UNIVERSITI PUTRA MALAYSIA

IDENTIFICATION OF LIGAND DEPENDENT NUCLEAR RECEPTOR COREPRESSOR LIKE (LCORL) GENE POLYMORPHISM IN SUMBAWA PONIES (*EQUUS CABALLUS*) USING PCR AND DNA SEQUENCING ANALYSIS METHOD

FRETELDY FREDDY

IDENTIFICATION OF LIGAND DEPENDENT NUCLEAR RECEPTOR COREPRESSOR LIKE (LCORL) GENE POLYMORPHISM IN SUMBAWA PONIES (*EQUUS CABALLUS*) USING PCR AND DNA SEQUENCING ANALYSIS METHOD

FRETELDY FREDDY

A project paper submitted to the

Faculty of Veterinary Medicine, Universiti Putra Malaysia

In partial fulfilment of the requirement for the

DEGREE OF DOCTOR OF VETERINARY MEDICINE

Universiti Putra Malaysia

Serdang, Selangor Darul Ehsan

2020/2021

CERTIFICATION

It is hereby certified that we have read this project paper entitled "Identification of Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) Gene Polymorphism in Sumbawa Ponies (*Equus caballus*) using PCR and DNA Sequencing Analysis Method" by Freteldy Freddy and in our opinion it is satisfactory in terms of scope, quality and presentation as partial fulfilment of the requirement for the course VPD 4999 – Final Year Project.

_____

DR. MOHD SHAHROM BIN SALISI

D.V.M. (UPM), PhD (UPM)

Senior Lecturer

Department of Veterinary Pre-Clinical Science

Faculty of Veterinary Medicine

Universiti Putra Malaysia

(Supervisor)

ASSOC. PROF. DR. NURUL HAYAH KHAIRUDDIN

D.V.M. (UPM), PhD (Glasgow, UK)

Associate Professor

Department of Farm and Exotic Animal Medicine and Surgery

Faculty of Veterinary Medicine

Universiti Putra Malaysia

(Co-Supervisor)

DR. NOR YASMIN ABDUL RAHMAN

D.V.M. (UPM), PhD (UPM)

Senior Lecturer

Department of Veterinary Laboratory Diagnosis

Faculty of Veterinary Medicine

Universiti Putra Malaysia

(Co-Supervisor)

DR. ANNAS SALLEH

D.V.M. (UPM), PhD (UPM)

Senior Lecturer

Department of Veterinary Laboratory Diagnosis

Faculty of Veterinary Medicine

Universiti Putra Malaysia

(Co-Supervisor)

DR. NUR MAHIZA MD ISA

D.V.M. (UPM), PhD (Glasgow, UK)

Senior Lecturer

Department of Veterinary Pathology and Microbiology

Faculty of Veterinary Medicine

Universiti Putra Malaysia

(Co-Supervisor)

ACKNOWLEDGEMENTS

I would like to give praise to the Lord, my Almighty God for giving me courage, strength and patience to carry out my final year project and see it to the end.

Firstly, my heartfelt gratitude to my supervisor, Dr Mohd Shahrom Bin Salisi for giving me an opportunity to conduct this project under his guidance. My utmost appreciation goes to my co-supervisors, Associate Professor Dr Nurul Hayah Khairuddin, Dr Nor Yasmin Abdul Rahman, Dr Annas Salleh and Dr Nur Mahiza Md Isa for being the backbone of the project. Also, I would like to thank my academic advisor, Associate Professor Dr Md. Sabri Yusoff and DVM graduate, Dr Iskandar Shah B Hazlan Shah for lending a helping hand whenever I was feeling anxious.

Next, I truly appreciate the staffs and practitioners in the Large Animal Ward, staffs in the Theriogenology Lab, and especially staffs in the Virology Lab for their provisions and laboratory training as well as assistance with our laboratory work. Not forgetting Associate Professor Datuk Dr Bashir Ahmad Fateh Mohamed for his big help with all things related to equine especially during sampling.

A big shout out to my final year project mates, Siti Norain, Abdul Aziz Rahman and Zul Khidir for their absolute support throughout this project. The time spent on this project will not be as entertaining without our shenanigans, Thank you. Special thanks to Sharlina who had accompanied and aided me in laboratory work.

My forthright acknowledgement towards my family for their endless love, motivation and support. Lastly, I'm giving my sincerest thanks to all my beloved friends, classmates and people who have willingly tried helping me with their abilities. Start from small steps, gain self-confidence and self-esteem. And one day, you will fly.

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

LCORL	Ligand Dependent Nuclear Receptor Corepressor Like
NCAPG	Non-SMC Condensin I Complex Subunit G
GWAS	Genome-Wide Association Study
DVS	Department of Veterinary Services
NCBI	National Center for Biotechnology Information
FEI	Fédération Équestre Internationale
DBKL	Kuala Lumpur City Hall
PCR	Polymerase Chain Reaction
DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
mRNA	Messenger Ribonucleic Acid
ECA 3, 9	Chromosome 3, 9
QTL	Quantitative Trait Locus
SNP	Single Nucleotide Polymorphism
TFIID	Transcription Factor IID
EDTA	Ethylenediaminetetraacetic Acid
PBS	Phosphate Buffered Saline
ddH_2O	Double-Distilled Water
dNTP	Deoxyribonucleotide Triphosphate

ddNTP	Dideoxynucleotide Triphosphates
UV	Ultraviolet
cm	Centimetre
bp	Base Pair
Mb	Mega Base Pair
rpm	Revolution per Minute
ml	Milliliter
µL	Microliter
µM	Micromolar
g	Gram
V	Volt
mA	Milliampere
ºC	Celsius
%	Percentage
∞	Infinity

ABSTRAK

Abstrak daripada kertas projek yang dikemukakan kepada Fakulti Perubatan Veterinar untuk memenuhi sebahagian daripada keperluan kursus VPD 4999 – Projek Tahun Akhir.

PENGENALPASTIAN POLIMORFISME GEN RESEPTOR SEPERTI NUKLEAR LIGDAN BERGANTUNG COREPRESIL PADA KUDA SUMBAWA MENGGUNAKAN KAEDAH PCR DAN ANALISIS PENJUJUKAN DNA

Oleh

FRETELDY FREDDY

2020

Penyelia: Dr. Mohd Shahrom Bin Salisi

Penyelia bersama: Prof. Madya Dr. Nurul Hayah Khairuddin, Dr. Nor Yasmin Abdul Rahman, Dr. Annas Salleh dan Dr. Nur Mahiza Md Isa

Reseptor Seperti Nuklear Ligdan Bergantung Corepresil (LCORL) yang mengekodkan gen protein adalah dicadangkan mempunyai perhubungan dengan saiz rangka tulang dan ukuran ketinggian melalui faktor transkripsi yang dikodkan itu. Kajian Persatuan di Seluruh Genom (GWAS) telah menemui banyak polimorfisme di dalam dan di luar gen ini yang melibatkan ekspresinya bersama dengan perkaitan

dengan ciri-ciri fizikal seperti morfologi badan dalam kepelbagaian baka kuda. Kajian ini bertujuan untuk mengenal pasti kehadiran polimorfisme gen LCORL dalam kuda Sumbawa (*Equus caballus*) menggunakan kaedah reaksi berantai polimerase (PCR) dan kaedah analisis penjujukan DNA. Sebanyak 16 sampel darah dikumpulkan daripada kuda yang sihat dengan komposisi badan yang berbeza-beza. Kemudian pengekstrakan DNA dilakukan terhadap sampel dan kualiti diperiksa dengan biofotometer. Sampel yang diekstrak kemudian digandakan melalui kaedah PCR menggunakan primer yang direka dari hulu urutan gen LCORL. Semua sampel menunjukkan visualisasi jalur PCR dengan ukuran yang dijangkakan antara 300-400 pasangan asas (bp). Setelah pemurnian, penjujukan dan analisis amplikon yang berdekatan dengan LCORL berada pada 320 bp menggunakan pasangan kajian primer. Hasil kajian menunjukkan polimorfisme nukleotida tunggal (SNP) di lokus berhampiran LCORL yang merangkumi alel T dan alel C menunjukkan dua belas (12) kuda adalah TT homozigot dan empat (4) lagi TC heterozigot. SNP ini mempengaruhi proses transkripsi gen yang terlibat dalam perkembangan rangka tulang yang menyebabkan peningkatan ketinggian bahu (wither). Kesimpulannya, penemuan kajian menunjukkan polimorfisme gen LCORL yang jelas dalam kuda Sumbawa tersebut dan boleh dicadangkan sebagai calon penanda ~~calon~~ untuk pemilihan strategik dalam program pembiakan.

