD]	0.682	0.978	-0.114	1.131	1.211	0.233	-0.603	1.702	[illegible]	[illegible]	[illegible]
Probability(H>He)	0.2930	0.1880	0.3740	0.0550	0.1130	0.5120	0.2520	0.01[illegible]	[illegible]	[illegible]	[illegible]
TPM											
He	0.710	0.512	0.600	0.836	0.397	0.800	0.65[illegible]	0.504	[illegible]	[illegible]	[illegible]
Std. Deviation(SD)	0.084	0.141	0.121	0.044	0.167	0.056	0.109	0.149	[illegible]	[illegible]	[illegible]
[(Ho-He)/SD]	0.116	0.668	-0.857	0.789	0.897	-0.713	-1.469	1.527	[illegible]	[illegible]	[illegible]
Probability(H>He)	0.4710	0.2930	0.1690	0.2150	0.1960	0.1930	0.0860	0.01[illegible]	[illegible]	[illegible]	[illegible]
SMM											
He	0.771	0.589	0.670	0.873	0.466	0.845	0.728	0.58[illegible]	[illegible]	[illegible]	[illegible]
Std. Deviation(SD)	0.052	0.103	0.089	0.026	0.132	0.033	0.067	0.108	[illegible]	[illegible]	[illegible]
[(Ho-He)/SD]	-0.979	0.166	-1.944	-0.098	0.612	-2.604	-3.535	1.349	[illegible]	[illegible]	[illegible]
Probability(H>He)											

Table 17b. Testing heterozygosity excess at 23 polymorphic microsatellite loci in Surti buffaloes using Standardised deviation test

Locus	CSSM06	CSSM19	CSSM57	ILSTS38	ILSTS29	ILSTS30	ETH152	CSSM47	CSSM33	CSSM08	CSRM60	CSSM43
Sample size (Haploid)	96	96	96	96	96	96	96	96	96	96	96	[illegible]
Ho	0.724	0.694	0.660	0.502	0.383	0.776	0.707	0.864	0.648	[illegible]	[illegible]	[illegible]
No. of Alleles (ko)	10	6	9	2	7	8	7	12	9	5	9	[illegible]
IAM												
He	0.733	0.578	0.703	0.189	0.629	0.675	0.629	0.779	0.704	[illegible]	[illegible]	[illegible]
Std. Deviation(SD)	0.097	0.143	0.110	0.165	0.130	0.115	0.134	0.080	0.107	[illegible]	[illegible]	[illegible]
[(Ho-He)/SD]	-0.094	0.814	-0.389	1.898	-1.889	0.874	0.580	1.059	-0.520	[illegible]	[illegible]	[illegible]
Probability(H>He)	0.37[illegible]	0.2240	0.2700	0.0230	0.0640	0.1700	0.3390	0.0790	0.2440	[illegible]	[illegible]	[illegible]
TPM												
He	0.801	0.657	0.774	0.211	0.713	0.745	0.707	0.837	[illegible]71	[illegible]	[illegible]	[illegible]
Std. Deviation(SD)	0.051	0.104	0.064	0.170	0.084	0.073	0.091	0.044	[illegible]67	[illegible]	[illegible]	[illegible]
[(Ho-He)/SD]	-1.500	0.361	-1.792	1.705	-3.928	0.418	-0.002	0.592	-1.839	[illegible]	[illegible]	[illegible]
Probability(H>He)	0.08[illegible]	0.42[illegible]	0.0590	0.0500	0.0070	0.3860	0.3960	0.3120	0.0620	[illegible]	[illegible]	[illegible]
SMM												
He	0.844	0.728	0.825	0.240	0.770	0.799	0.771	0.875	0.824	[illegible]	[illegible]	[illegible]
Std. Deviation(SD)	0.032	0.065	0.036	0.170	0.053	0.045	0.049	0.024	0.036	[illegible]	[illegible]	[illegible]
(Ho-He)/SD]	-3.701	-0.517	-4.569	1.543	-7.268	-0.527	-1.298	-0.477	-4.863	[illegible]	[illegible]	[illegible]
Probability(H>He)	0.0040	0.2580	0.0030	0.0460	0.0000	0.2540	0.1080	0.2710	0.0010	0.2240	0.0780	0.27[illegible]

Probability(H>He)	0.1510	0.4930	0.0550	0.3810	0.2940	0.0310	0.0080	0.0190	0.0940	0.4920	0.1900

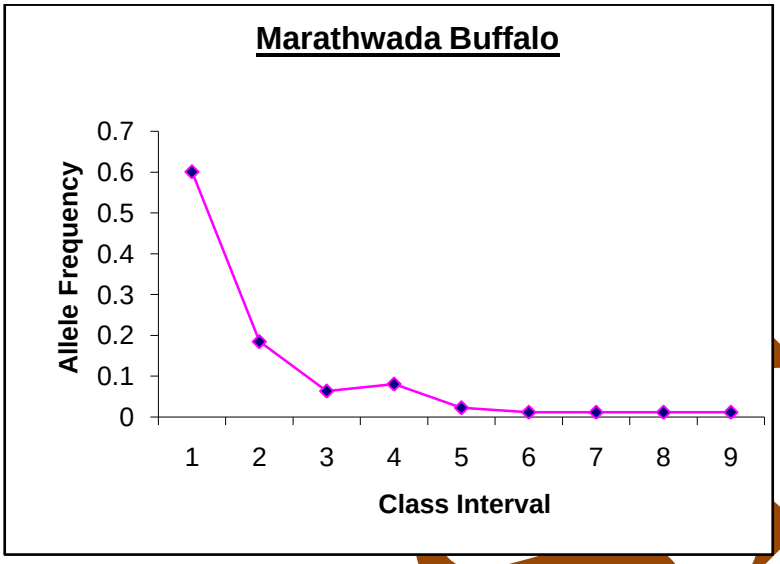


Figure 9: Normal L shaped curve depicting no Mode Shift of allelic distribution in Marathwada Buffalo population.

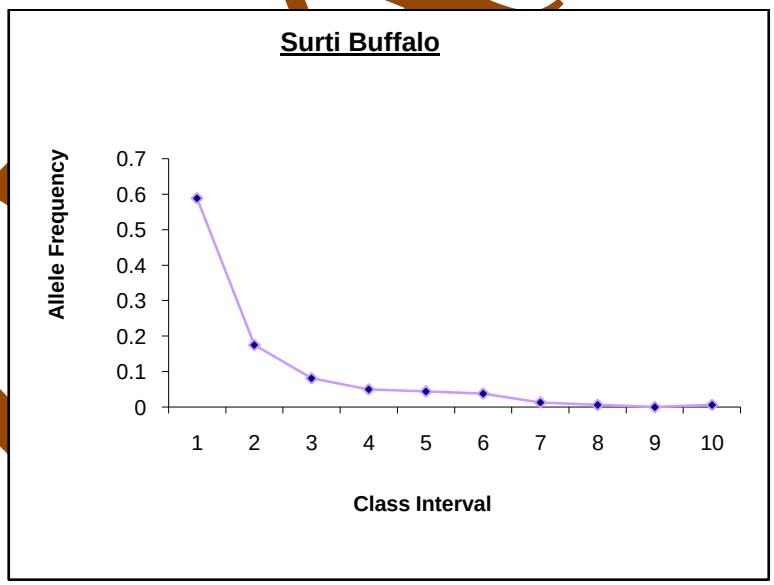


Figure 10: Normal L shaped curve depicting no Mode Shift of allelic distribution in Surti Buffalo population.

