



## Distribution of satellite markers in water buffalo genome

UPASNA SHARMA<sup>1</sup>, PRIYANKA BANERJEE<sup>2</sup>, JYOTI JOSHI<sup>3</sup>, INDU KHATRI<sup>4</sup> and RAMESH K VIJH<sup>5</sup>

National Bureau of Animal Genetic Resources, Karnal, Haryana 132001 India

Received: 3 June 2011; Accepted: 18 October 2011

### ABSTRACT

Genomic markers have wide range of applications and are of several types depending upon the size and the utility. These markers are preferred over several other markers like protein or blood groups. The genomic markers are highly polymorphic, ubiquitously distributed and easier to automate. Microsatellites are short, tandem-repeated DNA sequences which are widely distributed among genomes. Advances in sequence technologies and bioinformatics have completely changed the way the markers are identified. The buffalo whole genome sequence assembly was utilised using bioinformatics tools for extraction of genomic markers, viz. microsatellites, minisatellites and big tandem repeats. Large numbers of microsatellites (3.84 lakh) were identified on all the autosomes and sex chromosome. More than 85% of the microsatellites were dinucleotides while 14% were trinucleotides. The total number of minisatellites and big tandem repeats were 4655 and 752 respectively. The number of minisatellites and big tandem repeats are very small compared to the number of microsatellites.

**Key words:** Big Tandem repeats, Buffalo, Microsatellites, Minisatellites

Cattle and buffalo genome has large synteny (Amaral *et al.* 2008). Cattle (*Bos taurus*) and river buffalo (*Bubalus bubalis*) chromosomes can be matched arm to arm at the cytogenetic level. While the cattle genome consists of 29 acrocentric autosomes and a pair of XY sex chromosomes, the river buffalo genome has 5 biarmed and 19 acrocentric autosomes and XY sex chromosomes (Iannuzzi 1994). All buffalo chromosome arms have homology to single bovine acrocentric chromosome. Bbu1 is reported to be a fusion of Bta1 and 27, Bbu2 equals Bta2 and 23, Bbu3 equals Bta8 and 19, Bbu4 equals Bta5 and 28, and Bbu5 equals Bta16 and 29 at the cytogenetic level with state of the art banding procedures (Iannuzzi 1994). All the other chromosomes have a one-to-one correspondence between the 2 species (Nahas *et al.* 2001). Similar results were obtained using radiation hybrid panel (Amaral *et al.* 2008).

Tandem-repeat loci are composed by clusters of different copy numbers of tandem-repeat sequences, which are classified according to the unit size of the repeats: microsatellites or simple sequence repeats (SSR) for unit of 1 to 6 bp, minisatellites for 11 to 100 bp, and big tandem satellites for sequence periods longer than 100 bp (Chambers and MacAvoy 2000). Microsatellites are ubiquitous and are highly polymorphic. The tandem-repeat loci can be used to study gene duplication and identification as marker for certain disease conditions, genome wide association studies, linkage

analysis, quantitative trait loci (QTL) identification and marker assisted selection (MAS). They are also implicated as regulators of gene expression e.g. at levels of transcription, alternative splicing or impact control. As the buffalo genome sequence has been published (Tantia *et al.* 2011), an immediate need was felt to analyze the genome sequence to identify genomic markers. Thus, this effort has been made in order to find out the number and location of minisatellites, microsatellites and the longer repeats chromosome-wise in genome of water buffalo.

### MATERIALS AND METHODS

The whole genome assembly of the water buffalo was downloaded from the web server <http://210.212.93.84/cgi-bin/gb2/gbrowse/bovine/> (Tantia *et al.* 2011) for analysis of various types of repeats which included microsatellites, minisatellites and big tandem repeats. We utilized the software Unipro Ugene (<http://ugene.unipro.ru/>), and utilized the plugin Repeat Finder and carried out the analysis of each of the 24 autosomes and X chromosome for microsatellites, minisatellites and big tandem repeats. The minimum tandem size was 2, 10 and 100 for microsatellites, minisatellites and big tandem repeats. The occurrence of tandem repeat was kept at minimum of 6 to qualify as microsatellite as the repeats less than 6 are unlikely to be polymorphic. The number of microsatellites, minisatellites and long satellites thus obtained were analyzed chromosome-wise in the buffalo genome.