***Kata kunci**: kuda sumbawa, gen Reseptor Seperti Nuklear Ligdan Bergantung Corepresil (LCORL), polimorfisme, PCR, analisis penjujukan DNA, polimorfisme nukleotida tunggal (SNP).*

ABSTRACT

An abstract of the project paper presented to the Faculty of Veterinary Medicine in partial fulfilment of the course VPD 4999 – Final Year Project.

IDENTIFICATION OF LIGAND DEPENDENT NUCLEAR RECEPTOR COREPRESSOR LIKE GENE POLYMORPHISM IN SUMBAWA PONIES USING PCR AND DNA SEQUENCING ANALYSIS METHOD

By

FRETELDY FREDDY

2020

Supervisor: Dr. Mohd Shahrom Bin Salisi

Co-supervisor: Assoc. Prof. Dr. Nurul Hayah Khairuddin,

Dr. Nor Yasmin Abdul Rahman, Dr. Annas Salleh and Dr. Nur Mahiza Md Isa

The Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) is a protein coding gene suggested to have an association with the skeletal frame size and height measurements through its encoded transcription factor. Genome-wide association studies (GWAS) have discovered numerous polymorphisms within and outside of this gene involving its expression together with association towards physical traits such as body conformation across horse breeds. This study aims to identify the presence of LCORL gene polymorphisms in Sumbawa ponies (Equus caballus) using

polymerase chain reaction (PCR) and DNA sequencing analysis method. A total of sixteen (n=16) blood samples were collected from healthy ponies with varying body composition. Samples were then subjected to DNA extraction and quality checked with a biophotometer. The extracted samples were then amplified via PCR method using primers designed from upstream of LCORL gene sequence. All samples showed the visualization of PCR bands with expected size between 300-400 base pairs (bp). After purification, sequencing and analysis the amplicons adjacent to LCORL was at 320 bp using the primer pair of study. Results showed the single nucleotide polymorphism (SNP) at the locus near LCORL which included T allele and C allele revealing twelve (12) ponies were homozygous TT and the other four (4) heterozygous TC. This SNP affects the transcription process of genes involved in the development of skeletal bone leading to an increased wither height. Overall, this study ~~our findings~~ demonstrated evident polymorphisms of the LCORL gene in these Sumbawa ponies and can be proposed as a candidate marker for strategic selection in breeding programs.

Keywords: *sumbawa ponies, ligand dependent nuclear receptor corepressor like (LCORL) gene, polymorphism, PCR, DNA sequencing analysis, single nucleotide polymorphism (SNP).*

1.0 INTRODUCTION

1.1 Background

Part of our human history had a connection up to a certain degree with the history of horses. According to a well-known English naturalist Darwin (1859), domestication had demonstrated a dynamic modification and diversification whereby deemed 'artificially' induced, nonetheless comparable in numerous paths towards evolution in the wild by natural selection. This is evident with the variety of horse breeds across the globe within regions of human inhabitants. In Asia, a significant contribution from Mongolian horse ancestors practically gave rise to the diverse breeds of horses (Turk, 2017). Horses were used differently in Malaysia during pre-independence and post-independence. Prior to independence, horses were used mainly as transportation, created a new line of work to aid in economy, and exclusively owned by royalties or elites. Post-independence, horses were involved in many sectors like equestrian sports, recreational activities, tourism, health and therapy, veterinary, and used by government bodies for patrolling and enforcement purposes (Department of Veterinary Services, 2019).

Malaysia imported numerous horse breeds from different countries, however, one that is close to the Bornean region that is derived from the spread of Mongolian horse genetics would be the Sumbawa ponies (Edwards, Langrish, & Houghton, 2000). From this literature, it is mentioned that these ponies hailing from Indonesia, particularly on Sumatra are small at about 127 cm of wither height, exceptionally hardy and co-operative, versatile, strong, and agile. The key feature of a pony as a distinction from the prominent horse would be its height whereby when measured on a smooth

level surface, does not exceed 150 cm without shoes or 151 cm with shoes (Fédération Équestre Internationale, 2007).

Contemporarily, the main reason of researching within the animal genetic field would be identifying genes with major impact on the expression of quantitative traits. One of the likely important genes is that of association with wither height namely the Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene. Quite a number of literatures have recognized association between polymorphism adjacent to the LCORL gene and wither height of horses (He, Zhang, Li, & Liu, 2015; Okuda et al., 2016; Tozaki et al., 2016), for the most part, there have been analogous findings. These results were possible through previous studies that were done to identify chromosome number, location, and polymorphism of the gene with significant association towards its expression for morphological traits (Signer-Hasler et al., 2012; Tetens, Widmann, Kühn, & Thaller, 2013; Metzger, Schrimpf, Philipp, & Distl, 2013). Though many literatures had covered across breeds of horses, the same cannot be said for genetic research on horse breeds within Malaysia and for this study focusing Sumbawa ponies.

LCORL is a protein coding gene and also a transcription factor. Mostafavi et al. (2019) explained its location on the horse chromosome 3 (ECA 3) that is mapped to the quantitative trait locus (QTL) region nearby the vital single-nucleotide polymorphism (SNP). Polymorphisms in relation to this gene were reported to have association with skeletal frame size and height measurements in livestock animals such as cattle (Setoguchi et al., 2011), sheep (La et al., 2019), pigs (Rubin et al., 2012), and horses (Signer-Hasler et al., 2012). These researches were possible through genome-wide association studies (GWAS). This may imply a genetic potential for an exceptionally predictive marker on body size.

1.2 Hypothesis

The hypotheses of this study were:

H_0: There are no polymorphisms of the Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene in Sumbawa ponies as potential markers for selection of equine in breeding programs

H_a: There are polymorphisms of the Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene in Sumbawa ponies as potential markers for selection of equine in breeding programs

1.3 Objective

To identify presence of Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene polymorphism in Sumbawa ponies (*Equus caballus*).

2.0 LITERATURE REVIEW

2.1 Sumbawa Ponies (*Equus caballus*)

Throughout the years, the equine industry has always been distinctively growing steadily across countries. Generally, within the equine industry horses are involved in many sectors such as feed production for potential export, education and training for employment, veterinary as an equine specialist, agrotourism as an attraction or recreational activities, and breeding for improving breeding strategies (Vial & Evans, 2015). It is documented that over the past few years Malaysia had imported quite a number of horses comprising variety of breeds from different countries. Within records from 2010 – 2016 the total number of equines that had been imported was at 4,261. Among the countries of origin for those importation, there were around three in 2016 that originated from Indonesia and possibly even more so at present time (DVS, 2019).

With that being said, a breed of horse that is near to Malaysia particularly within Bornean region with derivation from the spread of Mongolian horse genetics would be the Sumbawa ponies (Edwards et al., 2000). According to Noor (2008), this breed's population are able to withstand harsh conditions making the Sumbawa ponies favourable for development. A historical and scientific investigation conducted by Turk (2017) mentioned that these ponies hailing from Indonesia, particularly on Sumatra are small at about 122 – 127 cm, exceptionally hardy and co-operative, versatile, strong, and agile. This study further explained the significant contribution from Mongolian horse ancestors practically being causal to the diverse breeds of horses comprising numerous Asian and North-western horse breeds.

In horses, the size of their body is an essential point of evaluation for different breeds with regards to function and appearance and is vital for horse classification. The key feature of a horse as a distinction from the pony would be its height whereby when measured on a smooth level surface, is at or exceeds 150 cm without shoes or 151 cm with shoes (FEI, 2007). Therefore, if circumstances are in vice versa then the animal would be classified as a pony (van de Pol & van Oldruitenborgh-Oosterbaan, 2007). However, this may generally differ from other organizations. Chromosomal studies have contributed with the delineation of the equine member species by identifying differences in chromosome numbers. Consequently, the typical domestic horse (*Equus caballus*) possesses a total number of 64 chromosomes (Bailey & Binns, 1998). Moreover, this provides somewhat a platform for continuation of other studies such as gene mapping.

