Exploring SNPs (Single Nucleotide Polymorphisms) of Myostatin gene in coding region in Bali cattle

W L N Aliyya, R R Noor and Jakaria

1 Post Graduate Student, Animal Production and Technology, Faculty of Animal Science, IPB University, Indonesia
2 Animal Production and Technology, Faculty of Animal Science, IPB University, Indonesia

E-mail: ladhunka@gmail.com

Abstract. The gene MSTN is a member of the superfamily growth gene transforming growth factor β (TGF-β) that functions to suppress muscle growth. The purpose of this study was to analyze the polymorphisms of MSTN genes in the coding region (exon 1, 2, and 3) in Bali cattle kept at the Breeding Center for Bali cattle (BPTU-HPT) in Denpasar. The number of samples used was 52 heads male of Bali cattle aged 2-3 years. SNP identification was done by amplifying the coding region, exon 1, exon 2 and exon 3 MSTN genes using direct sequencing method. Analysis of MSTN gene sequences in Bali cattle was carried out by using the BioEdit and MEGA7 programs. The present study found a total of 18 SNP mutations namely point mutation that are 4 SNP (g.313 C>A, g.324 T>A, g.330 T>G, g.400 G>A) in exon 1, 1 SNP (g.2609 G>A) in exon 2, and 13 SNP (g.4821 C>A, g.4838 C>T, g.4842 A>C, g.4849 T>C, g.4868 C>A, g.4873 C>G, g.4900 T>G, 4905 C>G, g.4957 C>G, g.4959 C>A, g.4969 T>A, g.4974 T>C, g.5044 C>A) in exon 3. SNPs found in the MSTN gene in Bali cattle might be used as candidates for Marker Assisted Selection, especially in Bali cattle.

1. Introduction

The gene MSTN or Growth Differentiation Factor-8 (GDF-8) is a member of the superfamily transforming growth factor β (TGF-β) that plays a role in the nature of growth. The myostatin gene in Bos indicus is located on chromosome 2 with 3 exons and 2 introns. The exon region is the region in which nucleotide sequences of genes are expressed to synthesize proteins and express genes and where gene transcription occurs. The gene MSTN has 375 amino acids [1]. Amino acids function as producers of proteins MSTN with a molecular weight of 26 kDa [1]. Myostatin is specifically expressed during the development of the embryonic and adult skeletal muscles, which function as negative regulatory proteins for the growth of body muscles. Some cattle breeds such as Piedmontese, Belgian Blue and Blond'Aquitaine exhibit polymorphisms in this gene. Where this polymorphism is directly related to the double muscled phenotype [2]. Myostatin gene mutations in Belgian Blue cattle are natural mutations due to the presence of recessive homozygous alleles. The gene MSTN can be used as an efficient selection model with shorter time, especially in the selection of quality breeders and prospective males, for example for selection of Bali cattle, because the study of MSTN genes in Bali cattle is still very limited, especially in SNP exploration myostatin gene.

Bali cattle (Bos javanicus) is one of the genetic resources of native Indonesian cattle produced from domestication of the Banteng (Bos sondaicus) which is recognized by the FAO as a nation of separate...
cattle outside Bos taurus and Bos indicus [3]. Based on National standard, Bali cattle family is determined through Minister of Agriculture Decree Number 325/Kpts/OT.140/1/2010 with population distribution in all regions of Indonesia. Bali cattle have the potential to be used as the main domestic meat producer because of their superior production and reproduction, easy to adapt to low-quality feed, low subcutaneous fat, a high percentage of carcasses which is 52.72-57.59% [4].

One of the efforts to improve the genetic quality of livestock (Bali cattle) is through the selection of the nature of growth. It aims to obtain genetic characteristics that can significantly increase meat production. Therefore genetic information related to the exploration of the SNP (Single Nucleotide Polymorphisms) MSTN gene in the coding area (exon 1, 2 and 3) in Bali cattle is needed. The results of this exploration are expected to be used as information and a basis for determining candidate genes in conducting Marker Assisted Selection (MAS) and as a direction in determining breeding strategies, especially in Bali cattle.

