To Explore the Evolutionary Relationships of Different Species of Livestock Populations with Respect to the Nucleotide Sequences of SRY Gene

Ishaan Gupta*, Jayakumar Sivalingam

Abstract – We obtain curetted nucleotide sequences of SRY gene of four animals from Bovidae family (*Bos indicus*, *Bos grunniens*, *Bos frontalis*, *and Capra hircus*) and Sus scrofa form Suidae family, and then align the sequences to further analyze the sequences for phylogenic inference and similarity among the sequences.

Keywords – SRY gene, phylogenetic, sex determination.

1 INTRODUCTION

Sex determination system refers to the mechanism that decides the development of sexual characters in a living organism. Sex determination can be of different types in different organisms. While in some species of reptiles sex depends on temperature [1], some species of flowers and fish are to known to change sex while going through different life cycles. Contemporaries had predicted the existence of such systems hundreds of years ago. Until the discovery sex chromosomes in 1900, however, only environmental theories of sex determination were popular. While the sex might not be externally distinguishable in the initial stages of life, it can be determined even at the time of fertilization itself. In fact, most animals exhibit genetic sex determination or chromosomal sex determination.

On the basis of sex chromosomes in the organism, Chromosomal sex determination is of three different types – XX/X0 type, where females have two copies of a sex chromosome (XX) and males hove only one sex chromosome (X0, where 0 denotes the absence of allosome); XX/XY type, where females have isogamous pair (XX) but males have heterogametic pair (XY); ZZ/ZW type where females are heterogametic (ZW) and males have a isogamous pair (ZZ). Humans and most mammals have XX/XY type of sex determination.

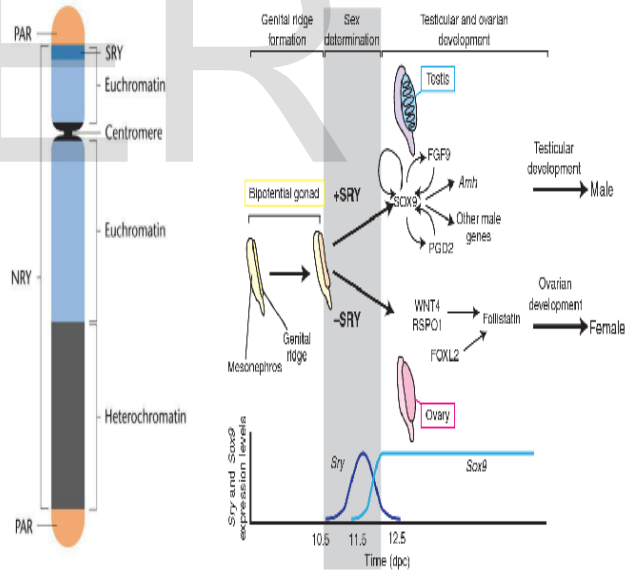


Figure 1 Structure, location and role of SRY gene in animals with Y-centred XX/XY type of sex determination system

Furthermore, XX/XY sex determination can be X-centered or Y-centered. In Y-centered sex determination, presence of Y chromosome ascertains the development male characters. When sex determination is Y – centered, SRY gene located on the Y chromosome suppresses the development of female characteristics and promotes the development of male characteristics. Y chromosome has

two types of distinct regions – PAR (Pseudo-autosomal region) and NRY (non recombination region of Y chromosome) (95%). The NRY has hyperchromatin and euchromatin portions. Functional genes like SRY gene are in euchromatin.

Located near the PAR on the p arm (shorter arm), SRY gene (sex-determining region of Y chromosome) is responsible for development of male phenotype characters. This intronless gene produces a single high mobility group DNA-binding domain, SRY protein, a transcription factor that can bind to regions of testis-specific DNA- the sequence (A/T)ACAA(T/A) in minor groove- and consequently induce a 60-85^{o} bend [2] to activate or enhance ability to promote testis formation [3]. SRY gene expression leads to expression of SOX9 gene which in turn leads to development of wolffian ducts that later form seminiferous tubules and undifferentiated precursor of testes. Then the Leydig cells start secreting testosterone, while Sertoli cells secrete anti-mullerian hormone [4]. Thus SRY protein marks the first step in development of male sexuality. In the absence of SRY gene, mullerian ducts get developed and female sexuality development is initiated. For the experimental confirmation of SRY gene, XX mice were converted to male by the introduction of SRY gene [5].

SRY is a highly conserved gene [6]. This means that nature prevents development of non-silent mutations in the nucleotide sequence of SRY gene sequence for a species.

Nucleotide sequence of any gene can be analyzed for its inter-species and intra-species sequence similarity. Alignment is used to infer sequence homology, which is a function of evolutionary relationship. Alignment can be of two types – global alignment and local alignment. While global alignment is a form of global optimization that forces the alignment to span the entire length of all query sequences [7], local alignment appends a vital step to that-identifying the regions of similarity [8]. Though this additional challenge makes local alignment more complex, it also makes it more preferable as we often need to deal with long sequences that are widely divergent overall. Local alignment finds the point mutations, insertion mutations (indels), deletion mutations (gaps). Alignment inserts gaps between the residues so that identical or similar characters are aligned in successive columns.

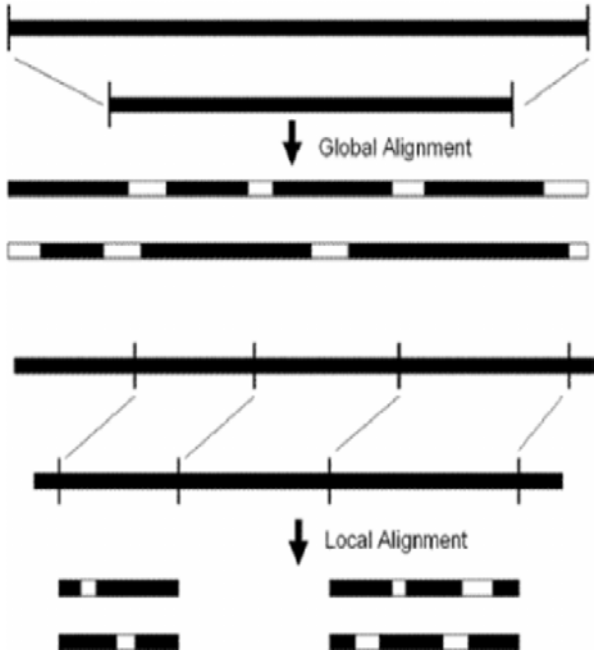


Figure 2 A comparison of local and global sequence alignments [8]

There are a myriad of bioinformatics tools at our disposal that can be used to align given set of sequences.

Developed in 1990, NCBI BLAST is one of the most popular bioinformatics tool for alignment of single reference sequence with multiple sequences [9].

MEGA can carry out multiple sequence alignment and can further perform various useful statistical representation tasks-compute nucleotide frequencies, highlight conserved sites, construct distance matrix, and make phylogenic inferences [10]. MEGA offers 2 options for alignment algorithms- CLUSTAL W and MUSCLE – both of which are standard algorithms.

2 REVIEW OF LITERATURE

SRY gene sequence is inherited without non-silent mutations under normal circumstances. After all, a change in SRY protein sequence, i.e. non-silent mutation in SRY gene sequence-could adversely affect the organism's sexual development. This is practicable in part because SRY gene lies in NRY region and does not face recombination. It is thus an ideal gene for comparing mammalian species for molecular phylogeny. Moreover, databases like nucleotide database of NCBI contain multiple sequences of SRY gene of most mammals that can be saved in FASTA format. For molecular phylogeny, we need to compare multitude of bases at each position. However, for that sequences need to be aligned first.

For beginner-level projects, MEGA software is recommended for alignment. MEGA can construct different types of phylogenetic trees. The UPGMA tree's algorithm uses the distance values to repeatedly cluster and estimate branch length.

BLAST is another local alignment tool that can align given query sequence with sequences from the entire database or given set of reference nucleotide sequences. It allows users to set parameters for protein or nucleotide sequence alignment on the basis of similarity of organisms and gives coverage and identity score.