2.2 Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) Gene

Further study on the horse genome was possible around 2007 when the domestic horse's whole genome sequence was established publicly (Wade et al., 2009). This sequence would then become the primary resource for the developed single nucleotide polymorphism (SNP) array of a commercial horse. Some great findings from genome-wide association studies (GWAS) up to present day were identification of SNPs on horse chromosomes within and proximal to genes (Binns, Boehler, & Lambert, 2010; Hill, McGivney, Gu, Whiston, & MacHugh, 2010; Tozaki et al., 2012). Possibility of the large intergenic region without any coding sequences may act as the chromatin regulatory domains (Libioulle et al., 2007). So, within the quantitative trait

locus (QTL), the causative nucleotide variant, that is not a direct neighbour to a gene may have extended regulatory effect towards it.

The Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) is a protein coding gene and also a transcription factor which may have specific binding sites for activation of certain biological processes. A study carried out by Makvandi-Nejad et al. (2012) reported that through GWAS identified four loci containing genetic variation explaining a great number (83%) for size variation in horses. One of the four loci comprised typically the LCORL gene whereby located on chromosome 3 (ECA 3) is recognized constantly as a transcription factor associated with the human height (Lango Allen et al., 2010). Horses, similar to other domestic mammals, possess a small number size loci containing alleles of impactful effect. This is in contrast in human studies nonetheless parallel for the results in domestic dogs (Boyko et al., 2010). Further research conducted by Signer-Hasler et al. (2012) discovered via GWAS as well, two QTLs for morphological traits on ECA 3 and ECA 9. Particularly within ECA 3, the association signal is upstream to the Ligand Dependent Nuclear Receptor Corepressor Like/ non-SMC condensin I complex subunit G (LCORL/ NCAPG) gene which is within a large intergenic region, also referred to as gene desert, at about 1.7 Mb. In addition to this study, the QTL in ECA 3 presented significant association with a number of conformation measurements which also includes the wither height.

The overlapping of genes for both LCORL/ NCAPG were shown as a major QTL mapped on ECA 3 that demonstrated around 18% explanation towards phenotypic variance (Tetens et al., 2013). Taking into consideration of previous literatures for these results, the genetic structure of varying heights in horses is denoted by few genes accompanied by numerous and substantial loci and minimal effects. This finding is similar to Makvandi-Nejad et al. (2012) and Signer et al. (2012). To put into

perspective, this implies body size (wither height) is regarded as a complex trait due to it undefined by a single or a small number of genomic variants be that as it may relatively by minor contributions of numerous variants throughout the genome (Raudsepp, Finno, Bellone, & Petersen, 2019). The understanding of these loci together with its action towards affecting growth or morphological traits has yet to be extensively grasped. Based from the whole genome shotgun sequence for equine on ECA 3 of EquCab 3.0 assembly published in the online database of National Center for Biotechnology Information (NCBI) regarding this particular LCORL gene is approximately 151,555 base pairs (bp) and consists of seven exons and six introns (Signer et al., 2012; Metzger et al., 2013), and is most likely mapped to just about 3q25 chromosome region (Godard et al., 1997).

The first functional study carried out by Metzger et al. (2013) verified via GWAS the highest associated QTL for wither height on ECA 3 proximal to the LCORL gene along with the confirmed single nucleotide polymorphism (SNP) of BIEC2-808543 (Broad institute nomenclature; EquCab3_107,374,136 T>C) whereby there is substitution of thymine (T) to cytosine (C) which is highly deemed as the potential candidate for body size. The understanding of this polymorphism adjacent to the LCORL gene together with its mechanism on affecting growth or morphological traits has yet to be extensively grasped. In short, this polymorphism is located within the putative DNA consensus sequence element (TATA box element) that disrupts the transcription factor binding site which acts as a main component in the transcription process of genes involved in the development of skeletal bone implicating significant alterations of LCORL expression levels whereby when reduced leads to an increased wither height (Abrishami-Moghaddam, Miresmaeli, Sani, & Seifati, 2020). That being so, this first functional study became a foundational framework for the following

genetic researches for this gene and its association towards morphological traits (He et al., 2015; Okuda et al., 2016; Kim et al., 2018; Chandra, 2019; Bai et al., 2020).

2.3 Polymerase Chain Reaction (PCR) and DNA Sequencing Analysis

In general, a scientific technique namely the polymerase chain reaction (PCR) in molecular biology allows amplification of a single or a few copies of a DNA piece, producing thousands to millions of copies of a particular DNA segment or sequence. This method was developed by an American biochemist, Kary Mullis in 1984 that had received the Nobel Prize as well as the Japan Prize in 1993 (Bartlett & Stirling, 2003). During that time, progression towards betterment of this technique was limited due to primer creation and polymerase purification issues (Kleppe et al., 1971). Currently, PCR is commonly used in research labs both medical and biological for varying applications (Joshi & Deshpande, 2010). In view of the fact that it is simple, quick and relatively inexpensive. The amplification can also take place with little amount of the source DNA material, even in the circumstance of a debatable poor-quality source DNA (Paxson, 2008). To simplify, the whole procedure is done in a single tube filled with nuclease free water, primers, thermostable DNA polymerase, nucleotides, and patient sample or DNA template. This tube is then run through three temperature steps, to be specific, denaturation (90 – 97 °C), annealing (50 – 60 °C) and extension (approximately 72 °C) in multiple cycles (at about 25 – 30) by which it is essential to follow specific temperatures for each reaction to occur (Maurya et al., 2005). The amplification produces a resultant PCR product for a specific target DNA, known as an amplicon.

This indispensable scientific technique aids in the diagnosis and investigation of numerous diseases and genes. PCR becomes a necessary preliminary step or requirement even for competing techniques such as DNA chips (Ishmael & Stellato, 2008). Amplicons would then be further analysed by other techniques such as in determining gene defects via hybridization, restriction mapping, or sequencing. Usually more than one polymorphism can be analysed simultaneously (Wrischnik et al., 1987; Wellinghausen et al., 2009). DNA sequencing is the process of identifying the order or sequence of nucleotides in DNA. It pertains to any technology or method used to determine these four bases order which are adenine (A), thymine (T), guanine (G), and cytosine (C). The Sanger method is a mixed-mode technique that revolves around synthesis of a complementary DNA template by usage of natural 2'-deoxynucleotides (dNTPs) and termination of synthesis by utilizing defective dNTPs called 2',3'-dideoxynucleotides (ddNTPs) by DNA polymerase (Sanger, Nicklen, & Coulson, 1977). This repeated cycle of synthesis and termination together with a balanced ratio of dNTP/ddNTP produces a set of nested fragments that can be separated based on size using gel electrophoresis to reveal the DNA sequence. Advancements in fluorescence dyes and detection, and capillary array electrophoresis has led to an automated Sanger sequencing (Metzker, 2005). This method is almost similar differing by usage of terminating ddNTPs, each tagged with specific fluorescent dye that is passed through a laser detection region which excites fluorophores resulting fluorescence emissions of four different colours (Smith et al., 1986; Prober et al., 1987). These colours would then be assigned to base calls and the arrangement of the fluorescent fragments reveals the DNA sequence.

The aligned data set of DNA sequences contain important information on phylogeny, molecular ecology, and population genetic. Throughout the years

numerous statistical and graphical software were developed to extract evolutionary information from DNA sequence data (Rozas & Rozas, 1999; Schneider, Roessli, & Excoffier, 2000). A few software commonly used such as 'ProSeq' (Filatov, 2002) and 'SeqTrace' (Stucky, 2012) provide features which allow showing translation, editing of chromatogram files, contig assembly, sequence alignment and other utilities. Gleeson and Hillier (1991) explained that initially, the DNA sequence file would be checked on its chromatogram for base calls requiring any correction on erroneous bases. Then it may be automatically assembled as a sequence contigs to be aligned (both forward and reverse strands) via 'Sequence/ Pairwise alignment' function to compute a single consensus sequence (Filatov, 2002; Stucky, 2012). The finalized file could then be confidently analysed on certain regions for any polymorphism (Tajima, Misawa, & Innan, 1998).

2.4 Gene Polymorphisms

The term polymorphism is a specific term in genetics and biology which relates to the multiple forms of a gene that can exist. Typically, the words 'poly' and 'morph' together means many forms. Genetic polymorphism is a type of genetic diversity that exists within a population's gene pool (Hartl & Clark, 2007). With that being said, there would be two or more possible forms for a particular phenotype in the same population. Based on Singh (2001) which stated this phenomenon can occur within coding or non-coding part of the DNA sequence for any genetic trait, be it physiological or phenotypic. Polymorphism that had resulted in a certain phenotypic trait among individuals in a population is heritable (Farkas & Holland, 2009). Ismail and Essawi (2012) added to their claim that the occurrence of genetic variation is often

regarded normal variant in a population, however, when a specific allele has a frequency approximately 1% or greater, it is asserted to be genetic polymorphism. Variation of DNA sequence among individuals is probable by the result of chance processes or induced by external agents (Jahromi, Ahmed, Behbehani, & Mohammad, 2014). Genetic polymorphism comprises single nucleotide polymorphisms (SNPs), sequence repeats, insertions, deletions, and recombination. These may also be present outside of genes, within the intergenic region with DNA that does not code for protein (Karki, Pandya, Elston, & Ferlini, 2015). The change for a single base pair such as substituting the nucleotide thymine (T) into cytosine (C) in the genomic DNA of a particular gene is referred to as SNP, common form of polymorphism, that can affect gene functions (Aerts, Wetzels, Cohen, & Aerssens, 2002).