2. Materials and methods

2.1. Materials

DNA samples used were collected from 52 blood samples from Bali cattle aged 2-3 years at BPTU-HPT Denpasar. Blood samples were taken from the jugular vein as much as ± 5 mL and stored at ± 4°C. DNA extraction was carried out in the Molecular Genetics Laboratory of the Faculty of Animal Science, Bogor Agricultural University based on Kit DNA extraction procedure GeneAid. The primers used for MSTN gene fragment amplification are sequential data from the NCBI (National Center for Biotechnology Information) genBank with access number AY794986.1 designed using primer 3 (bioinfo.ut.ec/primer3-0.4.0) and primer stats (www.bioinformatics.org). The primary sequences are presented in table 1.

| Table 1. Primary sequences of MSTN gene. |
|-------------------------------|-----------------|-----------|------|
| Fragment | Primer Sequence | Temperature | Product |
| Exon 1 | F: 5’-CAAGTTGTCTCTCAGACTGG-3’  
R: 5’-CTCCTCCTTACATAACAAGCC-3’ | 55ºC | 608 bp |
| Exon 2 | F: 5’-GATTGATATGGAGGTGTTCG-3’  
R: 5’-TAGGATGTGAAATGGGACAC-3’ | 56ºC | 622 bp |
| Exon 3 | F: 5’-CTCCTTCTTCTTCTTCCATAACAG-3’  
R: 5’-AGGGGAAGACCTTCCATGTT-3’ | 60ºC | 451 bp |

F = forward, R = reverse.
Source: GenBank AY794986.1

DNA amplification (PCR) was carried out in thermocycler machine (Applied Biosystem 9700) according to primers. DNA amplification reagents were consisted of 12.5 µL PROMEGA 2× Green Master Mix, 9.9 µL Nuclease Free Water, 0.3 µL primer forward, 0.3 µL primer reverse, put into a 1.5 µL tube and then homogenized. The PCR reagent mixture was then distributed as much as 23 µL to each tube containing 2 µL of DNA samples and then put into a PCR machine. Furthermore, DNA is ready to be amplified with adjusted temperatures. Electrophoresis is an advanced stage of the research procedure after PCR amplification. The materials used are PCR product DNA, agarose gel, 0.5×solution, TBEpeqgreen DNA/RNA stain and 100 bp ladder. Sequencing method in this research was done by 1st Base Laboratory, Selangor Malaysia, on position of forward or reverse. SNP identification data analysis was performed using the Bioedit [5] and MEGA7 [6].

2.2. Genotype and allele frequencies

Genotype and allele frequencies (Xi) are calculated based on methods [7]:

Genotype frequencies

\[ X_i = \frac{\sum_{n=1}^{N} n_i}{N} \]
### Allele frequencies

\[ X_i = \frac{(2n_{ii} + \sum n_{ij})}{2N} \]

- \( X_i \) = genotype frequency
- \( \sum n_{ij} \) = allele frequency
- \( n_{ii} \) = number of genotyped individuals ii
- \( n_{ij} \) = number of genotyped individuals ij
- \( N \) = number of individuals sampled

#### 2.3. Heterozygosity

Genetic diversity will be estimated using observational heterozygosity (Ho) and expected heterozygosity (He) by using of the Nei methods [7];

\[ H_o = \sum n_{ij}^2 \quad H_e = 1 - \sum X_i^2 \]

- \( H_o \) = heterozygosity observation (population)
- \( H_e \) = value of heterozygosity expectations
- \( n_{ij} \) = number of heterozygous individuals
- \( N \) = number of individuals observed
- \( X_i \) = allele frequency
- \( q \) = number of alleles.