Present address: <sup>1</sup> Research Associate (upasna30@gmail.com), <sup>2, 3, 4</sup> Senior Research Fellows and <sup>5</sup> Principal Scientist.

## RESULTS AND DISCUSSION

The next generation sequencing reads were utilized to make a genome guided assembly of buffalo against cattle with a coverage of 90% of the cattle genome. All the cattle chromosomes were covered with >30X coverage and had no ambiguous mapping in consensus sequence (Tantia *et al.* 2011). Gaps however remained for only 7% of the assembly against cattle genome Btau 4.0. 52 million nucleotide variations were observed in the aligned assembly of buffalo genome. All the autosomes and X chromosome of the *Bubalus bubalis* were screened for the presence of the microsatellites, minisatellites and big tandem repeats. The number of microsatellites per chromosome has been shown through a bar diagram in Fig. 1. As the diagram depicts, chromosome 1, exhibit the maximum number of microsatellites followed by chromosome 2, whereas, the least number of microsatellites were observed in chromosome 24. Almost equal numbers of microsatellites are present in the chromosome 3 and X. The distribution of various types of microsatellites (di, tri, tetra, penta or hexa) has been given in Table 1.

As the chromosome 1 to 5 are banded (resulting from probable centric fusion events), the number of repeats found in them far exceed other autosomes. Dinucleotide

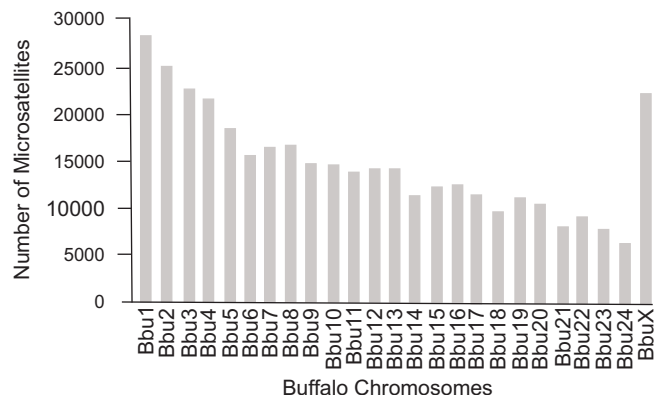


Fig. 1. Distribution of number of microsatellites in buffalo genome.

microsatellite repeats were maximum in chromosome Bbu1 followed by Bbu2, Bbu3 and BbuX. This may be attributed to relative contribution of these chromosomes to total buffalo genome. Similar pattern of abundance was observed for trinucleotide microsatellite marker repeats. Tetranucleotide repeats are maximum in chromosome 1 and 5 with maximum 153 of repeats. Pentanucleotide repeats were found to be maximum i.e. 118 in chromosome 5. The hexanucleotide repeats were very few in number, maximum 14 has been observed in chromosome 4 and X. Chromosome 20 and 23 were not found to have even a single hexanucleotide microsatellite marker. The total numbers of microsatellite markers were maximum in chromosome 1, ascribed to its longer size and minimum in chromosome 24 due to its smaller size. Pentanucleotides and hexanucleotides were very less as compared to dinucleotides, trinucleotides and tetranucleotides. Among these repeats around 85% of repeats were dinucleotides, 14% trinucleotides and rest attributed tetra-, penta- and hexanucleotides markers.

The tetra, penta and hexanucleotide repeats are not only scarcely abundant but also less useful since they exhibit very little polymorphism as they have low mutation rate. However assuming tetranucleotide repeats follow step-wise mutation model these repeats are best available markers for certain demographic parameters of the populations like genetic bottleneck studies (Cornuet and Luikart 1996).

The minisatellite markers are distributed in tandem in the buffalo chromosomes. Mutations in these repeats often have fascinating phenotypic consequences (Gemayel *et al.* 2010). In buffalo genome 4655 minisatellites were observed, almost equally distributed in all the chromosomes of buffalo except the chromosome 3, that reveals the maximum number of minisatellites i.e. 1192. The minimum number of minisatellites i.e. 72 was present in buffalo chromosome 21 (Fig. 2).

The big tandem repeats were very few in number being maximum (77) in Bbu 2 and minimum (8) in Bbu 7 (Fig. 2). In total, 752 Big tandem repeats were present in buffalo genome covering around 2.25 lakh nucleotides.