Othman, Abdel-Samad, El Maaty, and Sewify (2012) indicated that molecular markers containing polymorphism at the DNA level is of utmost importance in animal genetics. According to He et al. (2015), genetic markers along with polymorphisms are noted as a tool for improving horse features since it is able to affect morphological traits. Therefore, genetic variation determination in potential marker loci with significant association for certain economic traits of interest is useful in breeding strategies (Ozturk & Altinok, 2017). On horse chromosome 3 (ECA 3) nearby the region of Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene had been demonstrated by researchers possessing highly significant quantitative trait locus (QTL) for equine conformation traits that include frame, head, neck and development (Schröder, 2010). The association of LCORL towards body size, particularly wither height is reported in other horse breeds and populations (Signer-Hasler et al., 2012; Tetens et al., 2013; Staiger et al., 2016; Tozaki et al., 2016). The initial functional study carried out by Metzger et al. (2013) discovered via genome-wide association

studies (GWAS) the highest associated QTL for wither height on ECA 3 proximal to the LCORL gene along with the confirmed SNP of BIEC2-808543 (Broad institute nomenclature; EquCab3_107,374,136 T>C) whereby there is substitution of nucleotide T to C which is highly deemed as the potential candidate for body size. Genetic polymorphisms contribute towards genetic characterization of populations, consequently, assists identifying possible hybridization in the past (Sutarno, 2010). For many years, morphological traits are important in horse breeding as they are linked with specific performance and longevity (Koenen, Van Veldhuizen, & Brascamp, 1995; Saastamoinen & Barrey, 2000). Accordingly, genetic marker-assisted (MAS) program development would improve progress in economic traits.

3.0 MATERIALS AND METHOD

3.1 Samples Collection and Storage

Blood samples were obtained from a total number of 16 healthy Sumbawa ponies reared at Equestrian Unit, DBKL, Titiwangsa, Kuala Lumpur, Federal Territory of Kuala Lumpur. An amount of 6 ml of venous blood was collected aseptically as possible from the jugular vein of each pony into two EDTA tubes, size 3 ml. The blood tubes were labelled and maintained in an ice-block-filled cool box throughout transportation until arrived at the Virology Laboratory. The EDTA tubes were centrifuged at 3000 rpm for 5 minutes and buffy coats were collected by using P1000 micropipette to remove plasma then carefully aspirating the middle layer at about 200 µL. The buffy coats were transferred into respective 1.5 ml microcentrifuge tubes and reserved at -20 °C prior to further processing. Genomic DNA extraction is done using DNeasy® Blood & Tissue Kit (QIAGEN Biotechnology Malaysia Sdn Bhd) and the extracted samples are subjected to amplification for Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene study via polymerase chain reaction (PCR) and DNA sequencing analysis method.

3.2 Genomic DNA Extraction

The genomic DNA was isolated and obtained from nucleated blood specifically the buffy coat using DNeasy® Blood & Tissue Kit (QIAGEN Biotechnology Malaysia Sdn Bhd) according to the manufacturer's protocol with minor optimization. Lysing step is done by adding 20 µL Proteinase K into 200 µL anticoagulant-treated blood to achieve the volume 220 µL. Samples that were insufficient were adjusted with phosphate buffered saline (PBS). The mixture is then

added with 200 µL Buffer AL that is then mixed thoroughly by vortexing and incubated at 56 °C for 10 minutes in a digital dry bath incubator (heating block). After incubation, the solution was then added with 200 µL ethanol (96%), again mixed thoroughly by vortexing. The mixture was pipetted into a DNeasy mini spin columns placed in a 2 ml collection tubes followed by centrifugation at 8000 rpm for 1 minute. Flow-through and collection tubes were discarded. The spin columns were transferred into new collection tubes then added with 500 µL Buffer AW1, again centrifuged at 8000 rpm for 1 minute. Flow-through and collection tubes were discarded. The spin columns were transferred into new collection tubes then added with 500 µL Buffer AW2 followed by centrifugation at 13000 rpm for 3 minutes. Flow-through and collection tubes were discarded. The spin column is lastly transferred into 1.5 ml microcentrifuge tubes. Elution is done by adding 200 µL Buffer AE to the center of the spin column membrane, incubated at room temperature (15 – 25 °C) for 1 minute and centrifuged at 8000 rpm for 1 minute. DNA quality check is conducted quantitatively by measuring DNA concentration and purity using Eppendorf BioPhotometer plus (Eppendorf AG, Germany).

3.3 Polymerase Chain Reaction (PCR) and DNA Sequencing Analysis

Amplification of the Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene fragments was done via polymerase chain reaction (PCR) method. The primer pair of study (Table 1) would yield amplicons with target length 347 bp of the LCORL gene using both forward (5’- GCC ATC TAT TTG CAT GTT CTT G -3’) and reverse (5’- GGC AAG TTC ATA GGC TGG TTC -3’) primer sequences as stated by Metzger et al. (2013) and Chandra (2019).

Gene	Primer pairs (5' to 3')	Target length	References
LCORL	**F:** GCC ATC TAT TTG CAT GTT CTT G **R:** GGC AAG TTC ATA GGC TGG TTC	347 bp	Metzger et al. (2013) and Chandra (2019)

Table 1: The sequences and information for primer pair of study.

The PCR protocol is conducted by using MyTaq™ Red Mix (Bioline Reagents Ltd, UK) in a total reaction mixture of 25 µL (Table 2) according to the manufacturer's instructions with optimization. It consists of 12.5 µL MyTaq Red Mix, 2x (contains Red dye, 200 units MyTaq DNA Polymerase with a novel buffer system comprised of dNTPs, $MgCl_2$, and enhancers), 3 µL template genomic DNA, 1 µL of each forward and reverse primers (20 µM), and 7.5 µL double-distilled water (ddH_2O).

Reagent	Volume (µL)
Double-distilled water (ddH_2O)	7.5
MyTaq Red Mix	12.5
Forward primer	1.0
Reverse primer	1.0
Template DNA	3.0
Total volume	25

Table 2: PCR reaction mixture with reagents used and their respective volumes.

PCR conditions (Table 3) are programmed into the LabCycler Gradient by SensoQuest Biomedizinische Elektronik (Goettingen, Germany). The PCR products of this gene were sent to 1st BASE DNA Sequencing Division for purifying and

sequencing. The CLC Main Workbench 20.0.4 software was used for the comparison and alignment analysis.

Gene	PCR conditions			
	Initial denaturation	**Denaturation, Annealing, Extension**	**Final extension**	**Holding temperature**
LCORL	95 °C for 3 minutes 1 cycle	95 °C for 15 seconds, 60 °C for 15 seconds, 72 °C for 10 seconds 35 cycles	72 °C for 5 minutes 1 cycle	15 °C ∞ 1 cycle

Table 3: Program applied on the LabCycler Gradient by SensoQuest Biomedizinische Elektronik (Goettingen, Germany) for gene amplification.

3.4 Gel Electrophoresis and Documentation

The PCR products were electrophoresed on 2% agarose gel (prepared from 2 g of agarose powder, 100 ml 1X TAE buffer and 5 µL RedSafe) at constant setting of 90 V and 350 mA for 35 minutes. The results were visualized under the UV-transilluminator GENE GENIUS Bioimaging system (Syngene) and viewed using GeneSnap software (SynGene) for any observable banding particularly between 300 – 400 bp.

4.0 RESULTS

4.1 Amplification of LCORL Gene using Primer Pair of Study

The PCR products of Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene with the target fragment size of 347 bp, were successfully amplified with primer sequences (Table 1) as stated by Metzger et al. (2013) and Chandra (2019). The product was electrophoresed on 2% agarose gel (containing RedSafe nucleic acid staining solution) at constant setting of 90 V and 350 mA for 35 minutes. Lane M is indicated as the 100 bp DNA ladder (100 – 1,000 bp plus an additional band at 1,500 bp). All amplicons were visualized under UV-transilluminator light and viewed using GeneSnap software (SynGene) as single bands of expected size by which referring to the DNA ladder were between 300 – 400 bp (Figure 1 and 2).