#### 2.4. Hardy-Weinberg

Hardy-Weinberg equilibrium and allele frequencies obtained will be analysed using the methods [8];

\[ \chi^2 = \sum \frac{(Obs - Exp)^2}{Exp} \]

- \( \chi^2 \) = chi Square value
- \( Obs \) = observation value of genotype ii
- \( Exp \) = expected value of genotype ii

### 3. Results and discussion

#### 3.1. PCR amplification

PCR amplification of the MSTN gene using the machine (Applied Biosystem 9700) has been successfully carried out with a product length of 608 bp at a temperature condition of annealing 55°C in exon 1, a product length of 622 bp at a temperature condition annealing 56°C in exon 2, and product length 451 bp at a temperature condition of annealing 60°C in exon 3. The success of amplification is determined by several factors, e.g., the sample used, the amplification time, the denaturation time, the temperature and the primary concentration used. Visualization MSTN gene amplification results can be seen in figure 1.
3.2. Polymorphisms of the myostatin gene

The polymorphisms of the myostatin gene are identified using the direct sequencing method. Sequencing results found four polymorphic SNPs in exon one, e.g. in the base position g.313 C> A, g.324 T> G, g.400 G> A. One polymorphic SNP at exon two namely at base position g.2609 G> A, and 13 polymorphic SNP at exon three at base position g.4821 C> A, g.4838 C> T, g.4842 A> C, g.4849 T> C, g.4868 C> A, g.4837 C> G, g.4900 T> G, g.4905 C> G, g.4957 C> G, g.4959 C> A, g.4969 T> C, g.5044 C> A. These results are more than the SNP found by [2] in exon one, two, and three, 14 polymorphic SNPs in Nellore cattle. The visualization of partial mutations in the MSTN gene can be seen in figure 2.

Genotype frequency is the ratio of the number of a genotype to a population by calculating the ratio between the number of certain genotypes in each population while the allele frequency is the ratio of alleles to all alleles in an SNP in the population [9]. SNP is said to be polymorphic if it has an allele frequency ≤0.99 for large populations and ≤0.95 for smaller populations [8]. Table 2 presents the allele frequencies and genotypes of the MSTN gene. All SNP MSTN gene found polymorphic have two homozygous genotypes and heterozygous genotypes. Eighteen mutations were found with 8 points of transversion and 10 points of transition.
Base mutations c.313 C>A, c.330 T>G, c.4821 C>A, c.4842 A>C, c.4868 C>A, c.4900 T>G, c.4959 C>A, c.5044 C>A shows a transversion mutation that is the change in pyrimidine base to purine. Base mutations c.324 T>A, c.400 G>A, c.2609 G>A, c.4838 C>T, c.4849 T>C, c.4873 C>G, c.4957 C>G, c.4969 T>C, c.4905 C>G is a transition mutation. Transition mutations are base substitutions that have the same structure, whereas transversion mutations are base substitutions that have different structures [10]. It is stated in humans that exon regions or coding protein occur mostly in transitional mutations, and few transversion mutations occur [11,12]. Transversion mutations are more influential on changes in the structure and function of regulation of a gene than with transition mutations.

Mutations that show changes in amino acids (non synonymous mutation) namely g.313 stop> tyrosin, g.324 glutamine> valine, g.4821 aspartic acid> tyrosin, g.4842 stop> glutamic acid, g.4849 stop> serine, g.4868 glutamic acid> aspartic acid, g.4873 arginine> proline, g.4957 arginine> proline, g.4959 aspartic acid> tyrosine, g.4969 histidine> leucine, g.4974 asparagine> aspartic acid, g.5044 glycine > valine. While 6 mutations showed no change in bases (synonymous mutation), namely in the base mutation g.330 asparagine, g.2609 valine, g.4838 proline, g.4900 asparagine, and g.4905 glutamic acid. Structurally, synonymous mutations do not cause changes in the structure of amino acids. Still, synonymous mutations can affect conformation (structural stability) and protein function by influencing the transcription mechanism and regulation of RNA, changing the structure of mRNA, and affecting the speed of translation [13,14].

### Table 2. Genotype frequency and MSTN gene allele in Bali cattle genotype.

| Exon | SNP    | Genotype frequency | Allele frequency |
|------|--------|--------------------|------------------|
|      |        | AA     | AB     | BB     | A     | B     |
| 1    | c.313  C>A | 0.90  | 0.1    | -      | 0.95  | 0.05  |
|      | c.324  T>A | 0.96  | 0.04   | -      | 0.98  | 0.02  |
|      | c.330  T>G | 0.96  | 0.04   | -      | 0.98  | 0.02  |
|      | c.400  G>A | -     | 0.21   | 0.79   | 0.11  | 0.89  |
| 2    | c.2609 G>A | 0.96  | 0.04   | -      | 0.98  | 0.02  |
| 3    | c.4821 C>A | 0.94  | 0.06   | -      | 0.97  | 0.03  |
|      | c.4838 C>T | 0.94  | 0.06   | -      | 0.97  | 0.03  |
|      | c.4842 A>C | 0.96  | 0.04   | -      | 0.98  | 0.02  |
|      | c.4849 T>C | 0.94  | 0.06   | -      | 0.97  | 0.03  |
|      | c.4868 C>A | 0.83  | 0.17   | -      | 0.91  | 0.09  |
|      | c.4873 C>G | 0.96  | 0.04   | -      | 0.98  | 0.02  |
|      | c.4900 T>G | 0.92  | 0.08   | -      | 0.96  | 0.04  |
|      | c.4905 C>G | 0.94  | 0.06   | -      | 0.97  | 0.03  |
|      | c.4957 C>G | 0.9  | 0.1    | -      | 0.95  | 0.05  |
|      | c.4959 C>A | 0.92  | 0.08   | -      | 0.96  | 0.04  |
|      | c.4969 T>A | 0.96  | 0.04   | -      | 0.98  | 0.02  |
|      | c.4974 T>C | 0.96  | 0.04   | -      | 0.98  | 0.02  |
|      | c.5044 C>A | 0.96  | 0.04   | -      | 0.98  | 0.02  |

AA=genotype reference (GenBank), AB = genotype heterozygote, BB = genotype mutant, A = allele reference (GenBank), B = allele mutant.

### 3.3. Heterozygosity and Hardy-Weinberg equilibrium

Genetic diversity is an important component of populations. Heterozygosity states the genetic diversity of a population that can be used for selection programs. Noor [9] explains that the diversity of genes can be used as a reference in determining breeding programs that are selected if the population is diverse and the crossing is done if the population is uniform. Heterozygosity values are presented in table 3.
SNP found in the MSTN gene has a difference in Ho and He is small. A lower Ho value than He indicates inbreeding [15]. The results of the HW (χ²) analysis in the MSTN gene are generally balanced (18 SNP). According to [9], factors that affect gene balance in a population are non-random mating, selection, migration, mutation and genetic drift.

Table 3. Heterozygosity and Hardy-Weinberg equilibrium SNP MSTN gene.

| Gene | SNP       | Ho    | He    | χ²  |
|------|-----------|-------|-------|-----|
| MSTN | c.313 C>A | 0.0962| 0.0924| 0.133  |
|      | c.324 T>A | 0.038 | 0.038 | 0.020  |
|      | c.330 T>G | 0.038 | 0.038 | 0.020  |
|      | c.400 G>A | 0.212 | 0.189 | 0.720  |
|      | c.2609 G>A| 0.038 | 0.038 | 0.020  |
|      | c.4821 C>A| 0.058 | 0.056 | 0.046  |
|      | c.4838 C>T| 0.058 | 0.056 | 0.046  |
|      | c.4842 A>C| 0.038 | 0.038 | 0.020  |
|      | c.4849 T>C| 0.058 | 0.056 | 0.046  |
|      | c.4868 C>G| 0.173 | 0.158 | 0.467  |
|      | c.4873 C>G| 0.038 | 0.038 | 0.020  |
|      | c.4900 T>G| 0.077 | 0.074 | 0.083  |
|      | c.4905 C>G| 0.058 | 0.056 | 0.046  |
|      | c.4957 C>G| 0.096 | 0.092 | 0.133  |
|      | c.4959 C>A| 0.077 | 0.074 | 0.083  |
|      | c.4969 T>A| 0.038 | 0.038 | 0.020  |
|      | c.4974 T>C| 0.038 | 0.038 | 0.020  |
|      | c.5044 C>A| 0.038 | 0.038 | 0.020  |