Table 1. Microsatellite types in terms of repeats in buffalo chromosomes

| Buffalo chromosome | Di-   | Tri- | Tetra- | Penta- | Hexa- |
|--------------------|-------|------|--------|--------|-------|
| Bbu 1              | 26070 | 3567 | 153    | 75     | 8     |
| Bbu 2              | 22652 | 3261 | 129    | 53     | 7     |
| Bbu 3              | 20750 | 3002 | 127    | 46     | 9     |
| Bbu 4              | 19833 | 2829 | 97     | 56     | 14    |
| Bbu 5              | 16725 | 2235 | 153    | 118    | 10    |
| Bbu 6              | 13823 | 2005 | 67     | 39     | 3     |
| Bbu 7              | 14388 | 2112 | 62     | 47     | 1     |
| Bbu 8              | 14976 | 2061 | 94     | 40     | 6     |
| Bbu 9              | 12869 | 1836 | 79     | 22     | 3     |
| Bbu 10             | 12946 | 1969 | 69     | 34     | 7     |
| Bbu 11             | 11594 | 1767 | 56     | 27     | 4     |
| Bbu 12             | 12475 | 1692 | 77     | 29     | 3     |
| Bbu 13             | 12482 | 1547 | 64     | 37     | 3     |
| Bbu 14             | 9896  | 1292 | 54     | 17     | 1     |
| Bbu 15             | 10699 | 1432 | 48     | 19     | 11    |
| Bbu 16             | 10657 | 1452 | 56     | 28     | 4     |
| Bbu 17             | 9664  | 1267 | 55     | 20     | 5     |
| Bbu 18             | 8033  | 1034 | 42     | 13     | 1     |
| Bbu 19             | 9566  | 1258 | 46     | 16     | 2     |
| Bbu 20             | 8793  | 1201 | 61     | 22     | 0     |
| Bbu 21             | 6839  | 971  | 48     | 12     | 1     |
| Bbu 22             | 7645  | 1042 | 28     | 14     | 2     |
| Bbu 23             | 6351  | 922  | 35     | 16     | 0     |
| Bbu 24             | 5390  | 647  | 31     | 12     | 1     |
| Bbu X              | 20621 | 2408 | 113    | 50     | 14    |

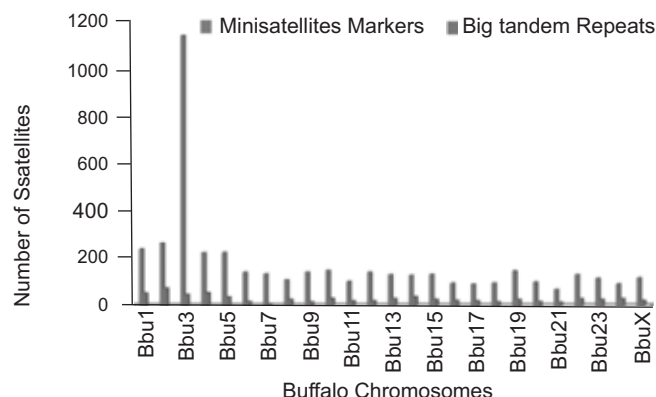


Fig. 2. The distribution and abundance of minisatellites and big tandem repeats in buffalo genome.

The comparative analysis of various repeats like microsatellites, minisatellites and big period tandems are correlated with the size of the chromosomes in buffalo. The first 5 autosomes and X chromosome exhibit the maximum number of repeats, while the minimum number were recorded in chromosome 24 (Table 2). The % distribution of tandem repeats are microsatellites (61%), minisatellites (26%) and big tandem repeats (13%) (Fig. 3).

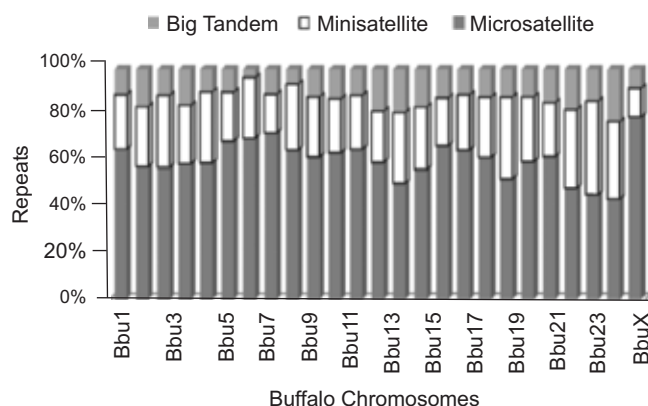


Fig. 3. Relative distribution (%) of microsatellites, minisatellites and big period tandems.