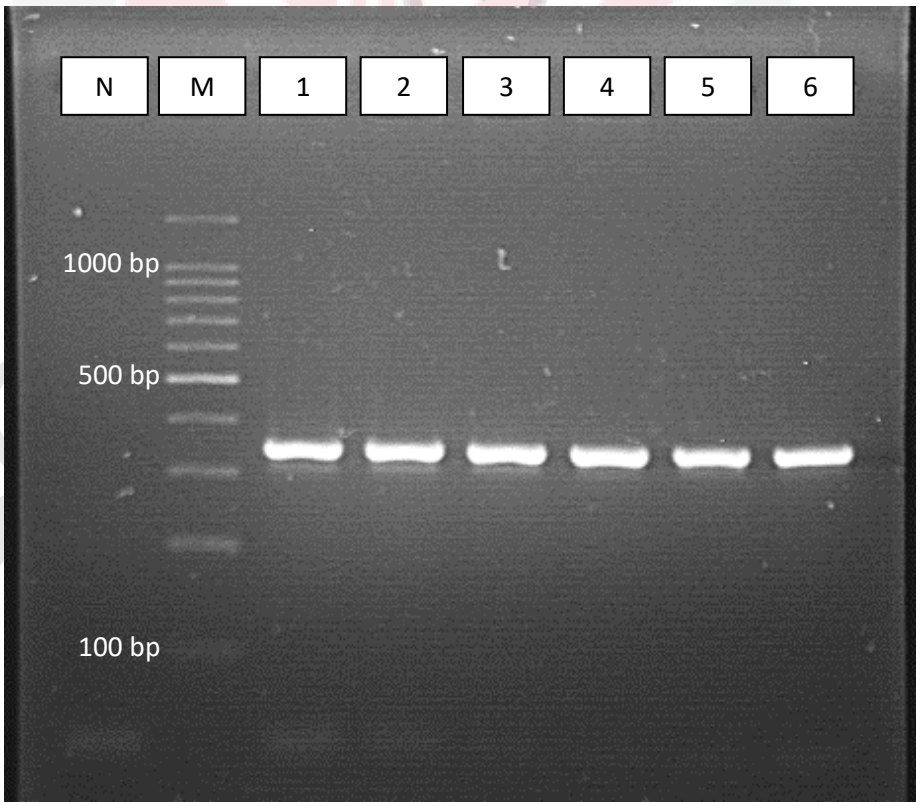

Figure 1: Visualization of PCR products for LCORL gene with target length of 347 bp, amplified using primer pair of study. Lane N is negative control, Lane M is DNA ladder (100 – 1500 bp), and Lane 3–8 are amplicons for each DNA samples no. 1 – 6 (first batch).

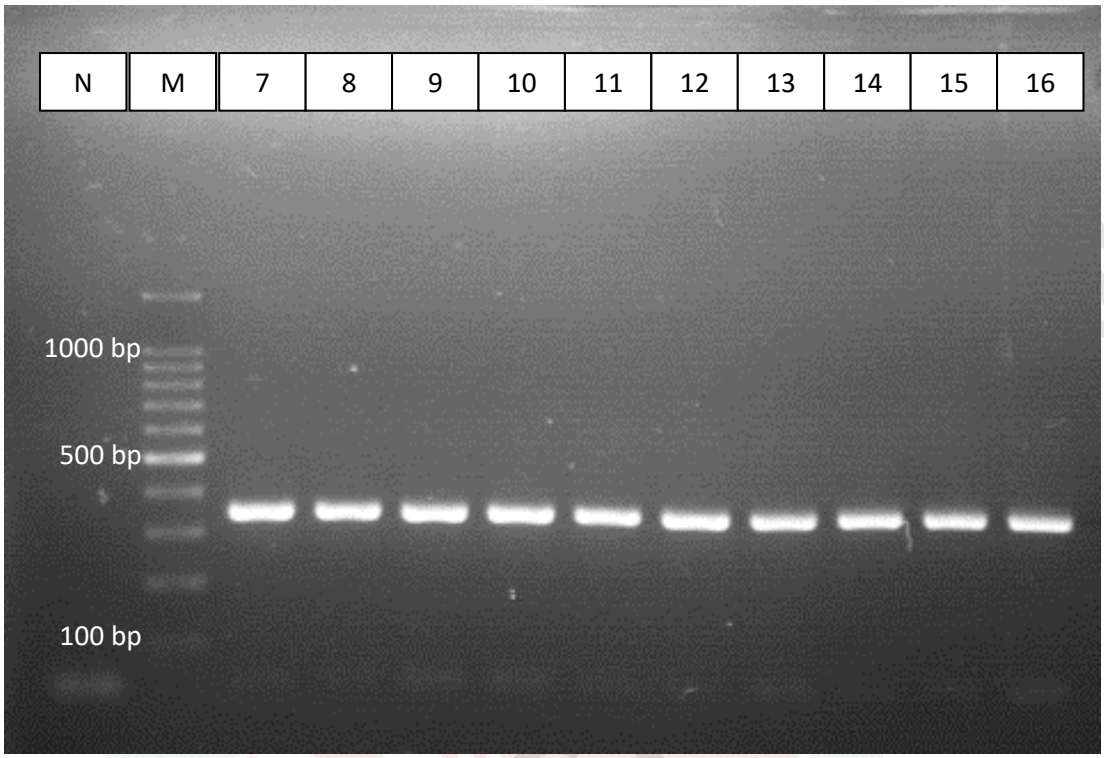

Figure 2: Visualization of PCR products for LCORL gene with target length of 347 bp, amplified using primer pair of study. Lane N is negative control, Lane M is DNA ladder (100 – 1500 bp), and Lane 3–12 are amplicons for each DNA samples no. 7 – 16 (second batch).

4.2 DNA Sequencing Analysis of PCR Products of LCORL Gene

The PCR products of each Sumbawa pony had been purified, sequenced and analysed revealing amplicons were approximately 320 bp using the primer pair of study. The chromatogram data showed that the BIEC2-808543 (Broad institute nomenclature; EquCab3_107,374,136 T>C) single nucleotide polymorphism (SNP) locus comprised of two genotypes for LCORL gene. This included the T allele and C allele showing twelve (12) ponies were homozygous TT and the other four (4) heterozygous TC (Figure 3).

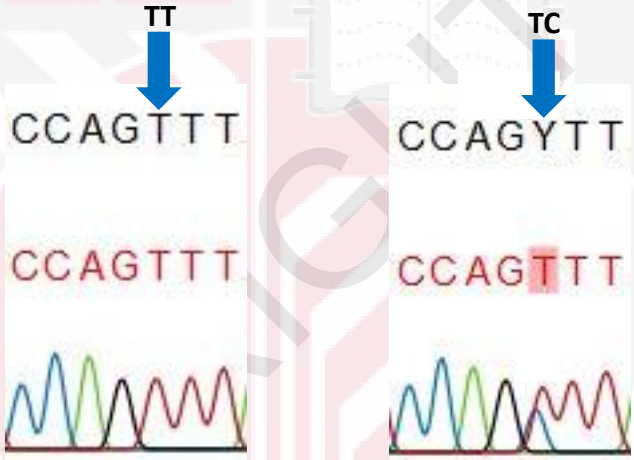

Figure 3: Genotyping of the BIEC2-808543 SNP in Sumbawa ponies based on chromatogram data adjacent to the LCORL gene.

5.0 DISCUSSION

5.1 Determination and Amplification of LCORL Gene Fragments

In both batches for lane 3 to 8 and 3 to 12, similar bands were successfully yielded on the 2% agarose gel and visualized under UV-transilluminator light. All fragment length produced were distinctively located between 300- and 400-bp. With the intention of confirming the DNA of Ligand Dependent Nuclear Receptor Corepressor Like (LCORL) gene fragment size, a semi-log graph was used, also, additionally analysed with GelAnalyzer software. Each of the LCORL gene amplicon size was determined by plotting the x-axis (inverse distance migrated) and y-axis (inverted log base pair) in Microsoft Excel resulting on average 320 bp when referred to the DNA ladder on Lane M, taking into consideration that it was also sent for sequencing. Further clarification for these fragment lengths of amplified products were by matching the alignment of the primer pair with LCORL gene sequence. Previous researches of LCORL gene polymorphisms in domestic horses (*Equus caballus*) conducted by Metzger et al. (2013) and Chandra (2019) had demonstrated the target length of 347 bp. This result is dissimilar from other publications by He et al. (2015) and Bai et al. (2020) which were 284 bp and 507 bp respectively. This is due to the fact that different designs of primers were used yet they amplified the same locus of interest containing BIEC2-808543 single nucleotide polymorphism (SNP).