3 METHODOLOGY

NCBI Nucleotide database was used as source of curetted data. For each of the five species (*Sus scrofa, Bos indicus, Bos grunniens, Bos frontalis, and Capra hircus*) complete or partial nucleotide sequences of SRY gene were collected. The sequences were saved in FASTA format.

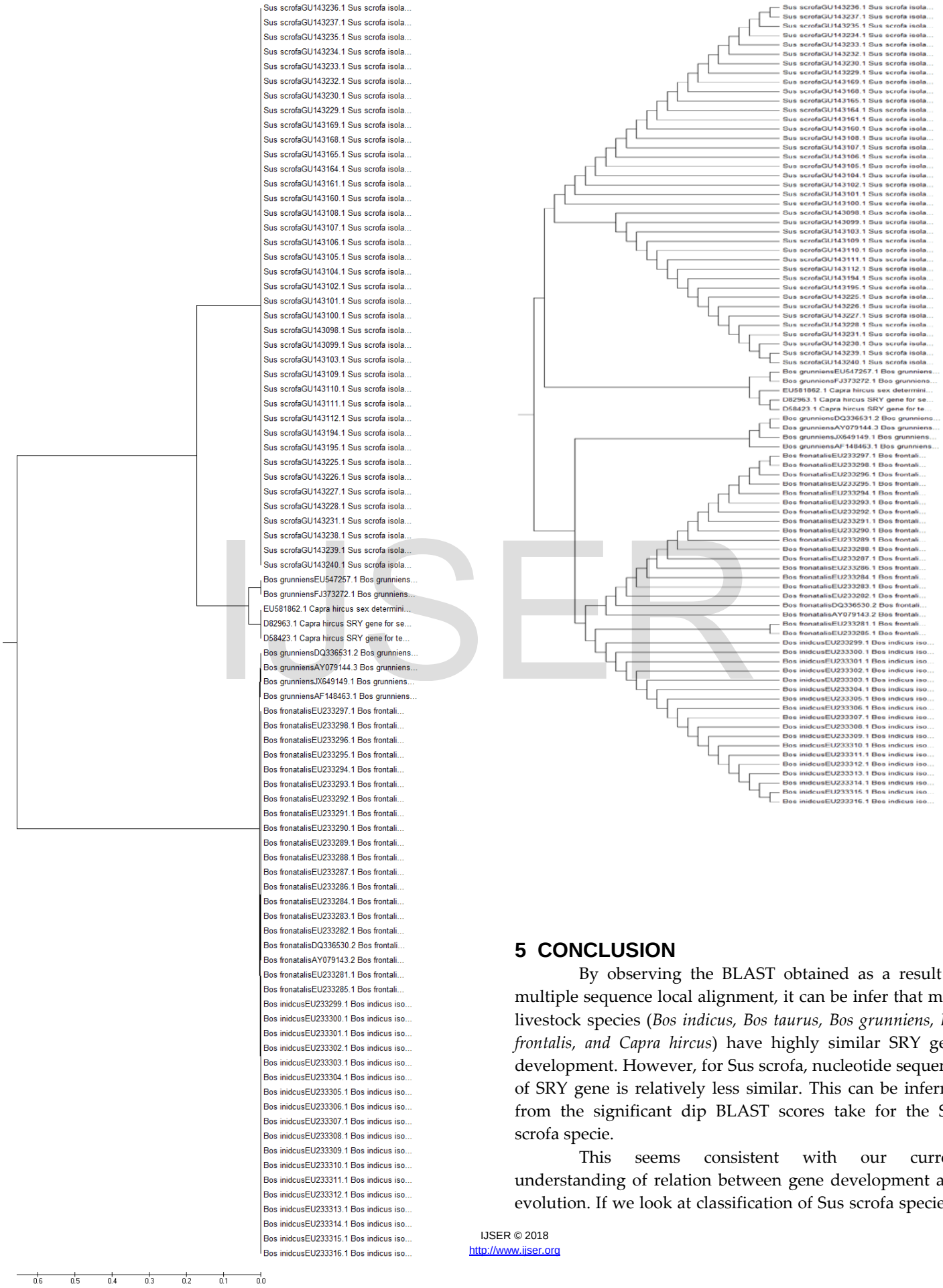
Next, alignment of all the sequences was performed using CLUSTAL W algorithm in MEGA software. While MEGA can build different types of phylogenetic trees, UPGMA tree was built for the study. Default parameters were used for multiple sequence alignment, distance matrix computation, and phylogenetic tree construction.