Thus, *Bubalus bubalis* genome scan revealed the presence of microsatellites, minisatellites and big period tandems to analyse the polymorphism in the buffalo genome. These results can be useful for the analysis of QTL, milk yielding genes and biotic and abiotic stress tolerance and many other studies like linkage disequilibrium analysis. There is usual genomic conservation within the family bovidae. The cattle microsatellite markers are ubiquitously distributed in cattle genome. A high density bovine genetic map was released having 3960 markers (Ihara *et al.* 2004). The genetic maps accelerate fine mapping and cloning of genes and mapping for various genetic diseases in livestock (Barendse *et al.* 1997). There is also a large amount of genomic conservation of cattle microsatellite including related wild species like

Table 2. Number of microsatellites, minisatellites and big tandem repeats in buffalo chromosomes

| Buffalo chromosomes | Microsatellites | Minisatellites | Big tandem repeats |
|---------------------|-----------------|----------------|--------------------|
| Bbu1                | 29873           | 250            | 55                 |
| Bbu2                | 26102           | 275            | 77                 |
| Bbu3                | 23934           | 304            | 49                 |
| Bbu4                | 22829           | 233            | 57                 |
| Bbu5                | 19241           | 235            | 38                 |
| Bbu6                | 15937           | 148            | 19                 |
| Bbu7                | 16610           | 140            | 8                  |
| Bbu8                | 17177           | 114            | 27                 |
| Bbu9                | 14809           | 148            | 16                 |
| Bbu10               | 15025           | 156            | 32                 |
| Bbu11               | 13448           | 107            | 27                 |
| Bbu12               | 14276           | 147            | 27                 |
| Bbu13               | 14133           | 137            | 31                 |
| Bbu14               | 11260           | 134            | 40                 |
| Bbu15               | 12209           | 139            | 27                 |
| Bbu16               | 12197           | 100            | 23                 |
| Bbu17               | 11011           | 94             | 21                 |
| Bbu18               | 9123            | 100            | 16                 |
| Bbu19               | 10888           | 154            | 28                 |
| Bbu20               | 10077           | 104            | 19                 |
| Bbu21               | 7871            | 72             | 15                 |
| Bbu22               | 8731            | 137            | 31                 |
| Bbu23               | 7324            | 121            | 27                 |
| Bbu24               | 6081            | 96             | 31                 |
| BbuX                | 23206           | 122            | 23                 |
| Total % of repeats  | 61%             | 26%            | 13%                |

gaur (*Bos gaurus*) (Nguyen *et al.* 2007). The cattle microsatellite have been reported to work well in zebu cattle (*Bos javanicus*) (Hishida *et al.* 1996), yak (Nguyen *et al.* 2005), buffalo (Moore *et al.* 1995, Navani *et al.* 2002 and Hishida *et al.* 1996) and African buffalo (*Syncerus caffer*) (Van hooft *et al.* 1999). However, the number of polymorphic loci are significantly reduced across the species within the family. Till date, the various research workers were utilising heterologous markers after testing them for polymorphism (Vijh *et al.* 2008, Navani *et al.* 2002). The present study and the publication of whole genome assembly has yielded 384172 microsatellite markers of which the major one are di- (85%) and tri- (14%) nucleotide repeats. Rest of the microsatellites contribute <1% of the repeats. The microsatellites loci are likely to be more polymorphic from the species from which they have been derived. The markers derived from cat did not show polymorphism in tiger while markers derived from tiger exhibited a great deal of polymorphisms (Mondal *et al.* 2009). This present study provides a very large number of microsatellites which are ubiquitously distributed throughout the genome on all the chromosomes of buffalo and thus in future there shall not be any compulsion to use heterologus markers in buffalo.

## ACKNOWLEDGEMENTS

The financial support provided from NAIP ICAR, Component IV (C-1050) is gratefully acknowledged.