The results of the study found 18 SNP s in MSTN gene namely 4 SNPs at exon 1 in base position g.313 C>A, g.324 T>A, g.330 T>G, g.400 G>A. 1 polymorphic SNP on exon 2 ie at base position g.2609 G>A, and 13 polymorphic SNP at exon three namely at base position g.4821 C>A, g.4838 C>T, g.4842 A>C, g.4849 T>C, g.4868 C>A, g.4873 C>G, g.4900 T>G, g.4905 C>G, g.4957 C>G, g.4959 C>A, g.4969 T>A, g.4974 T>C, g.5044 C>A. Based on research, from 18 mutations, there are 8 points of transversion mutations and 10 points of transition mutations. Mutations that show changes in amino acids (non-synonymous mutation) there are 12 mutations, while 6 mutations indicate no change in bases (synonymous mutation).

4. Conclusion
The study found a total of 18 SNP mutations and the SNPs found in the MSTN gene in Bali cattle might be used as candidates for Marker Assisted Selection, especially in Bali cattle.

Acknowledgments
This work was financially supported by PTM project from Directorate General of Strengthening Research and Development, Ministry of Research Technology and Higher Education of the Republic Indonesia Number: 4391/IT3.L1/PN/2019.

References
[1] Taylor W E, Bhasin S, Artaza J, Byhower F, Azam M, Willard D H, Cull F C and Gonzalez-Cadavid N 2001 Myostatin inhibits cell proliferation and protein synthesis in C2C12 muscle cells. Am. J. Physiol. Endo. Metab. 280 E221–8
[2] Grisolia A B, D’Angelo G T, Porto-Neto L R, Siqueira F and Garcia J F 2009 Myostatine (GDF8) single nucleotida polymorphisms in Nellore cattle Genet. Mol. Res. 8(3) 822-30
[3] Directorate of Generale of Livestock Services 2003 *National Report on Animal Genetic Animal Genetic Resources Indonesia* (Jakarta: Directorate of Generale of Livestock Services of Indonesia) p 14

[4] Hafid H and Rugayah N 2009 Percentage of carcass of bali cattle at various body weights and length of mastery before slaughtering *Nat. Sem. on Animal Husbandry and Veterinary Technology* (Bogor: Pusat Penelitian dan Pengembangan Peternakan)

[5] Hall T A 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT *Nucl. Acids. Symp. Ser.* 41 95-8

[6] Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood evolutionary distance and maximum parsimony methods *J. Mol. Biol. Evol.* 28 2731–9

[7] Nei M and Kumar S 2000 *Molecular Evolution and Phylogenetics* (New York USA: Oxford Univ Press)

[8] Allendrof F W, Luikart G H and Aitken S N 2013 *Conservation and the Genetics of Populations* 2nd Ed (Chicester UK : Wiley-Blackwell Publishing)

[9] Noor R R 2010 *Genetika Ternak* 6th Ed (Jakarta: Penebar Swadaya)

[10] Luo G H, Li X H, Han Z J, Zhang Z C, Yang Q, Guo H F and Fang J C 2016 Transition and transversion mutations are biased towards GC in transposons of chilo suppressalis (lepidoptera: pyralidae) *Genes* 7 1-12

[11] Cargill M *et al* Characterization of single-nucleotide polymorphisms in coding regions of human genes *Nat. Genet.* 22 231-9

[12] Kimura M 1980 A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences *J Mol. Evol.* 16 111–31

[13] Zuben E and Chave K 2011 Understanding the contribution of synonymous mutations to human disease *Nat. Rev. Genet.* 12 683–91

[14] Gonzales-Castejon M, Marin F, Soler-Rivas, Reglero G, Visioli F and Casado R 2011 Functional non-synonymous polymorphisms prediction methods: Current approaches and future developments *Curr. Med. Chemist.* 18 5095-103

[15] Nassiry M R, Javanmard A and Tohidi R 2009 Application of statistical procedures for analysis of genetic diversity in domestic animal populations *Am. J. Anim. Vet. Sci.* 4 136-41