To quantitatively analyze the similarity of a species with other species, NCBI BLAST was used.

4 RESULT

From the phylogenetic tree it can be seen that the species form distinct clusters. In fact, none of the species share a quaternary branch in the UPGMA tree constructed. High number of conserved sites is obtained. BLAST identity scores for all species but *Sus scrofa* are greater than 90%. Sus scrofa has relatively lower identity scores, but its sequence is still quite similar to the other species examined.

QUERY / SUBJECT	*Bos taurus*	*Bos indicus*	*Bos frontalis*	*Bos grunniens*	*Bubalus bubalis*	*Sus scrofa*	*Capra hircus*
Bos taurus	100	99	99	99	97	76	95
Bos indicus	99	100	99	99	97	78	95
Bos frontalis	99	99	100	99	95	78	95
Bos grunniens	99	99	99	100	96	73	92
Bubalus bubalis	97	97	97	93	100	78	94
Sus scrofa	76	78	78	75	78	100	82
Capra hircus	95	95	95	93	94	82	100



5 CONCLUSION

By observing the BLAST obtained as a result of multiple sequence local alignment, it can be infer that most livestock species (*Bos indicus, Bos taurus, Bos grunniens, Bos frontalis, and Capra hircus*) have highly similar SRY gene development. However, for Sus scrofa, nucleotide sequence of SRY gene is relatively less similar. This can be inferred from the significant dip BLAST scores take for the Sus scrofa specie.

This seems consistent with our current understanding of relation between gene development and evolution. If we look at classification of Sus scrofa specie, it

belongs to the family Suidae, while all the other species concerned belong to the family Bovidae. We have segregated the two families on the basis of phenotypic features like hooves-or-toes, but in general two species closer on evolutionary basis tend to have more similar nucleotide sequences of gene, especially conserved genes like SRY gene.

REFERENCES

[1] Claude Pieau, Marc Girondot, Noëlle Richard-Mercier, Gisèle Desvages, Mireille Dorizzi, Patrick Zaborski. Temperature sensitivity of sexual differentiation of gonads in the European pond turtle. Journal of Experimental Zoology 1994; 270: 86–93.

[2] Harley and Goodfellow, The biochemical role of SRY in sex determination, 1994.

[3] Berta P, Hawkins JR, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN, Fellous M. Genetic evidence equating SRY and the testis-determining factor 1990; 29; 348 (6300):448-50.

[4] Mittwoch U. The race to be male. New Scientist, 1988; 120 (1635): 38–42.

[5] Bullejos M, Koopman P. Spatially dynamic expression of Sry in mouse genital ridges. Dev. Dyn. 2001; 221: 201-205.

[6] Tomofumi Otake, Asato Kuroiwa. Molecular mechanism of male differentiation is conserved in the SRY-absent mammal, Tokudaia osimensis'2001.

[7] Pankaj Agarwal and S. A. M. Rizvi — Solving Sequence Alignment Problem using Pipeline Approach BIJIT, 2009; Vol. 1 No. 2; ISSN 0973 – 5658.

[8] Cheetham, James & K. H. A. Dehne, Frank & Pitre, Sylvain & Rau-Chaplin, Andrew & J. Taillon, Peter. Parallel CLUSTAL W for PC clusters. 2003; 2668: 300-309. 10.1007/3-540-44843-8_32.

[9] Ma J., Zhang L. (2010) Modern BLAST Programs. In: Heath L., Ramakrishnan N. (eds) Problem Solving Handbook in Computational Biology and Bioinformatics. Springer, Boston, MA.

[10] Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6 (Tamura, Stecher, Peterson, Filipski, and Kumar 2013)- www.megasoftware.net/