In Surti buffalos the number of individuals typed were 48 or 96 haploid genomes. The values of sample size, observed heterozygosity and number of alleles for each locus in Surti buffaloes are given in table no. 17a and 17b. The expected numbers of loci and observed loci with heterozygosity excess have been given in table no. 15. Using Sign test the number of loci with heterozygosity excess were 19, 17 and 11 respectively for the three-mutation models. The expected values for the three models were respectively 13.15, 13.27 and 13.24. In the Sign rank test the null hypothesis of mutation drift equilibrium was rejected in case of SMM in favour of heterozygosity deficiency and was accepted for TPM as well as IAM. The values obtained in Standardized difference test were 2.218, -0.943 and –6.994 respectively for IAM, TPM and SMM. However in IAM the null hypothesis is rejected for heterozygotic excess while it is rejected in SMM for heterozygotic deficiency. Wilcoxon rank test in which the null hypothesis was rejected under IAM probability values being 0.00558 and accepted in TPM and SMM probability values being 0.62301 and 0.99893 (<0.05). The shift mode test however reveal normal L-shaped curve (fig. 19) and there is no distortion from expected distribution of alleles.

The results of the bottleneck studies do not reveal any buffalo population to have experienced any recent genetic bottleneck if we use Stepwise mutation Model which is the most conservative model for the test of heterozygosity excess. IAM and SMM represent the two extreme models of mutation (Chakraborty and Jin, 1992). In the present study all the two populations exhibit bottleneck in all the three quantitative tests applied viz; Sign test, Standardized difference test and Wilcoxon rank test. It is however known that most of the loci are expected to evolve in an intermediate way (TPM model). In the present case the Surti buffaloes reveal rejection of null hypothesis of mutation drift equilibrium in IAM model of microsatellite evolution. However, the null hypothesis is accepted in TPM and SMM which being the most conservation model. The rejection of the null hypothesis in SMM however gives credence to the bottleneck. In case of Surti buffaloes such is not the case and the bottleneck if any must have occurred approximately 200 generations ago. The inference is backed by the allelic frequency graph, which do not reveal any bottleneck in the recent 50-100 generation.

6. SUMMARY

1. The 23 heterologus microsatellite loci selected for this study in Marathwada and Surti breeds of buffaloes were found to be highly polymorphic with allele numbers ranging from 3 to 15 in Marathwada breed and 2 to 12 in Surti breed respectively. This shows the utility of heterologus microsatellite loci of cattle in buffaloes for diversity analysis.

2. The two populations Marathwada and Surti buffaloes taken in this study showed variations in terms of allelic frequencies and the population were found to be quite distinct from one another.

3. The microsatellite studied in the present study were neutral in nature as evidenced by Ewens Watterson test of neutrality. This implies that the loci have not been subjected to selection and thus are fit for population genetic analysis at these loci.

4. The microsatellite loci selected were highly polymorphic with a large number of alleles at lower allelic frequencies. This is depicted by quite large variation between the number of alleles and effective number of alleles.

5. Most of the loci had very high heterozygosity values with mean 0.591 in Marathwada and 0.575 in Surti breed. The heterozygosity values in population ranged for 0.108 to 0.875 and 0.125 to 0.9167 in Surti population, except ILSTS87, which was homozygous for the alleles, observed.

6. The Hardy Weinberg equilibrium of the two buffalo populations was tested using chi square and G^2 statistics and 13 loci out of 23 loci were not in Hardy Weinberg equilibrium in Marathwada buffalo population and 9 loci in Surti buffalo population. It means that the gene and genotypic have changed with respect to the loci not in Hardy Weinberg equilibrium by factors like mutation, migration, selection and drift (small population size).

7. The F_{IS} values were significantly different from zero in both the Marathwada and Surti buffalo population meaning there by that the population structure exists in the two populations.

8. The F_{ST} value in two buffalo population were 0.0253 meaning there by that the two buffalo population were poorly differentiated at DNA level even though the two buffalo population Marathwada and Surti were morphologically distinct.