Currently there is unlikelihood for other studies to show LCORL gene base pair value of 320 bp as in this study. Although there is deviation for the fragment size produced but primer pairs matched to its original study (as stated by Metzger et al. (2013), F = 5’- GCC ATC TAT TTG CAT GTT CTT G -3’ and R = 5’- GGC AAG TTC ATA GGC TGG TTC -3’) and frequently used in determination of DNA

fragment length for LCORL gene sequence (Okuda et al., 2016; Chandra, 2019). The primer pair was specially designed using Primer3 (http://frodo.wi.mit.edu/primer3/) and for over a decade has been widely used for formulating PCR primers (Koressaar & Remm, 2007). Therefore, it is guaranteed for this study that the amplicons produced on 2% agarose gel electrophoresis were LCORL gene with 320 bp fragment size.

5.2 PCR and DNA Sequencing Analysis of LCORL Gene Polymorphism

In this study, LCORL gene polymorphism for these Sumbawa ponies were identified via PCR and DNA sequencing analysis method. This technique was made possible with the usage of CLC Main Workbench 20.0.4 software for the comparison and alignment analysis of sequence data. Based on the DNA sequences that were analysed (Figure 3) from the amplified segments of LCORL gene showed the BIEC2-808543 single nucleotide polymorphism (SNP) within individuals that contain the Y ambiguity with pyrimidine either thymine (T) or cytosine (C), whereby this finding is supported by He et al. (2015) and Bai et al. (2020) in their study on horse breeds available and native to their geographical location consequently upon visualization of the chromatogram and alignment analysis. According to the Ensembl genome database project for that particular target SNP is at position 107,374,136 (www.ensembl.org, Genome assembly: EquCab 3.0) by which the locus is comprised of three genotypes for LCORL gene. This included the T allele and C allele producing homozygous TT, heterozygous TC and homozygous CC (Chandra, 2019). The C allele is responsible for the reduced expression level of LCORL for body size (within and across horse breeds) leading to an increased wither height, whereas the T allele causes the normal or alternative form of expression at the same position (Metzger et al., 2013).

Determined from the chromatograms, a single peak position within a trace showed double peaks of different colours instead of just one by the base caller indicating heterozygous (SNP) peaks, while the latter with evenly-spaced peaks and the lack of baseline noise are homozygous (single) peaks (Stucky, 2012). Therefore, all Sumbawa ponies investigated in this study has presence of polymorphism and were genotyped as twelve (12) ponies being homozygous TT and the other four (4) heterozygous TC. Mostafavi et al. (2019) stipulated that in certain horse breeds the C allele is rare, because of selection pressure for particular body composition hence the absence of CC genotype Sumbawa ponies. The distribution of LCORL genotypes and alleles demonstrated genotype TT was close to complete association for all pony breeds limited up to 148 cm for wither height in contrast to the larger and heavier horses, for the most part showed the genotype CC (Metzger et al., 2013; Okuda et al., 2016). The understanding of this polymorphism adjacent to the LCORL gene together with its action towards affecting growth or morphological traits remains unclear.

The initial functional study carried out by Metzger et al. (2013) verified via GWAS the highest associated quantitative trait locus (QTL) for wither height on chromosome 3 (ECA 3) proximal to the LCORL gene along with the confirmed SNP of BIEC2-808543 (Broad institute nomenclature; EquCab3_107,374,136 T>C) whereby there is substitution of thymine (T) to cytosine (C) which is highly deemed as the potential candidate for body size. This polymorphism located within the putative DNA consensus sequence element which is the TATA box element (Figure 4) disrupts the binding site for transcription factor IID (TFIID) complex for pre-initiation of transcription apparatus by RNA polymerase II and causing lack of recognition by the core promoter elements (Tatarakis et al., 2008; Sawadogo & Roeder, 1985; Orphanides, Lagrange, & Reinberg, 1996). TFIID has been regarded to play a role in affecting

genes associated with skeletal development (Yang et al., 2011; Nakatani et al., 1990). This leads to the effect on indicating TFIID by a pro-motor nuclear element for the first step in mRNA transcription influences the bone-like AP-1, activator protein-1 transcription factor complex, and AP-1 is activated as a complex of mRNA thus subsequently transcription is up-regulated resulting to expansion of skeletal bones from chondrocytes, osteoblasts and osteoclasts cells (Abrishami-Moghaddam et al., 2020). The SNP T>C mutated transcription factor binding site implicates significant alterations of the LCORL expression levels whereby it is reduced and caused an increased wither height. This first functional study then became a foundational framework for the following genetic researches for this gene and its association towards morphological traits (He et al., 2015; Okuda et al., 2016; Kim et al., 2018; Chandra, 2019; Bai et al., 2020).

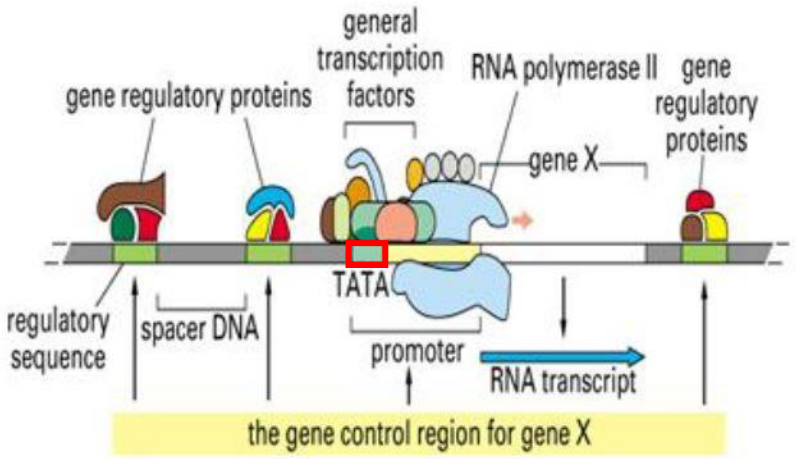

Figure 4: Controlling regions proximal to the LCORL gene in relation to gene expression effect. Adapted from: "Effect of BIEC2-808543 near LCORL on body size of Iranian Arab horse" by Abrishami-Moghaddam, Z., Miresmaeli, S. M., Sani, M. B., and Seifati, S. M., Spring 2020, *Iranian Journal of Animal Science Research*, *12*(1), p. 125 – 133. DOI: 10.22067/ijasr.v12i1.78988

6.0 CONCLUSION

To conclude, gene polymorphism was evident in the Sumbawa ponies investigated within stables of the selected Equestrian unit with regards to LCORL gene. They were identified as twelve (12) ponies with genotypes being homozygous TT and the other four (4) heterozygous TC. Gene diversity of LCORL is relatively low, yet polymorphism was still evident. This could be due to the limited number of Sumbawa ponies, absence of breeding program for that breed in the stables and possibly the high inbreeding frequency of those parent ponies from their original source of import. Therefore, the LCORL gene potential as candidate for genetic marker to be used in breeding program in these Sumbawa ponies can be considered especially when a higher gene diversity is determined.

7.0 RECOMMENDATIONS

For further studies, it is recommended to use a larger sample size and from different locations and stables to increase gene diversity therefore a more reliable and accurate result in terms of identification of all three genotypes pursued in this study could be obtained. Restriction fragment length polymorphism (RFLP) method should be considered due to its cost efficiency and relatively simpler conductivity. An extension of this study is perhaps to perform association analysis between polymorphism and body measurements.

REFERENCES

Abrishami-Moghaddam, Z., Miresmaeli, S. M., Sani, M. B., & Seifati, S. M. (2020). Effect of BIEC2-808543 near LCORL on body size of Iranian Arab horse. *Iranian Journal of Animal Science Research, 12*(1), 125-133. DOI: 10.22067/ijasr.v12i1.78988

Aerts, J., Wetzels, Y., Cohen, N., & Aerssens, J. (2002). Data mining of public SNP databases for the selection of intragenic SNPs. *Human Mutation, 20*(3), 162-173. DOI: 10.1002/humu.10107

Bai, H., Lu, H., Wang, L., Wang, S., Zeng, W., & Zhang, T. (2020). SNPs analysis of height traits in Ningqiang pony. *Animal Biotechnology*, 1-7. DOI: 10.1080/10495398.2020.1728288

Bailey, E., & Binns, M. M. (1998). The horse gene map. *ILAR Journal*, *39*(2-3), 171-176. DOI: 10.1093/ilar.39.2-3.171

Bartlett, J. M., & Stirling, D. (Eds.). (2003). *PCR Protocols* (Vol. 226). Totowa, NJ: Humana Press.

