



Indonesian River Buffalo Molecular Phylogeny Compared to Other Mammals Based on *STAT1* Sequence

(Filogeni kerbau Indonesia dibandingkan mamalia lain berdasarkan runutan nukleotida gen *STAT1*)

Fuad Hasan¹, Armyn Hakim Daulay¹, Ferdy Saputra², Isyana Khaerunnisa^{3*}

¹Department of Animal Science, Faculty of Agriculture, Universitas Sumatera Utara, Medan, Indonesia

²Indonesian Research Institute for Animal Production, Bogor, Indonesia

³Research Center for Applied Zoology, National Research and Innovation Agency (BRIN), Bogor, Indonesia

ABSTRACT. Genes differ in sequence, size, and functional domains among species. According to studies, *STAT1* provides information on the rate of evolution that correlates with its function in the immune system. *STAT1* is also considered a genetic marker for economic traits in mammals. Studying sequence comparison is an important issue in bioinformatic study and can explain phylogenetic. Therefore, this study aimed to identify the molecular phylogeny of river buffalo (*Bubalus bubalis*) and other mammals based on *STAT1* gene sequences. This study used 7 *STAT1* sequences from Ensembl (*Bos grunniens*, *Bos indicus*, *Bos Mutus*, *Capra hircus*, *Cervus hanglu yarkandensis*, *Moschus moschiferus*) and previous studies (*Bubalus bubalis*). The sequences were analyzed using the MEGA X 10.2.6 software to observe the nucleotide composition and the phylogeny (based on UPGMA). The adegenet package in the R 4.0.0 software is used to observe the *STAT1* sequence dimensionally among mammals. The *STAT1* sequence has almost similar diversity among the livestock of the same genus. Based on the *STAT1* sequence, *Bubalus bubalis* has closer genetic proximity to the genus *Bos* than to another genus. In conclusion, we found *STAT1* is more dynamic in evolution and more conserved and found in the similar related genus.

Keywords: Mammals, river buffalo, *STAT1*

ABSTRAK. Gen berbeda dalam urutan, ukuran, dan domain fungsional di antara spesies. Menurut penelitian sebelumnya, *STAT1* memberikan informasi tentang laju evolusi yang berkorelasi dengan fungsinya dalam sistem kekebalan. *STAT1* juga dianggap sebagai penanda genetik untuk sifat bernilai ekonomi pada mamalia. Studi perbandingan urutan merupakan isu penting dalam studi bioinformatika dan dapat menjelaskan filogenetik. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi filogeni molekuler kerbau sungai (*Bubalus bubalis*) dan spesies mamalia lain berdasarkan sekuens gen *STAT1*. Penelitian ini menggunakan 7 sekuens *STAT1* yang diambil dari Ensembl (*Bos grunniens*, *Bos indicus*, *Bos mutus*, *Capra hircus*, *Cervus hanglu yarkandensis*, *Moschus moschiferus*) dan penelitian sebelumnya (*Bubalus bubalis*). Sekuen dianalisis menggunakan program MEGA X 10.2.6 untuk melihat komposisi nukleotida dan filogeni (berdasarkan UPGMA). Adegnet *package* dalam program R 4.0.0 digunakan untuk mengamati urutan *STAT1* secara dimensional diantara mamalia. Sekuen *STAT1* memiliki keragaman yang hampir sama di antara ternak dari genus yang sama. Berdasarkan sekuens *STAT1*, *Bubalus bubalis* memiliki jarak genetik yang lebih dekat dengan genus *Bos* dibandingkan dengan genus lainnya. Sebagai kesimpulan, kami menemukan *STAT1* lebih dinamis dalam evolusi dan lebih terkonservasi serta ditemukan dalam genus terkait yang serupa.

Keywords: Kerbau sungai, mamalia, *STAT1*

INTRODUCTION

Buffaloes are livestock that is closely related to Indonesian society. In addition to producing meat and milk, in Indonesia, buffalo are also used as draught power, tourism attractions, and savings that are ready to be sold when needed. Most of the buffalo population in Indonesia is swamp buffalo, which is used as a meat producer. Meanwhile, the population of river buffalo in Indonesia is only less than 1% (Talib *et al.*, 2014). In Indonesia, the river buffalo only develops in North Sumatra as a milk producer. Based on

history and studies, the river buffalo (*Bubalus bubalis bubalis*) originates from and is domesticated in the Indian subcontinent (Kumar *et al.*, 2007; Colli *et al.*, 2018; Zhang *et al.*, 2020). It is believed that river buffalo were brought by the Indian people to North Sumatra (Talib *et al.*, 2014). River buffaloes are characterized by a black body and curved horns (Zhang *et al.*, 2020).