9. The Analysis of Molecular Variance (AMOVA) revealed that the between population variation was 2.53% while within the population variation was quite high 97.47% which means that the inter individual distances are more pronounced than the interpopulation differences.

10. The value of number of migrants between the populations was quite high, it being 9.62 migrants per generations depicting large amount of gene flow among the buffalo populations.

11. The interindividual genetic distance based on Nei's genetic measure were utilized for the assignment of the individuals to their respective populations using neighbour joining (NJ) and unweighted pair group method with arithmetic mean (UPGMA) algorithms. 17% wrong assignments were found in Marathwada and Surti buffaloes, which are attributed to similarity in allele frequencies.

12. The mutation drift equilibrium and bottleneck studies revealed that none of the populations were found to have a recent genetic bottleneck by graphical method as well as SMM of microsatellite evolution.

REFERENCES

Albert, B.; Bray, D.; Lewis, J.;Raff, M.; Robert, K. and Watson, J. D. (1994). Molecular Biology of the cell. Garland Publishing Inc.

Ali S, Gauri, Bala S. Detection of genome specific monomorphic loci in Bos Taurus and Bubalus bubalis with oligodeoxyribonucleotide probe. Anim Genet. 1993 Jun; 24(3): 199-202.

Allendorf F. W. 1986 Genetic drift and the loss of alleles versus heterozygosity. Zoo. Biol. 5, 653-662.

Amos, W., S.J. Sawcer, R.W. Feakes, and D.C. Rubinsztein. 1996. Microsatellites show mutational bias and heterozygote instability. Nature Genet. 13:390-391.

Bachtrog, D.; Weiss, S.; Zangerl, B.; Brem, G.; Schlotterer, C. (1999). Distribution of dineucleotide microsatellites in the Drosophila melanogaster genome. Molecular Biology and Evolutionary, 16: 602-610.

Barker, J. S. F., Moore, S. S. Hetzel, D. J. S.; Evans, D., Tan, S. G. and Byrne, K. (1997). Genetic diversity of Asian water buffalo: Microsatellite variation and a comparison with protein coding loci. Animal genetic. 28: 103-115.

Botstein, D., White, R. L. Skonick, M. and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Human Genetic. 32: 314-331.

Bowcock, A.M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R., and Cavalli- Sforza, L.L. 1994. High resolution of human evolution with polymorphic microsatellites. Nature 368: 455-457.

Budowle, B., Giusti, A.M., Waye, J.S., Baechtel, F.S., Fourney, R.M., Adams, D.E., Presley, L.A., Deadman, H.A., and Monson, K.L. 1991. Fixed BIN analysis for statistical evaluation of continuous distributions of allele data form VNTR loci for use in forensic comparisons. Am. J. Hum. Genet. 48: 841-855.

Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C., and Sutherland, G.R. 1993. Incidence and origin of 'null' alleles in the (AC) n microsatellite marker. Am. J. Hum. Genet. 52: 922-927.

Cornuet J M and Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-14.

Crow, J.F. 1993. How much do we know about spontaneous human mutation rates? Environmental Molecular Mutagen. 21: 122-129.

Deka, R. et al. 1995. Population genetics of dinucleotide (dC-dA) n (dG-dT). Polymorphisms in world populations. Am. J Hum Genet. 56:461-474.

Denniston C.1978. Small population size and genetic diversity. Implications for endangered species. Pp. 281-290. In: Endangered Birds: Management Techniques for Preserving Endangered Species, Edited by S. A. Temple. University of Wisconsin Press, Madison.

Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., and Freimer, N.B. 1994. Mutational processes of simple-sequence rep Dieringer, D., SchlÄotterer, C.: Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. Genome Research 13: 2242{2250 (2003).

Dib, C. et al. 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellties. Nature 380:152-154.

Dietrich, W. et al. 1996. A comprehensive genetic map of the mouse genome. Nature 380:149-152.