## REFERENCES

- Amaral M E J, Grant J R, Riggs P K, Stafuzza N B, Filho E A R, Goldammer T, Weikard R, Brunner R M, Kochan K J, Greco A J, Jeong J, Cai Z, Lin G, Prasad A, Kumar S, Saradhi G P, Mathew B, Kumar M A, Miziara M N, Mariani P, Caetano A R, Galvão S R, Tántia M S, Vijh R K, Mishra B, Kumar S T B, Pelai V A, Santana A M, Fornitano L M, Jones B C, Tonhati H, Moore S, Stothard P and Womack J E. 2008. A first generation whole genome RH map of the river buffalo with comparison to domestic cattle. *BMC Genomics* **9**: 631.
- Barendse W, Vaiman D, Kemp S J, Sugimoto Y, Armitage S M, Williams J L, Sun H S, Eggen A, Agaba M and Aleyasin S A. 1997. A medium-density genetic linkage map of the bovine genome. *Mamm. Genome* **8**: 21–28.
- Chambers G K and MacAvoy E S. 2000. Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology B* **126**: 455–76.
- Gemayel R, Vences M D, Legendre M and Verstrepen K J. 2010. Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences. *Annual Review of Genetics* **44**: 445–77.
- Hishida O, Hanotte O, Verjee Y, Tanaka K, Namikawa T, Teale A and Rege J E O. 1996. Crossspecies amplification and polymorphism of microsatellite loci in Asian bovidae. *Proceedings of the 8th AAAP Animal Science Congress* pp. 354–55. 13–18 October 1996; Tokyo, Japan. Japanese Society of Zootechnical Science.
- Iannuzzi L. 1994. Standard karyotype of the river buffalo (*Bubalus bubalis* L., 2n = 50). Report of the committee for the standardization of banded karyotypes of the river buffalo. *Cytogenetics and Cell Genetics* **67**(2): 102–13.
- Ihara N, Takasuga A, Mizoshita K, Takeda H, Sugimoto M, Mizoguchi Y, Hirano T, Itoh T, Watanabe T, Reed K M, Snelling W M, Kappes S M, Beattie C W, Bennett G L and Sugimoto Y. 2004. A Comprehensive Genetic Map of the Cattle Genome Based on 3802 Microsatellites. *Genome Res.* **(10a)**: 1987–98.
- Mondal S, Karanth K U and Ramakrishnan U. 2009. Why the Indian subcontinent holds the key to global tiger recovery. *Plos Genetics* **5** (8).
- Moore S S, Evans D, Byrne K, Barker J S, Tan S G, Vankan D and Hetzel D J. 1995. A set of polymorphic DNA microsatellites useful in swamp and river buffalo (*Bubalus bubalis*). *Animal Genetics* **26**: 355–59.
- Nahas S M, Hondt H A and Womack J E. 2001. Current status of the river buffalo (*Bubalus bubalis* L.) gene map. *Journal of Heredity* **92** (3): 221–25.
- Nguyen T T, Genini S, Bui L C, Voegeli P, Stranzinger G, Renard J P, Maillard J C and Xuan B. 2007. Genomic conservation of cattle microsatellite loci in wild gaur (*Bos gaurus*) and current genetic status of this species in Vietnam. *BMC Genet.* **8**: 77.
- Nguyen T T, Genini S, Menetrey F, Malek M, Vogeli P, Goe M R and Stranzinger G. 2005. Application of bovine microsatellite markers for genetic diversity analysis of Swiss yak (*Poephagus grunniens*). *Animal Genetics*. **36**: 484–89.
- Navani N, Jain P K, Gupta S, Sisodia B S and Kumar S. 2002. A set of cattle microsatellite DNA markers for genome analysis of riverine buffalo (*Bubalus bubalis*). *Animal Genetics* **33**(2): 149–54.
- Tántia M S, Vijh R K, Bhasin V, Sikka P, Vijh P K, Kataria R S, Mishra B P, Pandey A K, Yadav S P, Sethi R K, Joshi B K, Gupta S C and Pathak K M L. 2011. Whole-genome sequence assembly of the water buffalo (*Bubalus bubalis*). *Indian Journal of Animal Sciences* **81**(5):
- Van Hooft W F, Hanotte O, Wenink P W, Groen A F, Sugimoto Y, Prins H H and Teale A. 1999. Applicability of bovine microsatellite markers for population genetic studies on African buffalo (*Syncerus caffer*). *Animal Genetics* **30**: 214–20.
- Vijh R K, Tántia M S, Mishra B and Bharani Kumar S T. 2008. Genetic relationship and diversity analysis of Indian water buffalo (*Bubalus bubalis*). *Journal of Animal Sciences* **86** (7): 1495–1502.