Binns, M. M., Boehler, D. A., & Lambert, D. H. (2010). Identification of the myostatin locus (MSTN) as having a major effect on optimum racing distance in the Thoroughbred horse in the USA. *Animal Genetics*, *41*, 154-158. DOI: 10.1111/j.1365-2052.2010.02126.x

Boyko, A. R., Quignon, P., Li, L., Schoenebeck, J. J., Degenhardt, J. D., Lohmueller, K. E., ... Ostrander, E. A. (2010). A simple genetic architecture underlies morphological variation in dogs. *PLoS Biology*, *8*(8), e1000451. DOI: 10.1371/journal.pbio.1000451

Chandra, P. R. (2019). *Study on Genetic Diversity and Characteristics of Japanese Native Horse Populations* [Doctoral dissertation, Okayama University]. Okayama University Scientific Achievement Repository. Retrieved from http://ousar.lib.okayama-u.ac.jp/files/public/5/57439/20191208131956316185/K0006072_fulltext.pdf

Darwin, C. R. (1859). Chapter V. Laws of Variation. *The Origin of Species by Means of Natural Selection: or, The Preservation of Favoured Races in the Struggle for Life* (pp. 131-170). London, UK: John Murray.

Department of Veterinary Services (2019). *Overview of the equine industry in Malaysia – National Equine Forum 2019*. Retrieved from http://research.dvs.gov.my/opac/author/upload/Overview%20of%20the%20equine%20industry%20in%20Malaysia.pdf

Edwards, E. H., Langrish, B., & Houghton, K. (2000). *The New Encyclopaedia of the Horse*. London, UK: Dorling Kindersley.

Farkas, D. H., & Holland, C. A. (2009). Overview of molecular diagnostic techniques and instrumentation. *In Cell and Tissue based Molecular Pathology* (pp. 19-32). Churchill Livingstone.

Filatov, D. A. (2002). ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. *Molecular Ecology Notes*, *2*(4), 621–624. DOI: 10.1046/j.1471-8286.2002.00313.x

Fédération Équestre Internationale (2007). *Pony Measurement Rules: Updated Article 3111- Passports and certificates of pony measurement*. Retrieved from http://www.eurodressage.com/2007/01/10/new-pony-measurement-rules-announced-fei-website

Gleeson, T., & Hillier, L. (1991). A trace display and editing program for data from fluorescence-based sequencing machines. *Nucleic Acids Research*, *19*(23), 6481-6483. DOI: 10.1093/nar/19.23.6481

Godard, S., Vaiman, D., Oustry, A., Nocart, M., Bertaud, M., Guzylack, S., ... Guérin, G. (1997). Characterization, genetic and physical mapping analysis of 36 horse plasmid and cosmid-derived microsatellites. *Mammalian Genome*, *8*(10), 745-750. DOI: 10.1007/s003359900558

Hartl, D. L., & Clark, A. G. (2007). *Principles of Population Genetics* (4th ed.). Sunderland, Mass: Sinauer Associates.

He, S., Zhang, L., Li, W., & Liu, M. (2015). BIEC2-808543 SNP in the LCORL gene is associated with body conformation in the Yili horse. *Animal Biotechnology*, *26*(4), 289-291. DOI: 10.1080/10495398.2014.995303

Hill, E. W., McGivney, B. A., Gu, J., Whiston, R., & MacHugh, D. E. (2010). A genome-wide SNP-association study confirms a sequence variant (g. 66493737 C>T) in the equine myostatin (MSTN) gene as the most powerful predictor of optimum racing distance for Thoroughbred racehorses. *BMC Genomics*, *11*(1), 552. DOI: 10.1186/1471-2164-11-552

Ishmael, F. T., & Stellato, C. (2008). Principles and applications of polymerase chain reaction: basic science for the practicing physician. *Annals of Allergy, Asthma & Immunology*, *101*(4), 437-443. DOI: 10.1016/S1081-1206(10)60323-7

Ismail, S., & Essawi, M. (2012). Genetic polymorphism studies in humans. *Middle East Journal of Medical Genetics*, *1*(2), 57–63. DOI: 10.1097/01.mxe.0000415225.85003.47

Jahromi, M., Ahmed, A., Behbehani, K., & Mohammad, A. (2014). Critical association study of olfactory receptor Gene polymorphism in diabetic complications. *Immunome Research*, *10*(2), 1. DOI: 10.4172/1745-7580.1000079

Joshi, M., & Deshpande, J. D. (2010). Polymerase chain reaction: methods, principles and application. *International Journal of Biomedical Research*, *2*(1). DOI: 10.7439/ijbr.v2i1.83

Karki, R., Pandya, D., Elston, R. C., & Ferlini, C. (2015). Defining "mutation" and "polymorphism" in the era of personal genomics. *BMC Medical Genomics*, *8*(1), 37. DOI: 10.1186/s12920-015-0115-z

Kim, N. Y., Yang, Y. H., Park, N. G., Yang, B. C., Son, J. K., Shin, S. M., ... Park, H. B. (2018). Association of SNP markers on chromosomes 3 and 9 with body weight in Jeju horses. *Journal of Life Science*, *28*(7), 795-801. DOI: 10.5352/JLS.2018.28.7.795

Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I., & Khorana, H. G. (1971). Studies on polynucleotides: XCVI. Repair replication of short synthetic DNA's as

catalyzed by DNA polymerases. *Journal of Molecular Biology*, *56*(2), 341-361. DOI: 10.1016/0022-2836(71)90469-4

Koenen, E. P. C., Van Veldhuizen, A. E., & Brascamp, E. W. (1995). Genetic parameters of linear scored conformation traits and their relation to dressage and show-jumping performance in the Dutch Warmblood Riding Horse population. *Livestock Production Science*, *43*(1), 85-94. DOI: 10.1016/0301-6226(95)00010-I

Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, *23*(10), 1289-1291. DOI: 10.1093/bioinformatics/btm091

La, Y., Zhang, X., Li, F., Zhang, D., Li, C., Mo, F., & Wang, W. (2019). Molecular characterization and expression of SPP1, LAP3 and LCORL and their association with growth traits in sheep. *Genes*, *10*(8), 616. DOI: 10.3390/genes10080616

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., ... Ferreira, T. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, *467*(7317), 832-838. DOI: 10.1038/nature09410

Libioulle, C., Louis, E., Hansoul, S., Sandor, C., Farnir, F., Franchimont, D., ... Georges, M. (2007). Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genetics*, *3*(4), e58. DOI: 10.1371/journal.pgen.0030058

Makvandi-Nejad, S., Hoffman, G. E., Allen, J. J., Chu, E., Gu, E., Chandler, A. M., ... Sutter, N. B. (2012). Four loci explain 83% of size variation in the horse. *PLoS One*, *7*(7), e39929. DOI: 10.1371/journal.pone.0039929

Maurya, R., Singh, R. K., Kumar, B., Salotra, P., Rai, M., & Sundar, S. (2005). Evaluation of PCR for diagnosis of Indian kala-azar and assessment of cure. *Journal of Clinical Microbiology*, *43*(7), 3038-3041. DOI: 10.1128/JCM.43.7.3038-3041.2005

Metzger, J., Schrimpf, R., Philipp, U., & Distl, O. (2013). Expression levels of LCORL are associated with body size in horses. *PloS One*, *8*(2), e56497. DOI: 10.1371/journal.pone.0056497

Metzker, M. L. (2005). Emerging technologies in DNA sequencing. *Genome Research*, *15*(12), 1767-1776. DOI: 10.1101/gr.3770505

Mostafavi, A., Fozi, M. A., Koshkooieh, A. E., Mohammadabadi, M., Babenko, O. I., & Klopenko, N. I. (2020). Effect of LCORL gene polymorphism on body size traits in horse populations. *Acta Scientiarum. Animal Sciences*, *42*. DOI: 10.4025/actascianimsci.v42i1.47483

Nakatani, Y., Horikoshi, M., Brenner, M., Yamamoto, T., Besnard, F., Roeder, R. G., & Freese, E. (1990). A downstream initiation element required for efficient TATA box binding and in vitro function of TFIID. *Nature*, *348*(6296), 86-88. DOI: 10.1038/348086a0

Noor, R. R. (2008). *Genetika Ternak* (Edisi ke-4). Jakarta, Indonesia: Penebar Swadaya.