Estoup, A., Tailliez, C., Cornuet, J. -M., and Solignac, M. 1995a. Size homoplasy and the mutational processes of interrupted microsatellites in two bee species, Apis mellifera and Bombus terrestrus (Apidae). Mol. Biol. Ecol. 12:1074-1084.

FAO.1974.The husbandry and Health of The Domestic Buffalo. Ed. W.R.Cockrill. Food and Agriculture organization of the United Nations,Rome,Italy.

Feldman, M.W., A. Bergman, D.D. Pollock, and D.B. Goldstein. 1996. Microsatellite genetic distances with range constraints Analytic description and problems of estimation. Genetics 145:207-216.

Feldman, M.W., A. Bergman, D.D. Pollock, and D.B. Goldstein. 1996. Microsatellite genetic distances with range constraints Analytic description and problems of estimation. Genetics 145:207-216.

Field, D.; Wills,C. (1996). Long, polymorphic microsatellite in simple organisms. Proceeding of Royal Society of London B, Biological Sciences. 263: 209-215.

.

Goldstein, D.B., L.A. Zhivotovsky, K. Nayar, A. Ruiz Linares, L.L. Cavalli-Sforza, and M.W. Feldman. 1996. Statistical properties of the variation at linked microsatellite loci -- implications for the history of human Y-chromosome Mol. Biol. Evol. 13:1213-1218.

Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L., and Feldman, M.W. 1995a. An evaluation of genetic distances for use with microsatellite loci. Genetics 139: 463-471.

Hochgeschwender, V. and N.B.Brennan. 1991. Identifying the genome new ways for finding the needle in a haystack. Bioassays. 13: 139-144.

Iannuzzi L, Di Meo GP, Perucatti A, Schibler L, Incarnato D, Gallagher D, Eggen A, Ferretti L, Cribiu EP, Womack J. The river buffalo (Bubalus bubalis, 2n = 50) cytogenetic map: assignment of 64 loci by fluorescence in situ hybridization and R-banding. : Cytogenet Genome Res. 2003; 102(1-4): 65-75.

Imperial Council of Agriculture Research, 1941: Definition of the characteristic of cattle and Nili-Ravi and Surti buffaloes Miscellaneous Bulletin No.47: 20.1

Indian Council of Agriculture Research, 1960: Definition of the characteristic of cattle and buffalo breeds in India. New Delhi –Bulletin No. 86.

Jafferys, A.J., V.Wilson and S.L.Thein. 1985. Hypervariable 'minisatellite' regions in human DNA. Nature. 314: 67-73.

Jarne, Phillippe and Pierre J.L. Lagoda. 1996. "Microsatellites, form molecules to populations and back." Trends in Evolution and Ecology, vol. 11, no. 10, October 1996, pp. 424-429.

Jeffreys, A. F.; Murry, J.; Neumann, R. (1998). High resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. Molecular cell. 2:267-273.

Kimura M and Crow JF, 1964. The number of alleles that be maintained in a finite population. Genetics 49:725-738.

Kimura, M. and J.F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725-738.

Liu, L.; Dybvig, K.; Panangala, V. S.; Van, Santen, V. L.; French, C. T. (2000). GAA trinucleotide repeat region regulates M9*/pMGA gene expression in Mycoplasma gallsepticum. Infection and Immunity. 68: 871-876.

Liu, Z.; Tan, G.; Li, P. Dunham, R. A. (1999). Transcribed dinucleotide microsatellites and their associated genes from channel catfish Ictalurus punctatus. Biological and Biophysical Research Communication. 259: 190-194.

Luikart G, Allendorf F W, Cornuet J M and Sherwin W B. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity 89(3): 238-47.

Luikart G, Allendorf FW, Cornuet JM & Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89: 238-247.

Luikart G. 1997. Usefulness of molecular markers for detecting population bottlenecks and monitoring genetic change. Ph. D. Thesis. University of Montana, Missoula, USA.