Compared to swamp buffalo, river buffalo has not received special attention from the government for developing river buffalo as a milk source. Studies on the genetic diversity of swamp buffalo in Indonesia have been carried out based on morphological approaches (Anggraeni *et al.*, 2011), mtDNA cytochrome oxidase subunit I (Saputra *et al.*, 2013), microsatellite markers (Saputra *et al.*, 2020) and cytochrome b genes

*Email Korespondensi: isyana.khaerunnisa@brin.go.id

Diterima: 25 April 2021

Direvisi: 18 Agustus 2021

Disetujui: 1 Desember 2021

DOI: <https://doi.org/10.17969/agripet.v22i1.20889>

(Rusdin *et al.*, 2020). Swamp buffalo diversity in functional genes was also reported in GH and GHR genes (Andreas *et al.*, 2010), GHRH and Pit-1 (Sumantri *et al.*, 2010), and κ -casein (Rini *et al.*, 2014). Meanwhile, studies on the genetic diversity of Indonesian river buffalo are still very limited.

The Signal transducer and activator of transcription 1 (*STAT1*) gene is an important molecule in the interferon signaling pathway (Ramana *et al.*, 2000; Deng *et al.*, 2015). Together with *STAT2*, *STAT1* is a type I and type III IFN signaling mediator (Au-Yeung *et al.*, 2013). The *STAT1* gene is also a regulator of mammary gland differentiation and the diversity in this gene is reported to affect buffalo milk production (Deng *et al.*, 2016a; Du *et al.*, 2019). Information on the diversity of functional genes such as *STAT1* can be used for the construction of river buffalo phylogeny tree with other mammal species. This study aimed to identify the molecular phylogeny of river buffalo based on the *STAT1* gene sequence.

MATERIALS AND METHODS

Data Source

Table 1. Reference of *STAT1* sequences used in this study (according to <https://www.ensembl.org/index.html>)

No	Species	Popular Name	Ensembl ID
1.	<i>Bos grunniens</i>	Domestic yak	ENSBGRG00000001789
2.	<i>Bos indicus</i> -hybrid	Brahman	ENSBIXG00005011100
3.	<i>Bos mutus</i>	Wild yak	ENSBMUG00000015961
4.	<i>Capra hircus</i>	Goat	ENSCHIG00000009728
5.	<i>Cervus hanglu yarkandensis</i>	Yarkand deer	ENSCHYG00000021214
6.	<i>Moschus moschiferus</i>	Siberian musk deer	ENSMMSG00000009167

RESULTS AND DISCUSSION

River Buffalo Partial *STAT1* Gene Sequence

Figure 1 shows the alignment of the *STAT1* gene sequence for river buffalo and six other mammal species. The DNA sequences of intron 16 to intron 18 of the *STAT1* gene in river buffalo produced 1207 bp of nucleotides. The *STAT1* sequence of river buffalo is shorter than other mammals, which are more than 1500 bp (Figure 1). The difference of about 300 bp in river buffalo is due to deletion in intron 18. The partial sequences obtained include 2 exons, i.e exon 17 and exon 18. Both exons are conserved in river buffalo and other mammal species. Based on the *STAT1* sequence, river buffalo are close to cattle and are found in the same cluster as wild yak and domestic yak. Meanwhile, the *STAT1* sequence for goat, Siberian musk deer, and Yarkand deer

A total of 7 sequence data of *STAT1* collected from Ensembl (<https://www.ensembl.org/index.html>) were used in this study including *Bos grunniens*, *Bos indicus*, *Bos mutus*, *Capra hircus*, *Cervus hanglu yarkandensis*, and *Moschus moschiferus*. The river buffalo sequence of *STAT1* gene is based on Hasan *et al.* (2021), covering sequence area around 28159 to 29356 bp according to *Bos indicus* *STAT1* sequence (Ensembl ID: ENSBIXG00005011100).