Okuda, Y., Moe, H. H., Moe, K. K., Shimizu, Y., Nishioka, K., Shimogiri, T., ... Kunieda, T. (2016). Genotype distribution and allele frequencies of the genes associated with body composition and locomotion traits in Myanmar native horses. *Animal Science Journal*, *88*(8), 1198-1203. DOI: 10.1111/asj.12756

Orphanides, G., Lagrange, T., & Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development*, *10*(21), 2657-2683. DOI: 10.1101/gad.10.21.2657

Othman, O. E., Abdel-Samad, M. F., El Maaty, N. A. A., & Sewify, K. M. (2012). Evaluation of DNA polymorphism in Egyptian buffalo growth hormone and its receptor genes. *Journal of Applied Biological Sciences*, *6*(2), 37-42. Retrieved from http://jabsonline.org/index.php/jabs/article/view/282

Ozturk, R. C., & Altinok, I. (2017). Analysis of microsatellite data for population genetics of aquatic organisms. *Genetics of Aquatic Organisms*, *1*(2), 61-69. DOI: 10.4194/2459-1831-v1_2_04

Paxson, J. (2008). Polymerase chain reaction test interpretation. *Education for Veterinarians*, *186*, 196. Retrieved from https://pdfs.semanticscholar.org/4620/ea57ba2008f8ff028db767c149145172327c.pdf

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., ... Baumeister, K. (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, *238*(4825), 336-341. DOI: 10.1126/science.2443975

Raudsepp, T., Finno, C. J., Bellone, R. R., & Petersen, J. L. (2019). Ten years of the horse reference genome: insights into equine biology, domestication and population dynamics in the post-genome era. *Animal Genetics*, *50*(6), 569-597. DOI: 10.1111/age.12857

Rozas, J., & Rozas, R. (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics (Oxford, England)*, *15*(2), 174-175. DOI: 10.1093/bioinformatics/15.2.174

Rubin, C. J., Megens, H. J., Barrio, A. M., Maqbool, K., Sayyab, S., Schwochow, D., ... Archibald, A. L. (2012). Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences*, *109*(48), 19529-19536. DOI: 10.1073/pnas.1217149109

Saastamoinen, M. T., & Barrey, E. (2000). Genetics of conformation, locomotion and physiological traits. In *The Genetics of the Horse* (pp. 411-438). CABI, Oxford.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *74*(12), 5463-5467. DOI: 10.1073/pnas.74.12.5463

Sawadogo, M., & Roeder, R. G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell*, *43*(1), 165-175. DOI: 10.1016/0092-8674(85)90021-2

Schneider, S., Roessli, D., & Excoffier, L. (2000). Arlequin ver. 2.000. *A Software for Population Genetics Data Analysis.* Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schröder, W. (2010). *Athletic performance and conformation in Hanoverian warmblood horses-population genetic and genome-wide association analyses* [cumulative thesis]. Hannover: University of Veterinary Medicine. Retrieved from https://elib.tiho-hannover.de/servlets/MCRFileNodeServlet/etd_derivate_00001149/schroederw_ws10.pdf

Setoguchi, K., Watanabe, T., Weikard, R., Albrecht, E., Kühn, C., Kinoshita, A., ... Takasuga, A. (2011). The SNP c. 1326T> G in the non-SMC condensin I complex, subunit G (NCAPG) gene encoding a p. Ile442Met variant is associated with an increase in body frame size at puberty in cattle. *Animal Genetics*, *42*(6), 650-655. DOI: 10.1111/j.1365-2052.2011.02196.x

Signer-Hasler, H., Flury, C., Haase, B., Burger, D., Simianer, H., Leeb, T., & Rieder, S. (2012). A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PloS One*, *7*(5), e37282. DOI: 10.1371/journal.pone.0037282

Singh, R. S. (2001). Polymorphism. *Encyclopaedia of Genetics*, 1507–1509. DOI: 10.1006/rwgn.2001.1012

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., ... Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*, *321*(6071), 674-679. DOI: 10.1038/321674a0

Staiger, E. A., Al Abri, M. A., Pflug, K. M., Kalla, S. E., Ainsworth, D. M., Miller, D., ... Brooks, S. A. (2016). Skeletal variation in Tennessee Walking Horses maps to the LCORL/NCAPG gene region. *Physiological Genomics*, *48*(5), 325-335. DOI: 10.1152/physiolgenomics.00100.2015

Stucky, B. J. (2012). SeqTrace: A graphical tool for rapidly processing DNA sequencing chromatograms. *Journal of Biomolecular Techniques: JBT*, *23*(3), 90–93. DOI: 10.7171/jbt.12-2303-004

Sutarno, S. (2010). Genetic variations among Indonesian native cattle breeds based on polymorphisms analysis in the growth hormone loci and mitochondrial DNA. *Biodiversitas Journal of Biological Diversity*, *11*(1). DOI: 10.13057/biodiv/d110101

Tajima, F., Misawa, K., & Innan, H. (1998). The amount and pattern of DNA polymorphism under the neutral mutation hypothesis. In *Mutation and Evolution* (pp. 103-107). Springer, Dordrecht. DOI: 10.1023/A:1017011631572

Tatarakis, A., Margaritis, T., Martinez-Jimenez, C. P., Kouskouti, A., Mohan II, W. S., Haroniti, A., ... Talianidis, I. (2008). Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. *Molecular Cell*, *31*(4), 531-543. DOI: 10.1016/j.molcel.2008.07.013

Tetens, J., Widmann, P., Kühn, C., & Thaller, G. (2013). A genome-wide association study indicates LCORL/NCAPG as a candidate locus for withers height in

German Warmblood horses. *Animal Genetics*, *44*(4), 467-471. DOI: 10.1111/age.12031

Tozaki, T., Hill, E. W., Hirota, K., Kakoi, H., Gawahara, H., Miyake, T., ... Kurosawa, M. (2012). A cohort study of racing performance in Japanese Thoroughbred racehorses using genome information on ECA18. *Animal Genetics*, *43*(1), 42-52. DOI: 10.1111/j.1365-2052.2011.02201.x

Tozaki, T., Sato, F., Ishimaru, M., Kikuchi, M., Kakoi, H., Hirota, K. I., & Nagata, S. I. (2016). Sequence variants of BIEC2-808543 near LCORL are associated with body composition in Thoroughbreds under training. *Journal of Equine Science*, *27*(3), 107-114. DOI: 10.1294/jes.27.107

Turk, A. (2017). A scientific and historical investigation on Mongolian horses. *История: факты и символы*, *2*(11). DOI: 10.24888/2410-4205-2017-11-2-23-32

van de Pol, C., & van Oldruitenborgh-Oosterbaan, M. S. (2007). Measuring the height of ponies at the withers: Influence of time of day, water and feed withdrawal, weight-carrying, exercise and sedation. *The Veterinary Journal*, *174*(1), 69-76. DOI: 10.1016/j.tvjl.2006.10.023

Vial, C., & Evans, R. (Eds.). (2015). *The New Equine Economy in the 21st Century* (Vol. 136). Wageningen, Netherlands: Wageningen Academic Publishers.

Wade, C. M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., ... Blöcker, H. (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, *326*(5954), 865-867. DOI: 10.1126/science.1178158

Wellinghausen, N., Kochem, A. J., Disqué, C., Mühl, H., Gebert, S., Winter, J., ... Sakka, S. G. (2009). Diagnosis of bacteremia in whole-blood samples by use of a commercial universal 16S rRNA gene-based PCR and sequence analysis. *Journal of Clinical Microbiology*, *47*(9), 2759-2765. DOI: 10.1128/JCM.00567-09

Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H. A., Arnheim, N., & Wilson, A. C. (1987). Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucleic Acids Research*, *15*(2), 529-541. DOI: 10.1093/nar/15.2.529

Yang, M. Q., Laflamme, K., Gotea, V., Joiner, C. H., Seidel, N. E., Wong, C., ... Gallagher, P. G. (2011). Genome-wide detection of a TFIID localization element from an initial human disease mutation. *Nucleic Acids Research*, *39*(6), 2175-2187. DOI: 10.1093/nar/gkq1035