Mattapallil MJ, Ali S. Analysis of conserved microsatellite sequences suggests closer relationship between water buffalo Bubalus bubalis and sheep Ovis aries. DNA

Nair PG, Balkrishnan M, Yadu YR (1986) The Toda buffaloes of Nilgiris.Buffalo

Journal, 2, 167-178.

Neha (2006) Multilocus Genotyping for Diversity Analysis in Toda Buffaloes. Thesis submitted to Singhania Institute of Law Management Science and Technology, Pacheri Bari, Rajasthan.

Nei M, Maruyama T and Chakraborty R. 1975. The bottleneck effect and genetic

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89:583-590.

Ohta, T. (1982a) Linkage disequilibrium with the island model. Genetics. 101: 139-155.

Paetkau, D. and Strobeck, C. 1994. Microsatellite analysis of genetic variation in black bear populations. Mol. Ecol. 3: 489-495.

Paetkau, D. and Strobeck, C. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. Mol. Ecol. 4: 519-520.

Pollock, D.D., A. Bergman, M.W. Feldman and D.B. Goldstein. 1996. Microsatellite behavior with range constraints: Parameter estimation and improved distance estimation for use in phylogenetic reconstruction. Submitted to Theor. Popul. Biol.

Queller, D.C., Strassmann, J.E., and Hughes, C.R. 1993. Microsatellites and Kinship. TREE 8: 285-2 Reynolds, J., Weir, B.S., and Cockerham, C.C. (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics.

Richard, G. F.; Paques,F. (2000). Mini and microsatellite expansions: the recombination connection. EMBO Reports, 11: 122-126.

S.Kumar et al. (2006). Genetic variation and relationhip among eight Indian riverine

Shriver, M.D., Jin, L., Chakraborty, R., and Boerwinkle, E. 1993. VNTR allele frequency distributions under the SMM: A computer simulation approach. Genetics 134: 983-993lihood approach to tting equilibrium models of microsatellite evolution. Mol. Biol. Evol. 18: 413{417

Slatkin, M. 1991. Inbreeding coefficients and coalescence times. Genet. Res. 58: 167- 175.

Stalling, R. L.; Ford, A.F.; Nelson,D.; Torney, D.C.; Hildebrand, C. E.; Moyzis, R. K. (1991). Evolution and ditribution of (GT) n repetitive sequences in mammalian genomes. Genomics. 10: 807-815.

A E. 2000. Animal Genetic Resources of India Cattle and Buffalo. 6:19-215.

Tantia M S, Vijh P K, Vijh R K, Kumar P, Joshi B K, Nivsarkar A E abd Sahai R. 1996. Siri: The Cattle of Eastern Himalayas. Animal Genetic Resources Information 19: 41-49.

Tantia MS Vijh R K Mishra Bina Bharani K and Arora R(2006). Multilocus genotyping to study the population structure in three buffalo populations of India. Asian Austalsian Journal of Animal Sciences 19: 1070-1078.

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17: 6463-6471.

Tonhati, H., Vascocellos, F,B., Alburque,L.G.,2000.Genetic aspects of productive & reproductive traits in a Murrah buffalo herd in Sao Paulo,Brazil,J.Anim.Breed, genet,117,331-339.

Van Hooft WF, Hanotte O, Wenink PW, Groen AF, Sugimoto Y, Prins HH, Teale A. Applicability of bovine microsatellite markers for population genetic studies on African buffalo (Syncerus caffer). Anim Genet. 1999 Jun; 30(3): 214-20.

Weber, W. and Wong, C. (1993) Human Molecular Genetics, 2:1123-1128.

Weissenbach et al. (1992).”A second generation linkage map of human genome,” nature, 359(6390): 794-801

Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395-420.

Yeh F. C., Boyle T., Rongcai Y., Ye Z. and Xian J. M. 1999 POPGENE version 3.1 (http://www.ualberta.ca/-fyeh/fyeh).