Data Analysis

Chromatogram resulted from DNA sequencing, as previously reported by Hasan *et al.* (2021), were analyzed using FinchTV, MEGA X (Kumar *et al.*, 2018), and BioEdit (Hall, 2011). As a comparison, several data on the *STAT1* gene sequence of other mammals were used as shown in Table 1. The phylogenetic tree was constructed following UPGMA (unweighted pair grouping method of arithmetic averages) methods. The dendrogram was drawn by MEGA X software (Kumar *et al.*, 2018). Principal coordinate analysis was carried out using adegenet (Jombart, 2008) in the R 4.0.0 program (R Core Team, 2017).

formed a separate cluster from river buffalo (Figure 1).

Goats have the longest sequence (1861 bp) in the *STAT1* gene compared other species i.e. cows (1505 bp), yaks (1504 bp), deer (1576 bp), and river buffalo (1207 bp). River buffaloes have a lower G + C nucleotide composition than cattle (*B. indicus*) and yaks (*B. mutus* and *B. grunniens*) (Table 2). However, river buffaloes have a higher A + T nucleotide composition than cattle (*B. indicus*) and Yak (*B. mutus* and *B. grunniens*). In eukaryotes, the composition of the nucleotides often varies between regions of the genome. Genes that use G + C-rich codons also have introns and flanking sequences that are relatively rich in G + C (Sharp, 2001). In the leptin gene, red jungle fowl and Pekin duck have a high G + C composition of 67.7 and 73.3% (Beauclair *et al.*, 2019). Ross *et al.* (2013) suggests a long sequence

with a high G + C composition will complicate the amplification process and cause bias in the sequencing process. Therefore, the *STAT1* gene in mammals should be easily amplified and sequenced to obtain the entire nucleotide of the

gene. The *STAT1* gene itself functions to provide instructions for making proteins that are involved in various functions of the immune system (Boisson-Dupuis *et al.*, 2012).



Figure 1. UPGMA of river buffalo compared to other mammals based on *STAT1* gene. Grey boxes show the area covered by primers used in this study.

Table 2. Nucleotide composition of *STAT1* sequences

Species	Length	G+C Content	A+T Content	A	C	G	T	S
<i>B. indicus</i>	1505	42.19%	57.81%	26.45	20.33	21.86	31.36	0
<i>B. mutus</i>	1504	42.22%	57.78%	26.46	20.28	21.94	31.32	0
<i>B. grunniens</i>	1504	42.35%	57.65%	26.33	20.28	22.07	31.32	0
<i>M. moschiferus</i>	1576	40.93%	59.07%	26.84	19.73	21.19	32.23	0
<i>C. hanglu</i>	1580	40.70%	59.30%	27.47	19.68	21.01	31.84	0
<i>C. hircus</i>	1861	43.52%	56.48%	25.85	21.55	21.98	30.63	0
<i>B. bubalis</i>	1207	41.84%	58.08%	26.01	21.29	20.55	32.06	0.08

River Buffalo Phylogenetic Analysis Based on *STAT1*

Deng *et al.* (2015) stated that based on *STAT1*, phylogenetic analysis revealed that a close genetic relationship was found between buffalo and cattle. This is supported by the research we conducted using the UPGMA method and the Principal Coordinate Analysis. However, unlike UPGMA, the Principal Coordinate Analysis provides a dimensional picture, where *B. indicus*, *B. mutus*, and *B. grunniens* are located in a cluster adjacent to *B. bubalis* (Figure 2). *C. hircus*, *C. hanglu*, and *M. moschiferus* separated to form separate clusters. In this case, the same genus has

strong genetic affinities in functional genes such as *STAT1*. Deng *et al.* (2016b) found the *STAT1* gene on SNP2338 with genotype TC and SNP5558 with genotype GT had a positive correlation with high milk production 305 days. SNP2965, SNP3666, and SNP4007 are known to be associated with a high percentage of milk protein. Based on the CSMM66, ILSTS61, and LSTS17 microsatellite loci, it is known that river buffalo in North Sumatra have low genetic diversity because they are found in a monomorphic state (Saputra *et al.*, 2020). This is possible because of the use of the same male, so cement is needed from a superior male.

