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, Army Hakim Daulay1, Ferdy Saputra2, Isyana Khaerunnisa3*

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 sekuen gen STAT1. Penelitian ini menggunakan 7 sekuen 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). Adegenet package dalam program R 4.0.0 digunakan untuk mengamati urutan STAT1 secara dimensional diantara mamalia. Sekuen STAT1 memiliki keceragaman yang hampir sama di antara ternak dari genus yang sama. Berdasarkan sekuen 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
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 k-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>
<thead>
<tr>
<th>No</th>
<th>Species</th>
<th>Popular Name</th>
<th>Ensembl ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bos grunniens</td>
<td>Domestic yak</td>
<td>ENSBGRG00000001789</td>
</tr>
<tr>
<td>2</td>
<td>Bos indicus-hybrid</td>
<td>Brahman</td>
<td>ENSBIXG00000011100</td>
</tr>
<tr>
<td>3</td>
<td>Bos mutus</td>
<td>Wild yak</td>
<td>ENSBMUG00000015961</td>
</tr>
<tr>
<td>4</td>
<td>Capra hircus</td>
<td>Goat</td>
<td>ENSCHIG00000009728</td>
</tr>
<tr>
<td>5</td>
<td>Cervus hanglu yarkandensis</td>
<td>Yarkand deer</td>
<td>ENSCHYG00000021214</td>
</tr>
<tr>
<td>6</td>
<td>Moschus moschiferus</td>
<td>Siberian musk deer</td>
<td>ENSMMSG0000009167</td>
</tr>
</tbody>
</table>

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 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% (Beaulclairet 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).

Table 2. Nucleotide composition of STAT1 sequences

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

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 SNP 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 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.

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/TALENTA USU-2020).

REFERENCES


Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular