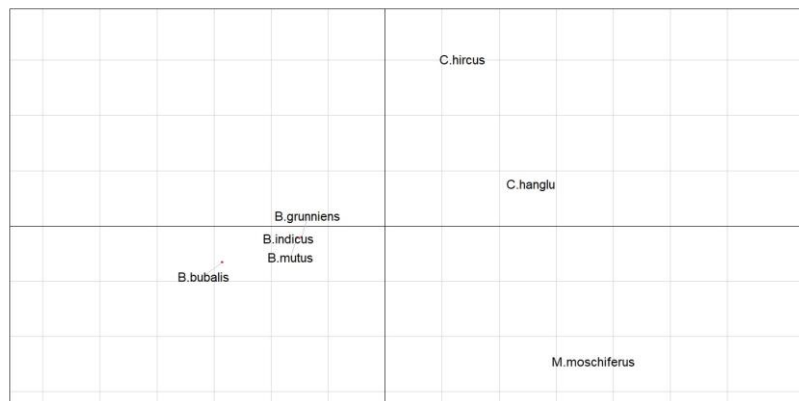


Figure 2. Principal Coordinate Analysis of river buffalo compared to other mammals based on *STAT1* gene

CONCLUSIONS

In conclusion, we found *Bos* cluster close to *Bubalus bubalis* than *Capra hircus*, *Cervus hanglu yarkandensis*, and *Moschus moschiferus*. This study also proved that *STAT1* is more dynamic in evolution and more conserved and found in the similar related genus.

ACKNOWLEDGEMENT

This research was funded by the Talenta USU Research Implementation April 28, 2020 (Contract No.: 278/UN5.2.3.1/PPM/SPP-TALENTA USU/2020).

REFERENCES

- Andreas, E., Sumantri, C., Nuraini, H., Farajallah, A., Anggraeni, A., 2010. Identification of GH|AluI and GHR|AluI genes polymorphisms in Indonesian buffalo. *J. Indonesian Trop. Anim. Agric.* 35(4):215-221.
- Anggraeni, A., Sumantri, C., Praharani, L., Dudi, Andreas, E., 2011. Estimasi jarak genetic kerbau rawa local melalui pendekatan analisis morfologi. *JITV.* 16(3):199-210.
- Au-Yeung, N., Mandhana, R., Horvath, C.M., 2013. Transcriptional regulation by *STAT1* and *STAT2* in the interferon JAK-STAT pathway. *JAKSTAT.* 2(3): e23931.
- Beauclair, L., Ramé, C., Arensburger, P., Piégu, B., Guillou, F., Dupont, J., Bigot, Y., 2019. Sequence properties of certain GC rich avian genes, their origins and absence from genome assemblies: case studies. *BMC Genomics* 20(1):734.
- Boisson-Dupuis, S., Kong, X.F., Okada, S., Cypowyj, S., Puel, A., Abel, L., Casanova, J.L., 2012. Inborn errors of human *STAT1*: allelic heterogeneity governs the diversity of immunological and infectious phenotypes. *Curr Opin Immunol.* 24(4):364-78.
- Colli, L., Milanesi, M., Vajana, E., Iamartino, D., Bomba, L., Puglisi, F., Del Corvo, M., Nicolazzi, E.L., Ahmed, S.S.E., Herrera, J.R.V., Cruz, L., Zhang, S., Liang, A., Hua, G., Yang, L., Hao, X., Zuo, F., Lai, S., Wang, S., Liu, R., Gong, Y., Mokhber, M., Mao, Y., Guan, F., Vlaic, A., Vlaic, B., Ramunno, L., Cosenza, G., Ahmad, A., Soysal, I., Ünal, E.O., Ketudat-Cairns, M., Garcia, J.F., Utsunomiya, Y.T., Baruselli, P.S., Amaral, M.E.J., Parnpai, R., Drummond, M.G., Galbusera, P., Burton, J., Hoal, E., Yusnizar, Y., Sumantri, C., Moiola, B., Valentini, A., Stella, A., Williams, J.L., Ajmone-Marsan, P., 2018. New Insights on Water Buffalo Genomic Diversity and Post-Domestication Migration Routes from Medium Density SNP Chip Data. *Frontiers in Genetics.* 9:53
- Deng, T., Pang, C., Lu, X., Zhu, P., Duan, A., Tan, Z., Huang, J., Li, H., Chen, M., Liang, X., 2016a. De Novo Transcriptome Assembly of the Chinese Swamp Buffalo by RNA Sequencing and SSR Marker Discovery. *J. Anim. Sci.* 94:927–935.
- Deng, T., Pang, C., Lu, X., Zhu, P., Duan, A.Q., Liang, X.W., 2016b. Associations between polymorphisms of the *STAT1* gene and milk production traits in water buffaloes. *J. Anim. Sci.* 94(3): 927–935.
- Deng, T., Pang, C., Zhu, P., Liao, B., Zhang, M., Yang, B., Liang, X., 2015. Molecular cloning and expression analysis of the *STAT1* gene in the water buffalo (*Bubalus bubalis*). *Trop Anim Health Prod.* 47(1):53-9.
- Du, C., Deng, T., Zhou, Y., Ye, T., Zhou, Z., Zhang, S., Shao, B., Wei, P., Sun, H., Khan, F.A., Yang, L., Hua, G., 2019. Systematic analyses for candidate genes of milk production traits in water buffalo (*Bubalus bubalis*). *Anim Genet.* 50(3):207-216.
- Hasan, F., Khaerunnisa, I., Daulay, A.H., 2021. Signal transducer and activator of transcription 1 gene polymorphism in Indonesian river buffalo. *IOP Conf. Series: Earth and Environ. Sci.* 782: 022087
- Hall, T., 2011. BioEdit: An Important Software for Molecular Biology. *GERF Bulletin of Bioscience* 2:6.
- Jombart, T., 2008. Adegenet: an R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11):1403-1405.
- Kumar, S., Nagarajan, M., Sandhu, J.S., Kumar, N., Behl, V., 2007. Phylogeography and domestication of Indian river buffalo. *BMC Evolutionary Biology.* 7:186.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular

- Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35(6):1547–1549.
- R Core Team, 2017. R: A Language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, <https://www.R-project.org/>
- Ramana, C.V., Chatterjee-Kishore, M., Nguyen, H., Stark, G.R., 2000. Complex roles of *STAT1* in regulating gene expression. *Oncogene*. 19:2619-2627.
- Rini, A.O., Sumantri, C., Damayanthi, E., 2014. K-casein gene polymorphisms in riverine and swamp buffalo in Indonesia. *J. Indonesian Trop. Anim. Agric.* 39(1):1-9.
- Ross, M.G., Russ, C., Costello, M., Hollinger, A., Lennon, N.J., Hegarty, R., Nusbaum, C., Jaffe, D.B., 2013. Characterizing and measuring bias in sequence data. *Genome Biol.* 14(5):R51.
- Rusdin, M., Solihin, D.D., Gunawan, A., Talib, C., Sumantri, C., 2020. Genetic Variation of Eight Indonesian Swamp-Bufferalo Populations Based on Cytochrome b Gene Marker. *Trop. Anim. Sci. J.* 43(1): 1-10.
- Saputra, F., Jakaria, J., Anggraeni, A., Sumantri, C., 2020. Genetic Diversity of Indonesian Swamp Buffalo Based on Microsatellite Markers. *Trop. Anim. Sci. J.* 43(3): 191-196.
- Saputra, F., Jakaria, J., Sumantri, C., 2013. Genetic Variation of mtDNA Cytochrome Oxidase Subunit I (COI) in Local Swamp Buffaloes in Indonesia. *Media Peternakan* 36(3):165-170.
- Sharp, P.M., 2001. Codon Usage Bias. Editor(s): Sydney Brenner, Jefferey H. Miller, Encyclopedia of Genetics, Academic Press.
- Sumantri, C., Diyono, R., Farajallah, A., Anggraeni, A., Andreas, E., 2010. Pemanfaatan Famili Gen Hormon Pertumbuhan (GH, GHR, GHRH dan PIT-1) untuk Mendeteksi Keragaman Genetik Kerbau di Kabupaten Pandeglang dan Lebak Provinsi Banten. *JITV*. 15(4): 286-296.
- Talib, C., Herawati, T., Hastono, 2014. Strategi Peningkatan Produktivitas Kerbau melalui Perbaikan Pakan dan Genetik. *Wartazoa*. 24(2):83-96.
- Zhang, Y., Colli, L., Barker, J.S.F., 2020. Asian water buffalo: domestication, history and genetics. *Anim Genet.* 51(2):177-191.